



Characterization of transcript enrichment and detection bias in single-nucleus RNA-seq for mapping of distinct human adipocyte lineages

Anushka Gupta, Farnaz Shamsi, Nicolas Altemose, et al.

Genome Res. 2022 32: 242-257 originally published online January 18, 2022
Access the most recent version at doi:[10.1101/gr.275509.121](https://doi.org/10.1101/gr.275509.121)

References This article cites 117 articles, 16 of which can be accessed free at:
<http://genome.cshlp.org/content/32/2/242.full.html#ref-list-1>

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Characterization of transcript enrichment and detection bias in single-nucleus RNA-seq for mapping of distinct human adipocyte lineages

Anushka Gupta,¹ Farnaz Shamsi,² Nicolas Altemose,¹ Gabriel F. Dorlhiac,³ Aaron M. Cypess,⁴ Andrew P. White,⁵ Nir Yosef,^{6,7,8,9} Mary Elizabeth Patti,¹⁰ Yu-Hua Tseng,^{2,11} and Aaron Streets^{1,3,8}

¹University of California at Berkeley–University of California at San Francisco Graduate Program in Bioengineering, Berkeley, California 94720, USA; ²Section on Integrative Physiology and Metabolism, Joslin Diabetes Center, Harvard Medical School, Boston, Massachusetts 02115, USA; ³Biophysics Graduate Group, University of California at Berkeley, Berkeley, California 94720, USA; ⁴Diabetes, Endocrinology, and Obesity Branch, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health, Bethesda, Maryland 20892, USA; ⁵Department of Orthopedic Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02115, USA; ⁶Center for Computational Biology, University of California, Berkeley, Berkeley, California 94720, USA; ⁷Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, California 94720, USA; ⁸Chan Zuckerberg Biohub, San Francisco, California 94158, USA; ⁹Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Cambridge, Massachusetts 02139, USA; ¹⁰Research Division, Joslin Diabetes Center, Boston, Massachusetts 02115, USA; ¹¹Harvard Stem Cell Institute, Harvard University, Cambridge, Massachusetts 02138, USA

Single-cell RNA sequencing (scRNA-seq) enables molecular characterization of complex biological tissues at high resolution. The requirement of single-cell extraction, however, makes it challenging for profiling tissues such as adipose tissue, for which collection of intact single adipocytes is complicated by their fragile nature. For such tissues, single-nucleus extraction is often much more efficient and therefore single-nucleus RNA sequencing (snRNA-seq) presents an alternative to scRNA-seq. However, nuclear transcripts represent only a fraction of the transcriptome in a single cell, with snRNA-seq marked with inherent transcript enrichment and detection biases. Therefore, snRNA-seq may be inadequate for mapping important transcriptional signatures in adipose tissue. In this study, we compare the transcriptomic landscape of single nuclei isolated from preadipocytes and mature adipocytes across human white and brown adipocyte lineages, with whole-cell transcriptome. We show that snRNA-seq is capable of identifying the broad cell types present in scRNA-seq at all states of adipogenesis. However, we also explore how and why the nuclear transcriptome is biased and limited, as well as how it can be advantageous. We robustly characterize the enrichment of nuclear-localized transcripts and adipogenic regulatory lncRNAs in snRNA-seq, while also providing a detailed understanding for the preferential detection of long genes upon using this technique. To remove such technical detection biases, we propose a normalization strategy for a more accurate comparison of nuclear and cellular data. Finally, we show successful integration of scRNA-seq and snRNA-seq data sets with existing bioinformatic tools. Overall, our results illustrate the applicability of snRNA-seq for the characterization of cellular diversity in the adipose tissue.

[Supplemental material is available for this article.]

Adipose tissue is a complex, heterogenous organ responsible for maintaining energy balance in animals by storing energy during nutritional excess and providing energy during nutritional deprivation. This regulation of whole-body energy homeostasis is primarily maintained by two functionally different types of fat: white adipose tissue (WAT), the primary site of lipid storage, and brown adipose tissue (BAT), which specializes in thermogenic energy expenditure. An imbalance in the metabolic activity or expansion of WAT and BAT is implicated in the pathogenesis of lipodystrophy or obesity and associated comorbidities like cardiovascular diseases and type 2 diabetes (Jo et al. 2009; Carobbio et al. 2011; Levelt et al. 2016). Further complexity arises from the het-

erogeneity within WAT, which also includes a cellular subtype called beige adipocytes with greater oxidative capacity (Pfeifer and Hoffmann 2015). Therefore, understanding the molecular pathways of adipose tissue expansion (adipogenesis) in humans and identifying resident cell types that regulate adipocyte activity are necessary for understanding the tissue's contribution in the pathology of such metabolic diseases. Consequently, recent studies are beginning to elucidate cellular heterogeneity and developmental pathways in distinct adipose tissue lineages at the single-cell level (Deutsch et al. 2020; Ferrero et al. 2020; Shamsi et al. 2021).

Corresponding author: astreet@berkeley.edu

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.275509.121>.

© 2022 Gupta et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Over the last decade, single-cell RNA sequencing (scRNA-seq), combined with microfluidic technologies, has proven to be a powerful tool for transcriptomic profiling of complex tissues in an unbiased fashion (Trapnell 2015; Birnbaum 2018; Chen et al. 2018). Indeed, multiple recent studies using microfluidic scRNA-seq approaches have investigated the heterogeneity of adipocyte precursors (referred to as preadipocytes in the text) in mice (Ferrero et al. 2020; Rondini and Granneman 2020). However, transcriptomic profiling of single mature adipocytes has been challenging (Shinoda et al. 2015; Xue et al. 2015; Min et al. 2016; Spaethling et al. 2016; Gao et al. 2017; Lee et al. 2019; Song et al. 2020), in part because of the technical barriers associated with isolating intact adipocytes, which are marked with fragile nature, high buoyancy, and large size (Deutsch et al. 2020). Consequently, single-nucleus RNA sequencing (snRNA-seq) is emerging as an alternative approach for transcriptomic profiling of cellular heterogeneity within adipose tissue (Rajbhandari et al. 2019; Sun et al. 2020; Sárvári et al. 2021). However, a single nucleus contains 10- to 100-fold less mRNA than whole cells, raising the question of whether the composition of mRNA transcripts in the nucleus is sufficient to enable identification of the same cell populations as whole cells. Indeed, previous comparisons of single-cell and single-nucleus approaches have shown that the relationship between nuclear and cytoplasmic mRNA is tissue specific (Habib et al. 2017; Lake et al. 2017; Bakken et al. 2018; Wu et al. 2019; Selewa et al. 2020; Thrupp et al. 2020). Therefore, there is a need to understand the transcriptomic similarities and differences between single-cell and single-nucleus profiles in the context of human adipose tissue, for which there is growing need to rely on snRNA-seq.

In this study, we explored the ability of snRNA-seq to recapitulate the transcriptional profiles observed by scRNA-seq in the human adipose tissue white and brown lineages. We focused our study on a well-controlled in vitro system of human white and brown adipogenesis (Fig. 1A; Xue et al. 2015; Kriszt et al. 2017). In this in vitro model, paired white and brown primary preadipocytes were isolated from a defined region (the neck depot) of a single individual. This system allowed us to measure cell-to-cell transcriptomic variations within and between lineages, while controlling for inter-individual variabilities that are typically associated with transcriptomic profiling of primary human adipose tissue, such as body mass index, genotype, and gender. Preadipocytes from both lineages were isolated while preserving their intrinsic cellular heterogeneity and were then immortalized to allow for long-term in vitro cell culture. Previously reported data showed that the preadipocyte populations could be differentiated into mature adipocytes with gene expression profiles that correspond to the adipogenic and thermogenic function of primary tissue from human neck BAT and WAT (Xue et al. 2015). Moreover, the in vitro cell-culture system allows for the isolation of intact nuclei as well as intact single cells across well-defined stages of adipogenesis, including mature, lipid-laden white and brown adipocytes. Overall, using this model system, we aimed to characterize the accuracy with which snRNA-seq can identify cell types and recover biology relevant to the adipose tissue, compared with scRNA-seq.

Results

scRNA-seq reveals transcriptional landscape of white and brown preadipocytes

Unsupervised clustering of the white and brown preadipocyte scRNA-seq library grouped the cells into three clusters, referred

to as populations 0, 1, and 2 (Fig. 1B). White preadipocytes organized into a single homogeneous cell population, cluster 0, whereas brown preadipocytes revealed two cell populations, cluster 1 and cluster 2 (Fig. 1B). As expected, clusters of white and brown preadipocytes were highly concordant with molecular features of respective primary preadipocytes (Supplemental Fig. S1C,D; Tews et al. 2014). All populations were devoid of endothelial (*CD31*) and hematopoietic marker genes (*PTPRC* [previously known as *CD45*]) (Supplemental Fig. S1E,F) and reflected a preadipocyte state on the basis of their high expression for the common mesenchymal stem cell markers *ITGB1* (also known as *CD29*), *THY1* (also known as *CD90*), *CD44*, and *ENG* (Supplemental Fig. S1G–J; Cawthorn et al. 2012; Mildmay-White and Khan 2017). All populations also had positive expression for adipogenesis regulators *CEBPB* & *PPARG* (Ghaben and Scherer 2019), as well as *ZEB1* (Gubelmann et al. 2014), further verifying an adipogenic fate for these cells (Supplemental Fig. S1K–M).

Differential gene expression (DGE) analysis confirmed that white preadipocytes showed enrichment of genes that are reported to be involved in establishing white preadipocytes' identity (Supplemental Table S1B) such as *TCF21* (de Jong et al. 2015), *PAX3* (Sanchez-Gurmaches et al. 2016), and *PDGFRA* (Berry and Rodeheffer 2013). The most up-regulated gene in white preadipocytes was *ID1* (Fig. 1C), which is known to maintain the progenitor state in preadipocytes by positively regulating the progression of cell cycle for sustained growth and proliferation (Satyanarayana et al. 2012; Patil et al. 2014). Consequently, enriched expression of *ID1* in white preadipocytes suggested ongoing signaling for the maintenance of cellular proliferation. In brown preadipocytes, the top up-regulated genes included *ANKRD1* and *CCN2* (Fig. 1C), which are well-characterized YAP target genes (Yu et al. 2012). YAP/TAZ are mechanosensitive transcriptional coactivators that regulate proliferation and differentiation at the precursor state (Dupont et al. 2011; Hansen et al. 2015; Zhang et al. 2018), while also maintaining thermogenic activity at mature adipocyte state in brown lineage (Tharp et al. 2018). Therefore, our results suggest that brown preadipocytes may have ongoing YAP/TAZ activity for maintenance of brown-lineage progenitor state. DGE analysis also revealed up-regulation of smooth-muscle lineage marker genes in brown preadipocytes, such as *TAGLN* (Fig. 1C), *ACTA2*, *MYL9*, and *CNN1* (Supplemental Table S1B). These findings are consistent with a recent study that showed abundant expression of smooth-muscle lineage-selective genes in clonal human brown preadipocytes (Shinoda et al. 2015), suggesting that brown preadipocytes derived from human neck depot may share this lineage.

Of note, we identified two distinct cell populations within brown preadipocytes (cluster 1 and cluster 2) (Fig. 1B). Gene Ontology (GO) analysis identified cellular adhesion and regulation of cellular motility as the most enriched terms in cluster 1 (Fig. 1D), suggesting the prevalence of stem cell-like migratory behavior in these cells. Transforming growth factor superfamily genes (*BMP4* and *TGFB2*) were also enriched in cluster 1 (Supplemental Table S1C), which play an important role in regulating adipocyte commitment in mesenchymal stem cells (Modica and Wolfrum 2017; Li and Wu 2020). Investigating differential activity of transcription factors (TFs) in cluster 1, transcription factor enrichment analysis (TFEA) identified *FOX* (*FOXC2* and *FOXL1*) and *FOSL1* TFs with high activity (Supplemental Table S1D). *FOXC2* participates in the early regulation of preadipocyte differentiation (Gerin et al. 2009; Lidell et al. 2013), whereas *FOSL1* proteins have been implicated as regulators of cell differentiation and transformation (Luther et al. 2011, 2014). Therefore, our results indicate that

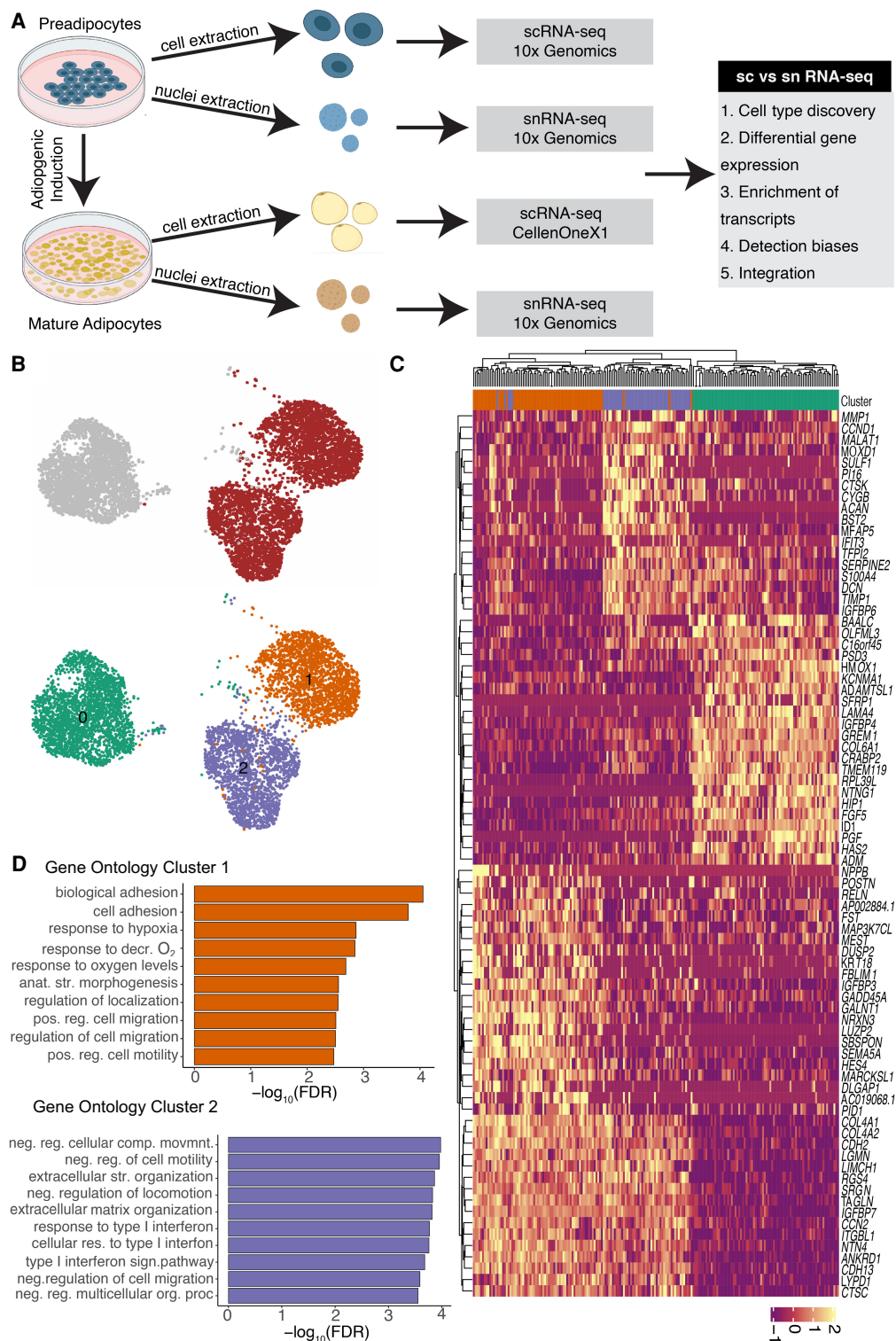


Figure 1. scRNA-seq reveals transcriptional and compositional landscapes of white and brown preadipocytes. (A) A schematic representation of scRNA-seq versus snRNA-seq characterizations performed in our study. Comparisons were performed at both the preadipocyte stage and the mature adipocyte stage using an in vitro system of human preadipocytes. For the cell culture maintenance and differentiation protocol, see the Supplemental Methods. (B) UMAP visualization of white and brown preadipocytes either annotated manually to reflect the sample of origin (*top panel*) or annotated based on unsupervised clustering (*bottom panel*); 2578 white and 4055 brown cells were detected. Of these cells, 2558 were in cluster 0, 2062 in cluster 1, and 2013 in cluster 2. (C) Heat map of the top 20 differentially expressed genes between white and brown preadipocytes, as well as cluster 1 and cluster 2 preadipocytes based on log fold change values. Both genes (rows) and cells (column) underwent unsupervised clustering, with the *top row* reflecting cluster assignment as in *B*. (D) Top 10 Gene Ontology biological processes terms enriched in brown preadipocyte cluster 1 (*top panel*) and cluster 2 (*bottom panel*). (decr.) Decreased, (anat. str.) anatomical structure, (pos. reg.) positive regulation, (neg. reg) negative regulation, (str.) structural.

cluster 1 cells may show migratory behavior with ongoing signaling similar to adipogenic fate commitment in mesenchymal stem cells, a behavior we refer to here as stem cell–like. Enrichment of multiple regulators of adipose tissue development was also detected in cluster 1, such as *SEMA5A* (Giordano et al. 2001), *NPPB* (Villarroya and Vidal-Puig 2013), *MEST* (Karbiener et al. 2015), and *FST* (Braga et al. 2014), further suggesting the existence of adipogenic commitment activity in this cell population.

Cluster 2 cells were marked by the expression of the *S100A4* gene, also known as the fibroblast-specific protein-1 (*FSP1*) (Fig. 1C; Supplemental Fig. S2D,F), which is considered a reliable marker of fibroblasts (Strutz et al. 1995). GO analysis showed enrichment of immune response, extracellular structure, and matrix organization, as well as negative regulation of cell migration terms, in this cell population (Fig. 1D). Multiple genes encoding for extracellular matrix (ECM) components such as *MFAP5*, *ECM1*, *COL6A2*, and *ACAN* were also enriched in cluster 2 (Supplemental Table S1C). Recent investigations have reported the presence of *S100A4*⁺ fibroblasts in the adipogenic niche, with a potential role in maintaining adipose homeostasis (Hou et al. 2018; Zhang et al. 2018; Vijay et al. 2020). The markers identified for fibroblasts in these investigations—*FBN1*, *IGFBP6*, *MFAP5*, *S100A4*, and *PI16*—were some of the most enriched markers of cluster 2 cells (Supplemental Fig. S2A–F). Therefore, these results indicate that cluster 2 cells are fibroblast-like, with negative regulation of cellular migration and an ongoing activity for ECM organization. The existence of two phenotypically distinct brown preadipocytes was further corroborated by performing single-molecule fluorescent in situ hybridization (smFISH) imaging of the cluster 2–enriched gene *MMP1* (Supplemental Note S2; Supplemental Fig. S3).

Recently, snRNA-seq of primary human WAT harvested from the same region as our model system (neck) identified a single white preadipocyte population (Sun et al. 2020), whereas multiple scRNA-seq studies identified two to three adipocyte progenitor populations in adult murine abdominal WAT (Ferrero et al. 2020). These differences in WAT preadipocyte composition could arise from species-specific and/or depot-specific variation. In contrast to our findings, snRNA-seq of primary human neck BAT also revealed a homogenous brown preadipocyte population (Sun et al. 2020), possibly because of differences during tissue biopsy collection or because of poor cell-capture efficiencies during single-nuclei isolation.

snRNA-seq identifies the same preadipocyte populations as scRNA-seq and detects biologically relevant differential expression

To evaluate the efficacy of snRNA-seq for recovering transcriptional heterogeneity, we sequenced the nuclear transcriptome of single preadipocytes from the white and brown lineages. Unsupervised clustering of the two lineages grouped nuclei into four clusters, referred to as populations 0, 1, 2, and 3 (Supplemental Fig. S4A,B). Cluster 3 nuclei, however, had enriched expression for stress-response genes and mitochondrial genes, along with high background RNA contamination (Supplemental Fig. S4D) and, hence, were removed from downstream analyses. In the remaining clusters, brown nuclei were primarily grouped into clusters 1 and 2, whereas white nuclei grouped into a single cluster, 0 (Fig. 2A). Similarity between clusters identified in snRNA-seq and scRNA-seq was assessed using the concept of transcriptional signatures (DeTomaso and Yosef 2016; Gaublot et al. 2019), defined as genes differentially expressed in either white vs brown preadipo-

cytes, or cluster 1 versus cluster 2, in the scRNA-seq data set (Supplemental Table S1B,C). As expected, the transcriptional signature scores, calculated using Vision (DeTomaso et al. 2019), were enriched in the corresponding preadipocyte-type/clusters in the snRNA-seq data set (Fig. 2B), thereby showing a high concordance between transcriptional features uncovered by the two techniques.

As was observed with scRNA-seq, white nuclei were enriched for the genes *TCF21*, *PAX3*, and *PDGFRA* (Supplemental Table S2A), and brown nuclei were enriched for the YAP/TAZ target genes *ANKRD1* and *CCN2* (Supplemental Table S2A) and the smooth-muscle lineage marker genes *TAGLN*, *MYL9*, *CNN1*, and *MYH11* (Supplemental Table S2A). Gene *IDI1*, however, was not differentially enriched in white nuclei because of a lack of differential enrichment in the nuclear compartment between white and brown preadipocyte (Supplemental Note S3; Supplemental Fig. S5). In scRNA-seq data set, we had classified certain DE genes as markers for white and brown preadipocytes based on their highly enriched and specific expression (Supplemental Note S6). All such white preadipocyte– and brown preadipocyte–specific marker genes were also enriched in white and brown nuclei, respectively (Supplemental Table S2A). Of the 50 genes with maximum enrichment (ordered by logFC) in white and brown preadipocytes in the scRNA-seq data set, >94% were also differentially expressed in white and brown nuclei, respectively (Fig. 2C). This analysis shows that snRNA-seq has sufficient sensitivity to recover the same molecular differences as scRNA-seq between white and brown preadipocytes.

GO analysis identified enrichment of cellular adhesion and regulation of cellular localization terms in brown cluster 1 nuclei, corresponding with the findings in scRNA-seq data set (Supplemental Fig. S4E). Transforming growth factor superfamily genes *BMP4* and *TGFB2* were also enriched in cluster 1, along with regulators of adipose tissue development *SEMA5A*, *MEST*, and *FST* (Supplemental Table S2B). All six cluster 1–specific marker genes (Supplemental Note S6) identified were also enriched in cluster 1 nuclei (Supplemental Table S2B). Of the 50 genes with maximum enrichment (ordered by logFC) in cluster 1 cells in the scRNA-seq data set, 94% were also differentially expressed in the nuclear data set (Fig. 2C). In cluster 2 brown nuclei, enrichment of *S100A4* was observed (Supplemental Table S2B), as well as regulation of ECM organization terms based on GO analysis (Supplemental Fig. S4F). Genes encoding for the ECM components *COL6A2*, *MFAP5*, *ACAN*, and *ECM1* were all up-regulated in cluster 2 (Supplemental Table S2B). Of the 50 genes with maximum enrichment (logFC) in cluster 2 brown preadipocytes (scRNA-seq data set), 80% were also differentially expressed in the nuclear data set (Fig. 2C). All cluster 2–specific marker genes (Supplemental Note S6) identified were also enriched in cluster 2 nuclei (Supplemental Table S2B). Overall, our snRNA-seq analyses indicated the emergence of stem cell–like behavior in cluster 1 and fibroblast-like behavior in cluster 2, in agreement with the whole-cell data set. Finally, preadipocyte-/cluster-specific transcriptional signatures now defined using the snRNA-seq data set revealed enrichment in the corresponding preadipocyte-type/clusters in the scRNA-seq data set, thereby validating that markers derived from snRNA-seq can be used to identify the same populations in whole-cell analysis (Supplemental Fig. S4H).

Gene length–associated detection bias in snRNA-seq

Typical scRNA-seq data analysis pipelines often filter intronic reads for downstream count matrix generation. More recently,

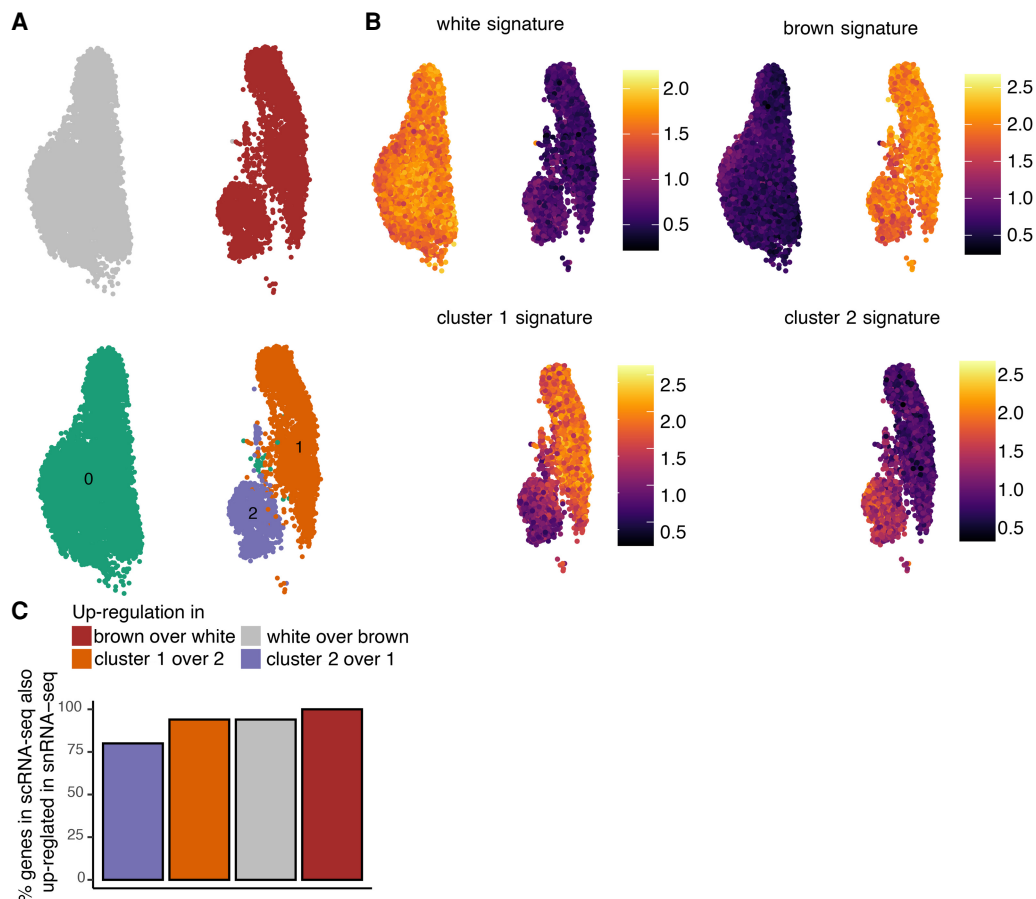


Figure 2. snRNA-seq identifies the same preadipocyte populations as scRNA-seq and detects biologically relevant differential expression. (A) UMAP visualization of white and brown preadipocytes either annotated manually to reflect the sample of origin (top) or annotated based on unsupervised clustering (bottom); 6556 white and 3891 brown nuclei were detected. Of these nuclei, 6578 were in cluster 0, 2716 in cluster 1, and 1153 in cluster 2. (B) Heat map of transcriptional signature scores for white preadipocyte (top left), brown preadipocyte (top right), brown preadipocyte cluster 1 (bottom left), and brown preadipocyte cluster 2 (bottom right) as plotted on the UMAP visualization of snRNA-seq data. (C) Bar plot of percentage of top 50 genes differentially enriched (DE) in the scRNA-seq data set that are also DE in the snRNA-seq data set. The top 50 genes were evaluated based on log fold change values using the scRNA-seq data set.

however, evidence has suggested that intronic reads originate from nascent transcripts (Ameur et al. 2011; Gray et al. 2014; Hendriks et al. 2014) and, hence, are informative about expression levels in single-cell data. Furthermore, the additional read counts improve gene detection sensitivity and can improve cell-cluster resolution (Bakken et al. 2018; Wu et al. 2019). Multiple recent studies have suggested internal hybridization of poly(T) RT-primer to intronic poly(A) stretches in nascent transcripts as the primary mechanism for the capture and detection of intronic reads (La Manno et al. 2018; Shulman and Elkon 2019; Patrick et al. 2020). Consequently, intronic reads are more readily detected in genes with more intronic poly(A) stretches, which are more likely to be longer in length (Supplemental Fig. S6A). This bias is increased in nuclear libraries where up to 40% of all the reads map to intronic regions compared with only 9% in scRNA-seq (Fig. 3A). Consequently, recent studies have reported enrichment of longer genes (Lake et al. 2017; Bakken et al. 2018) and poor detection of shorter genes (Thrupp et al. 2020) in nuclei.

To examine the enrichment of long genes in nuclei, we first performed DGE analysis between cells and nuclei in white preadipocytes. Using both intronic and exonic reads, our analysis identified 493 genes enriched in cells and 568 genes enriched in nuclei

(logFC > 1 and FDR < 0.05). This analysis revealed nuclear-enriched genes to be significantly longer than genes enriched in whole cells (two-group Mann-Whitney *U* test, *P*-value < 0.01) (Fig. 3B). Next, we performed DGE analysis between white and brown nuclei, with and without intronic reads, for genes that are enriched in white preadipocytes in the scRNA-seq data set. Certain long genes such as *KCNMA1* (99th percentile), *AK5* (96th percentile), and *BAALC* (85th percentile) were differentially expressed (in white nuclei over brown nuclei; DE) only upon inclusion of intronic reads (non-DE with only exonic reads), likely because of their preferential detection (Fig. 3C,D, highlighted in red). Conversely, we also identified certain short genes such as *CRABP2* (32nd percentile), *TM4SF1* (40th percentile), and *EVA1B* (17th percentile) that remain nondifferentially expressed, even with inclusion of intronic reads in the snRNA-seq data set (also non-DE with only exonic reads) (Fig. 3C,D, highlighted in blue). We also performed DGE analysis between white cells and white nuclei using only exonic reads (logFC > 0.25 and FDR < 0.05). The logFC differential enrichment for nuclear-enriched genes was poorly correlated with counting exons or exons and introns (Pearson *R* = 0.50, *P*-value < 0.01) (Fig. 3E). logFC values for some of the longest genes were artificially inflated, possibly because of their preferential

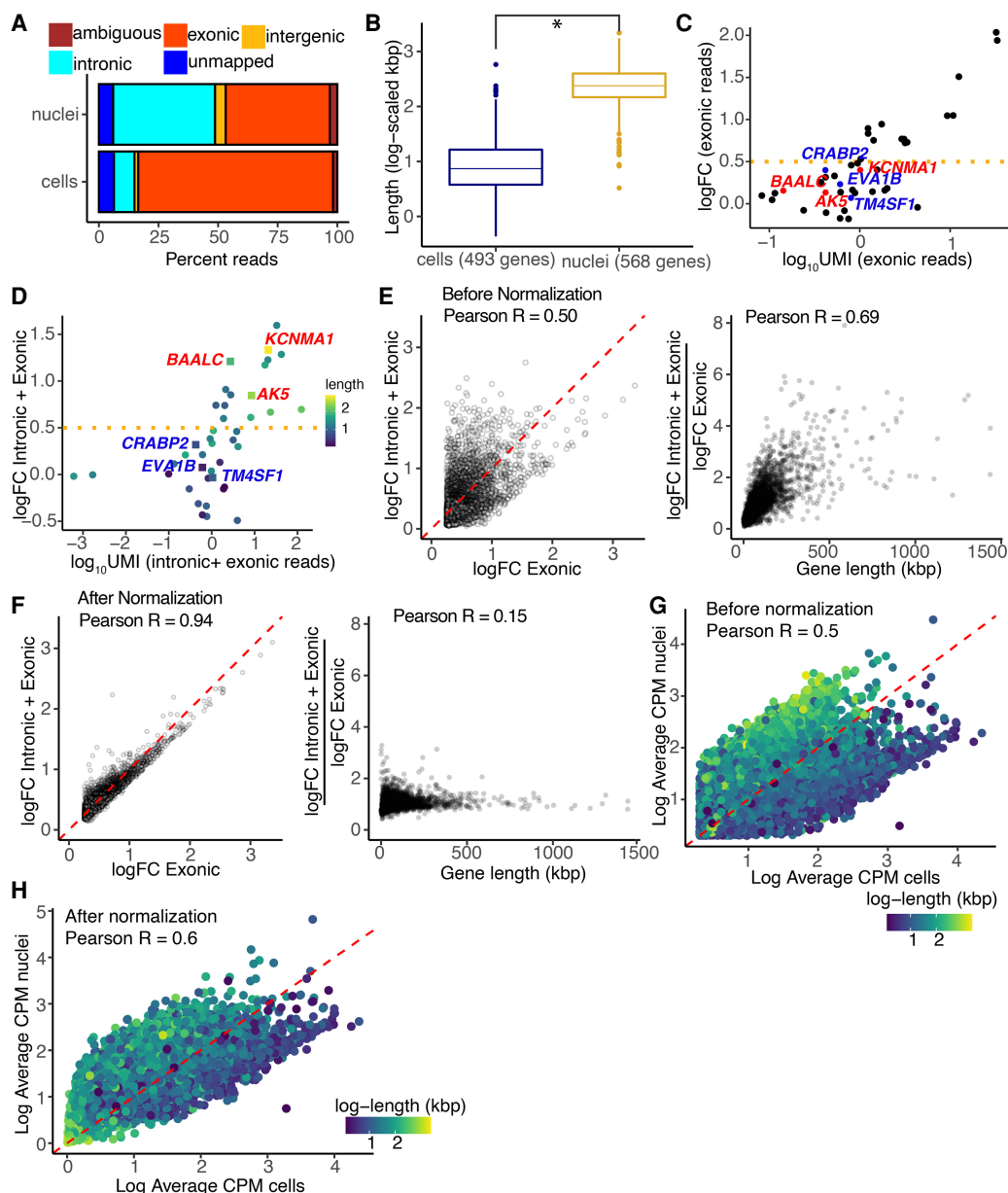


Figure 3. Gene length-associated detection bias in the nuclear transcriptome. (A) Distribution of reads in the scRNA-seq and snRNA-seq preadipocyte data sets. (B) Distribution of gene length for genes enriched in cells (in blue) and nuclei (in yellow) with log fold change > 1 and FDR < 0.05 , including both intronic and exonic reads. (C) logFC versus log-UMI counts in white nuclei when using only exonic reads, where each dot represents a white preadipocyte-enriched gene (white vs. brown DE test) detected using the scRNA-seq data set. Horizontal dotted line indicates a logFC cutoff value of 0.5 used as a threshold for DE testing. All genes had a logFC > 0.5 in the scRNA-seq data set. (D) logFC versus log-UMI counts in white nuclei when using both intronic and exonic reads. Each dot is the same as in panel C. Horizontal dotted line indicates a logFC cutoff value of 0.5 used as a threshold for DE testing. Highlighted genes are represented with a square symbol. (E, left) logFC for nuclear-enriched genes when using only exonic reads or both intronic and exonic reads before normalization. Each dot represents a gene enriched in nuclei using exonic-only reads with logFC > 0.25 and FDR < 0.05 . Red dotted line indicates $y = x$ axis. (Right) Ratio of y -axis value over x -axis value for genes in left panel, plotted as a function of their length. (F, left) logFC for nuclear-enriched genes when using only exonic reads or both intronic and exonic reads after normalization. Each dot is the same as in panel E. Red dotted line indicates $y = x$ axis. (Right) Ratio of y -axis value over x -axis value for genes in left panel, plotted as a function of their length. (G, H) Average expression of genes in white cells and white nuclei when using both intronic and exonic reads, without normalization (G) and with normalization (H). For normalization strategy, see Supplemental Note S5. Each dot represents a gene with average counts per million (CPM) > 1 when using both intronic and exonic reads, without normalization, in both cells and nuclei. White nuclei were randomly selected to have as many barcodes as white cells. Red dotted line has a slope of 1.

detection upon inclusion of intronic reads (Fig. 3E). Conversely, logFC values for some of the shortest genes were artificially deflated because of their poor detection (Fig. 3E). Consequently, the ratio of the logFC values with counting exons or exons and introns

was strongly correlated with gene length (Pearson $R = 0.69$, P -value < 0.01) (Fig. 3E). Overall, our results show technical artifacts induced by gene length-associated detection bias in snRNA-seq upon inclusion of intronic reads. We therefore developed a

normalization strategy to address this technical, length-associated detection bias (Supplemental Note S5; Supplemental Fig. S6). After normalization, the logFC differential enrichment of nuclear-enriched genes was highly correlated with counting exons or exons and introns (Pearson $R=0.94$, P -value <0.01) (Fig. 3F). Moreover, the ratio of the logFC values with counting exons or exons and introns, after normalization, was poorly correlated with gene length (Pearson $R=0.15$, P -value <0.01) (Fig. 3F). Nuclear and cellular transcriptomes were also better correlated after removal of technical biases using our normalization strategy (Fig. 3G,H).

DGE analysis, between white cells and white nuclei, with normalized read counts identified 382 enriched genes in cells and 249 enriched genes in nuclei (logFC >1 and FDR <0.05), with nuclear-enriched genes still significantly longer than whole cells (two-group Mann–Whitney U test, P -value <0.01) (Supplemental Fig. S7A). However, the genes enriched in nuclei were, on average, 14-fold longer than genes enriched in cells (compared with a 32-fold difference before normalization), which is comparable with the difference observed when using only exonic reads (11-fold difference) (Supplemental Fig. S7B), suggesting that after accounting for technical bias, there also exists biological enrichment of longer genes in nuclei. Overall, our observations show that length normalization removes artificial detection biases, thereby improving UMI count estimation accuracy while also preserving improved gene detection sensitivity afforded by inclusion of intronic reads.

To further understand differential transcript enrichment between whole-cell and nuclear transcriptomes, we next focused on genes enriched in whole cells after normalization. GO analysis identified protein translation-associated terms as most enriched in whole cells (Supplemental Fig. S7D). Genes contributing to the enrichment of translational terms primarily included mRNAs encoding for ribosomal proteins. This enrichment of ribosomal-protein mRNAs in whole cells is consistent with their very low cytoplasmic decay rates and selective nuclear export machinery (Wickramasinghe et al. 2014; Chen and van Steensel 2017). Yet, poor detection of ribosomal proteins in the nuclear transcriptome did not affect the ability to resolve cellular populations in snRNA-seq, as evident by the score of the transcriptional signature consisting of the top 100 genes enriched in cells based on logFC values (approximately 53/100 ribosomal protein genes) (Supplemental Fig. S7C).

Nuclear transcriptome is enriched for long noncoding RNAs that regulate adipogenesis and drive cell type differences

Long noncoding RNAs (lncRNAs) function in regulating diverse biological processes, including regulation of transcription, proliferation, pluripotency, and cellular differentiation (Quinodoz and Guttman 2014; Samata and Akhtar 2018; Sherstyuk et al. 2018; Del a et al. 2019). Because of their regulatory function, lncRNAs predominantly remain localized in the nucleus (Cabili et al. 2015; Wen et al. 2018). snRNA-seq intrinsically enriches for nuclear localized transcripts, and previous studies have reported enrichment of lncRNAs in snRNA-seq libraries over scRNA-seq (Grindberg et al. 2013; Zeng et al. 2016). We hypothesized that nuclear enrichment of lncRNAs could be advantageous for characterizing adipose tissue because multiple lncRNAs have also been implicated in regulating adipogenesis (Sun et al. 2013; Wei et al. 2016; Ding et al. 2018; Sun and Lin 2019; Zhou et al. 2020). We tested this hypothesis in our *in vitro* system by profiling adipogenic regulatory lncRNAs in our whole-cell and nuclear libraries derived from white preadipocytes after normalization. We identified significant enrichment of the lncRNAs *NEATI* (Wei et al.

2016), *MEG3* (Li et al. 2017), *MIR31HG* (Huang et al. 2017), and *PVT1* (Zhang et al. 2020) in white nuclei, which are previously reported regulators of adipogenesis (Fig. 4A). All four lncRNAs were also enriched in brown nuclei compared with brown whole cells (Supplemental Fig. S8A–D). Generally, snRNA-seq consistently detected a greater number of lncRNAs at all read depths than did scRNA-seq (P -value <0.01 , two-group Mann–Whitney U test) (Fig. 4B). Of the 111 differentially expressed lncRNAs between white nuclei and white cells, ~86% (96/111 genes) were up-regulated in nuclei, thereby validating a higher prevalence of this class of genes in the nuclear compartment. Seven out of 15 lncRNAs that were enriched in white cells were snRNA host genes (SNHG), which have been shown to have various functions in cytoplasm such as repressing mRNA translation, miRNA sponging, and protein ubiquitination (Zimta et al. 2020). Overall, our results suggest a higher likelihood to deconstruct the functional roles of adipogenic regulatory lncRNAs (and other lncRNAs in general) using snRNA-seq.

Next, we evaluated the sensitivity of snRNA-seq for detection of lncRNAs driving molecular heterogeneity between brown preadipocyte cluster 1 and 2, two cell types most closely related to each other. At approximately 50,000 reads per cell/nuclei, DGE analysis identified more than 40 lncRNAs distinctively regulated between cluster 1 and 2 in the snRNA-seq data set compared with only 15 lncRNAs in the scRNA-seq data set. Unsupervised hierarchical clustering in the snRNA-seq data set based on the expression of the top 20 up-regulated lncRNAs in cluster 1 and 2 each revealed sorting of nuclei into two distinct groups that predominantly reflected their original cluster assignment (Fig. 4C). Moreover, silhouette coefficient analysis (a method for evaluating clustering performance) revealed better cluster separation performance for snRNA-seq compared with scRNA-seq between cluster 1 and 2 for all down-sampled read depths (Fig. 4D). Silhouette coefficients were calculated based on the Euclidean distance between cells/nuclei in the principal component space generated using only lncRNAs (see Supplemental Material). To validate that the observed performance features were not metric dependent, we quantified two more indices, the Calinski–Harabasz index and the Davies–Bouldin index, to compute inter-cluster separation and found similar trends (Supplemental Fig. S8F,G). A similar analysis performed by normalizing for the same number of mean unique molecules (UMIs) per sample revealed a similar trend for the three separation indices (Supplemental Fig. S8H–J). Together, our results suggest that snRNA-seq is superior for learning heterogeneity governed by lncRNAs compared with scRNA-seq.

snRNA-seq detects relevant transcriptional regulation during adipogenesis in white preadipocytes

After identifying transcriptomic similarities and differences between scRNA-seq and snRNA-seq in the preadipocyte state, we next focused on evaluating molecular correspondence between the two techniques in mature adipocytes. We leveraged our *in vitro* model of white adipogenesis that enabled us to prepare a single-cell suspension of mature adipocytes without the need of implementing harsh tissue dissociation protocols (see Methods). Following single-cell suspension preparation, one of the most common ways to sort single cells is using flow cytometry. Recently, FACS gating strategies have been tailored to isolate mature adipocytes (Majka et al. 2014; Hagberg et al. 2018), although only a small percentage of adipocytes are able to survive the shear stress associated with flow sorting (Majka et al. 2014). Therefore, to

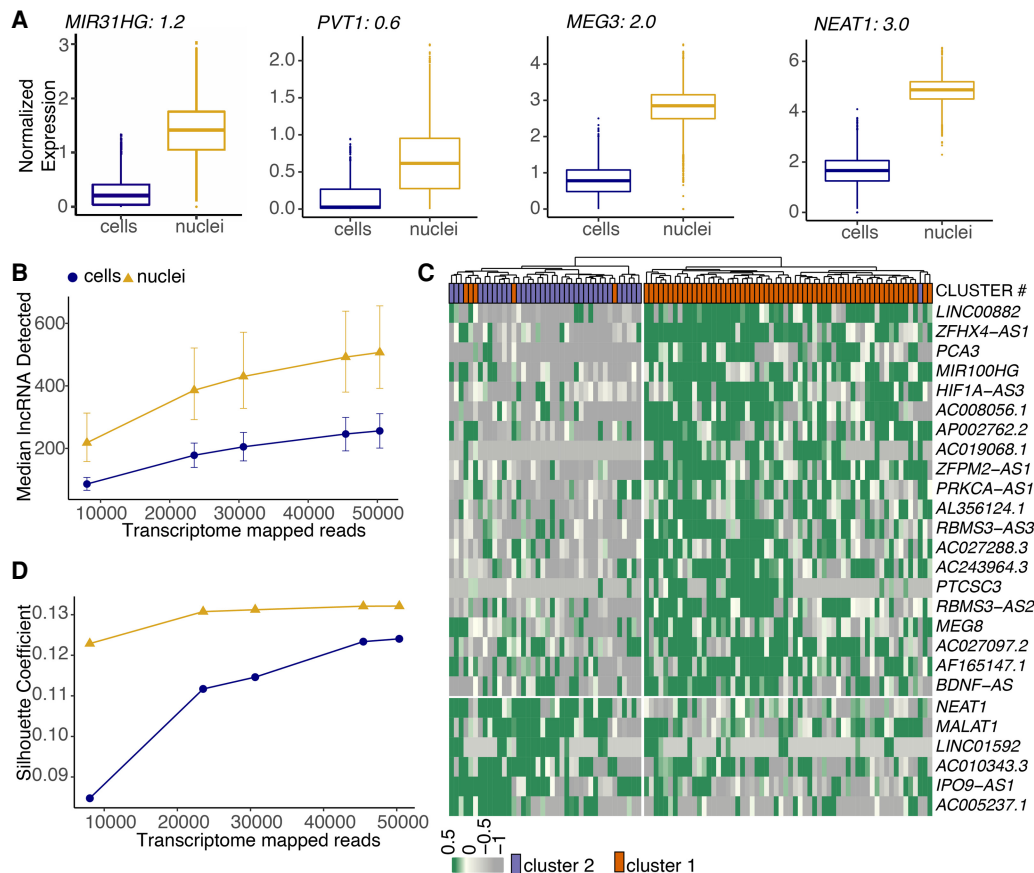


Figure 4. Nuclear transcriptome is enriched for lncRNAs that regulate adipogenesis and drive cell type differences. (A) Box plots of lncRNAs reported as regulators of adipogenesis. Black text indicates logFC value for white nuclei versus white cell DE test in preadipocytes with a FDR < 0.05 after normalization. (B) Median lncRNAs detected as a function of read depth across single cells and nuclei (both white and brown lineages). Error bars indicate the interquartile range. (C) Hierarchical clustering using scaled expression values of the top-20 up-regulated lncRNAs in brown cluster 1 and cluster 2 in the snRNA-seq data set. One-hundred random barcodes were chosen for this analysis. Top row reflects original cluster assignment for the selected barcodes. (D) Cluster separation resolution quantification between brown cluster 2 versus cluster 1 in the scRNA-seq and snRNA-seq data set. Only lncRNAs were considered for PCA manifold generation. Both data sets were subsampled to have the same number of cells/nuclei and same number of mean transcriptome mapped reads.

enable gentle sorting of single adipocytes for downstream scRNA-seq, we developed a new protocol using the cellenONE X1 single-cell isolation platform. This automated liquid-handling robot uses gentle piezo-acoustic technology for dispensing cells encapsulated in a picoliter-volume droplet, ensuring minimal cellular perturbation and background RNA contamination. To harvest adipocytes *in vitro*, human white preadipocytes were cultured and differentiated using a chemical adipogenic induction cocktail for 20 d (Shamsi and Tseng 2017). Coherent antistokes Raman imaging established successful differentiation of white preadipocytes, with distinctly visible signal from round lipid droplets (Supplemental Fig. S9A; Gupta et al. 2019). After creating a single-cell suspension of white adipocytes, 200 cells were spotted using the cellenONE X1 machine onto 96-well plates preloaded with lysis buffer and barcoded poly(T) primer. Library preparation was then performed using mcSCR-seq chemistry (Bagnoli et al. 2018). Transcriptomic profiles of these cells were then compared with a snRNA-seq library of approximately 12,000 nuclei isolated from 20-d differentiated white adipocytes.

Independent unsupervised clustering revealed organization of both cells and nuclei into primarily two clusters, referred to as clusters 0 and 1 (Fig. 5A). snRNA-seq identified an additional cluster 2,

which showed characteristics of mitotic preadipocytes with ongoing cell cycle progression, suggesting that these cells could be preadipocytes that never underwent growth arrest (Supplemental Note S4; Supplemental Fig. S10A–H). Cluster 0 in both data sets was marked by the expression of mesenchymal marker *THY1* (Fig. 5A), suggesting that these cells/nuclei were differentiating preadipocytes. Cluster 1, on the other hand, had high expression of the adipogenic gene *ADIPOQ*, indicating that cells/nuclei in this cluster were mature adipocytes (Fig. 5A). DGE analysis further identified enrichment of other adipogenic marker genes (along with *ADIPOQ*) in cluster 1 (Fig. 5B,C, highlighted in red), confirming a transition from differentiating preadipocytes to mature adipocytes from cluster 0 to cluster 1 in both data sets. GO analysis identified enrichment of ECM organization terms in cluster 0 and lipid metabolism in cluster 1, independently in both the scRNA-seq and snRNA-seq data sets (Supplemental Fig. S9B–E). Moreover, ~80% genes (106/133) up-regulated in cluster 1 in the scRNA-seq data set were also differentially expressed in the snRNA-seq data set. Notably, the remaining 20% genes (27/133) that were not differentially expressed in the snRNA-seq data set primarily included genes associated with the mitochondrial respiratory chain process (Supplemental Fig. S9F), suggesting that adipocytes' enhanced

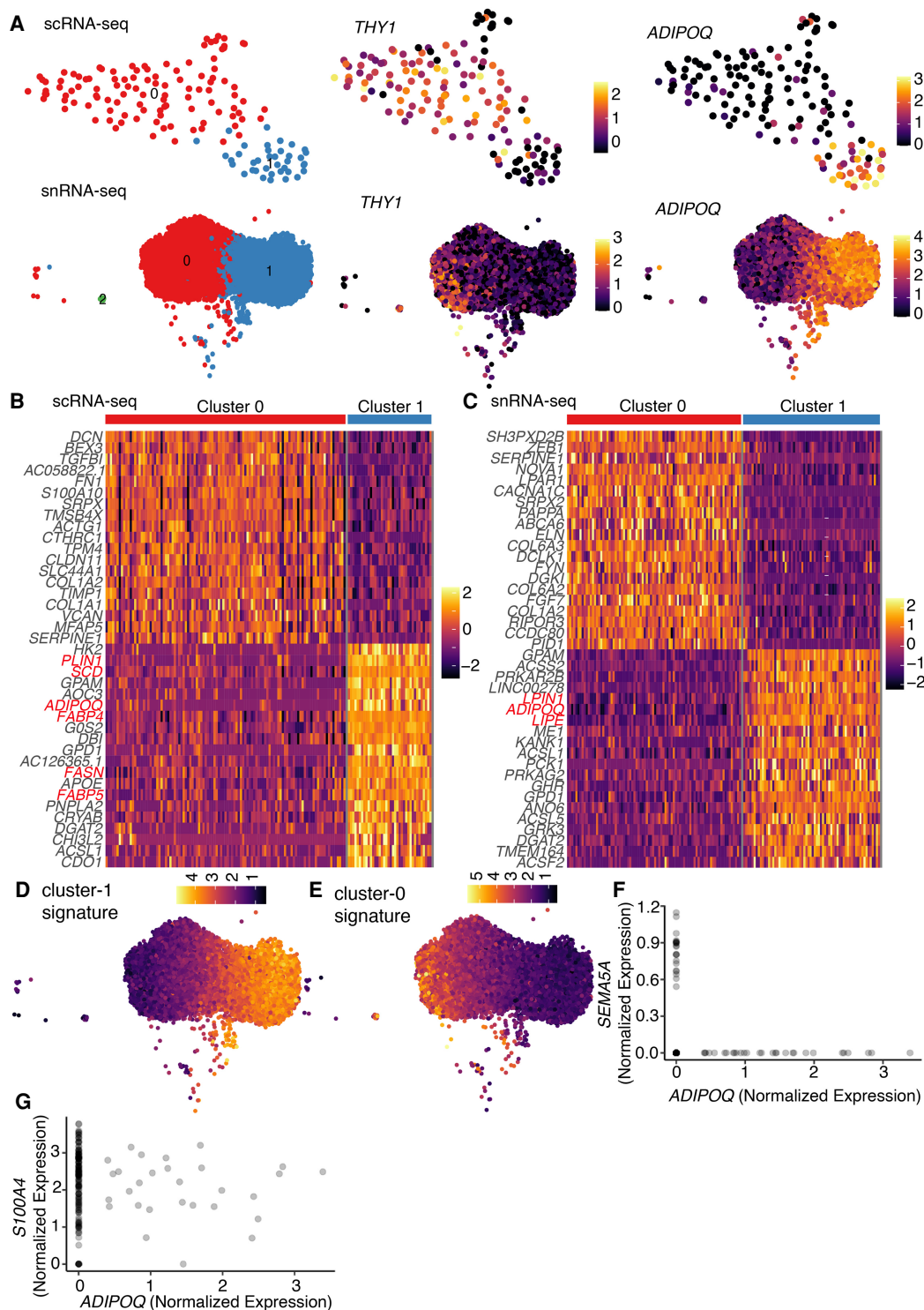


Figure 5. snRNA-seq detects important transcriptional regulation during adipogenesis in white preadipocytes. (A) UMAP visualization of scRNA-seq and snRNA-seq white adipocyte data sets (day 20) after unsupervised clustering (left panels). The scRNA-seq data set had 109 cells in cluster 0 and 38 cells in cluster 1. The snRNA-seq data set had 4859 nuclei in cluster 0, 4546 nuclei in cluster 1, and 31 nuclei in cluster 2. Expression profile for mesenchymal marker *THY1* and mature-adipocyte marker *ADIPOQ* in both scRNA-seq and snRNA-seq data sets (middle and right panels). (B) Heat map of z-scored expression of the top 20 differentially expressed genes between cluster 0 and cluster 1 in the scRNA-seq white adipocyte data set. Highlighted in red are markers of adipogenesis. (C) Heat map of z-scored expression of the top 20 differentially expressed genes between cluster 0 and cluster 1 in the snRNA-seq white adipocyte data set. A random subset of 150 barcodes was used for this visualization. Highlighted in red are markers of adipogenesis. (D) Heat map of transcriptional signature scores for cluster 1 as plotted on the UMAP visualization of snRNA-seq white adipocyte data. (E) Heat map of transcriptional signature scores for cluster 0 as plotted on the UMAP visualization of snRNA-seq white adipocyte data. (F, G) Normalized expression of genes *ADIPOQ* and *SEMA5A* (F) and *ADIPOQ* and *S100A4* (G) in differentiating brown preadipocytes (day 20) scRNA-seq data set. Also see Supplemental Figure S2, G–I.

mitochondrial activity may not be captured in the snRNA-seq data set. Correspondingly, the snRNA-seq data set lacked manifestation of mitochondrial biological processes such as oxidative phosphorylation and electron transport chain in cluster 1 upon GO (Supplemental Fig. S9B,D). This observation was also supported by the fact that these 27 genes had a median length of ~11 kb, the same order of magnitude as the length of genes with poor detection in nuclei over whole cells (Fig. 3B). As expected, scores of the cluster 0 and 1 transcriptional signatures in the scRNA-seq data set were observed to be highly conserved and enriched in corresponding cluster types (Fig. 5D,E), further validating the conservation of information in the nuclear transcriptome. Overall, our results reveal a comparable molecular landscape in white adipocytes between the scRNA-seq and snRNA-seq data sets.

We next looked to investigate any differential adipogenic capacity between the two brown preadipocyte clusters identified in our study (cluster 1 and 2) (Fig. 1B) by performing scRNA-seq on mature brown adipocytes using the cellenONE X1 for gentle isolation of intact mature adipocytes. These adipocytes were derived by differentiating brown preadipocytes for a period of 20 d. After sorting approximately 200 cells, library preparation was performed using the mcSCR-seq protocol (Bagnoli et al. 2018). Transcriptomic profiling revealed detection of mature brown adipocytes, along with recovery of multiple cells that were not terminally differentiated but rather distributed along a continuum of differentiation states (Supplemental Fig. S2G,H). Consequently, our analysis revealed a range of adipogenic gene expression (*ADIPOQ*) in our data set (Fig. 5F), which was mutually exclusive from the expression of the cluster 1-enriched gene *SEMA5A* (Fig. 5F). On the other hand, we identified multiple cells with shared expression of the cluster 2-enriched genes *S100A4* and *ADIPOQ* (Fig. 5G). These results supported the observation of two brown preadipocyte populations and indicate that cluster 2 cells are more likely to differentiate into mature brown adipocytes. Additionally, we also compared transcriptomic similarities between mature brown adipocytes (Supplemental Fig. S2H, highlighted in red) and cluster 1/cluster 2 cells using transcriptional signatures defined for respective clusters using the day-0 scRNA-seq data set (Supplemental Table S1C). Mature adipocytes had a significantly higher score for cluster 2 cells compared with those of cluster 1 (Supplemental Fig. S2I), thereby providing additional evidence that the former cell type is more likely to be adipogenic.

Integration of snRNA-seq and scRNA-seq data sets

A comprehensive cell atlas of the adipose tissue will require joint analyses of data sets generated using both scRNA-seq and snRNA-seq. However, technical biases and differential transcript enrichment in snRNA-seq lead to significant batch effects between snRNA-seq and scRNA-seq experiments, thereby reducing clusterability of cells from these two protocols (Mereu et al. 2020). Multiple bioinformatic tools are now available to remove covariates that lead to technical batch effects and to facilitate integration of scRNA-seq data sets generated across different days, laboratories, individuals, or technologies (Zappia et al. 2018). We used single-cell variational inference (scVI), a deep generative modeling-based tool (Lopez et al. 2018; Gayoso et al. 2021), to explore the possibility of integrating the snRNA-seq and scRNA-seq data sets for joint analysis. Four data sets of white preadipocytes were integrated in total: day-0 scRNA-seq and snRNA-seq (cluster 0 in Figs. 1B, 2A) and day-20 scRNA-seq and snRNA-seq (top and bottom left panels in Fig. 5A).

Without batch correction, all four data sets arranged into distinct individual clusters, with no shared population identified at the same time point across different techniques or the same technique but across different time points (Fig. 6A). A dendrogram, based on the Euclidean distance in dimensionally reduced space, grouped clusters first by sequencing chemistry (mcSCR-seq vs. 10x), followed by technique type (snRNA-seq vs. scRNA-seq), and finally by time point (day 0 vs. day 20) (Fig. 6A). After integration, matching adipocyte populations from day 20 and preadipocyte populations from both day 0 and day 20 in nuclear and whole-cell data sets were primarily nearest neighbors in a dendrogram based on the Euclidean distance in dimensionally reduced space (Fig. 6B). UMAP visualization further revealed proximal placements of similar cell populations (Fig. 6B,C). Of note, we observed that preadipocytes from both the day-0 snRNA-seq and day-0 scRNA-seq data sets localized into two distinct groups, which was driven by differences in proliferation state with one cluster composed of mitotic cells and another composed of growth arrested cells (Supplemental Note S4; Supplemental Fig. S10A–H). Unsupervised clustering of the integrated data set revealed adipocytes, day-20 preadipocytes, and the two groups of day-0 preadipocytes as distinct cell types, illustrating scVI's abilities to remove batch effects while retaining biological variation (Fig. 6C, right panel). Indeed, top marker genes for each cluster recovered previously reported expression trends such as enrichment of ECM components in day-0 preadipocytes (Weiner et al. 1989), enrichment of insulin-binding proteins in day-20 preadipocytes (Gleason et al. 2010), and enrichment of adipogenic genes in mature adipocytes (Supplemental Fig. S10J). Although multiple markers for each cluster had conserved expression across scRNA-seq and snRNA-seq, some markers were exclusively enriched in either one of the data sets (Supplemental Fig. S10J), thereby highlighting the importance of performing joint analysis. Finally, integration of the same four data sets using Seurat (Stuart et al. 2019) revealed minimal overlap of single-cell and single-nuclei data sets for both day 0 and day 20 (Supplemental Fig. S10I). Recently, benchmarking of distinct integration methodologies indeed revealed effective performance by scVI on complex integration tasks, with Seurat v3 performing well on simpler tasks with distinct biological signals (Luecken et al. 2022). Overall, our results show scVI's integration abilities by identifying functionally similar preadipocyte and adipocyte populations shared across scRNA-seq and snRNA-seq techniques.

Discussion

In this investigation, we evaluated the ability of snRNA-seq to recapitulate the molecular and compositional landscape of distinct lineages in human adipose tissue. We avoided confounding variability associated with inter-depot and inter-subject transcriptional variation by performing a direct comparison of snRNA-seq and scRNA-seq on a pair of immortalized white and brown human preadipocytes isolated from the neck region of the same individual. We found that snRNA-seq was able to recover the same cell types as scRNA-seq at both the preadipocyte and mature adipocyte states. Furthermore, we provided evidence for recovering similar expression profiles of biologically relevant genes and attributing similar functional annotations to cell types by nuclear transcriptome profiling compared with whole cells.

At the preadipocyte stage, brown preadipocytes were a heterogeneous mix of two distinct cell populations, cluster 1 and cluster 2. However, cell type enrichment followed by differentiation and metabolic assays will need to be further performed to identify their

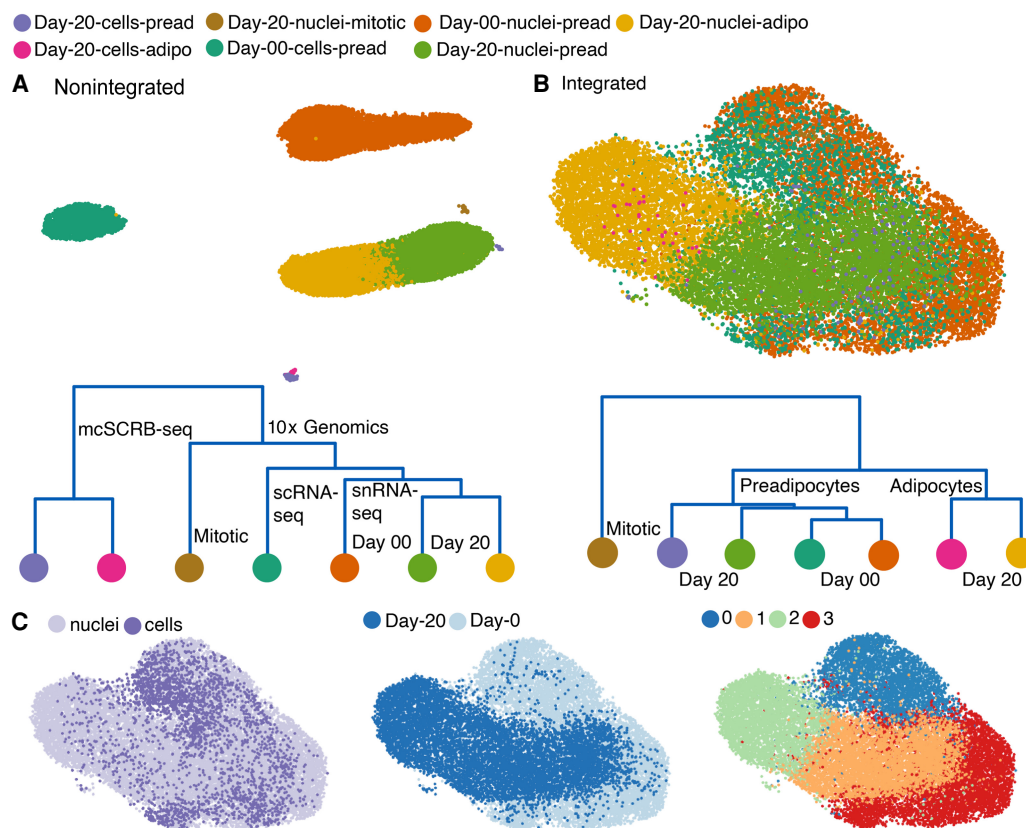


Figure 6. Integration of the snRNA-seq and scRNA-seq data sets. (A, top panel) UMAP visualization of nonintegrated scRNA-seq and snRNA-seq data sets for both white preadipocyte (day 0) and mature adipocyte (day 20), for a total of four batches (total 18,717 barcodes). (Bottom panel) Cluster dendrogram for nonintegrated data sets based on the eigenvalue-weighted Euclidean distance matrix constructed in latent-dimension space inferred using scVI. (B) UMAP visualization and cluster dendrogram of scRNA-seq and snRNA-seq data sets as in A after integration using scVI tools (total 18,717 barcodes). See also Supplemental Note S4 and Supplemental Figure S10. (C) UMAP visualization of scVI integrated data set with barcodes annotated by sequencing technique (left), harvest day (middle), and joint unsupervised clustering (right).

individual functions in maintaining adipose tissue homeostasis. To date, different scRNA-seq studies of mouse stromal vascular fraction (SVF) have identified multiple subpopulations of adipose progenitor cells (APCs) expressing distinct markers (Burl et al. 2018; Hepler et al. 2018; Schwalie et al. 2018; Merrick et al. 2019). Integrated analysis of these data sets primarily identified two common populations of APCs in mice referred to as Asc1 and Asc2 (Ferrero et al. 2020; Rondini and Granneman 2020). Like cluster 2 cells, Asc2 showed proinflammatory and profibrotic phenotype and positive expression of genes *P116* and *MFAP5*. Functional investigations into the two cell types revealed Asc2 cells inhibiting the differentiation of Asc1 cells in vitro (Rondini and Granneman 2020). This agrees with the proadipogenic nature of cluster 2 cells identified in our study. Therefore, it is plausible that cluster 1 and cluster 2 cells identified in our study may be functioning in a manner like Asc1 and Asc2 to maintain adipocyte turnover.

snRNA-seq is the preferred technique to study samples whose compositional landscape may be biased by the differential efficiency of cell type recovery when using scRNA-seq. Adipose tissue is one such sample in which isolation of intact, single adipocytes is complicated by their fragile nature. Indeed, most adipose scRNA-seq studies to date derive the transcriptomes of the cell types within SVF only, with minimal detection efficiencies for mature adipocytes either in animal models (Burl et al. 2018; Hepler et al.

2018; Schwalie et al. 2018; The Tabula Muris Consortium 2018; Cho et al. 2019; Gu et al. 2019; Jaitin et al. 2019; Merrick et al. 2019; Henriques et al. 2020) or humans (Acosta et al. 2017; Jaitin et al. 2019; Merrick et al. 2019; Vijay et al. 2020). Here, we developed a new single-adipocyte isolation protocol using piezo-acoustic-based gentle dispensing technology for improved recovery with downstream scRNA-seq. Using this strategy, ~26% barcodes recovered were annotated as adipocytes. However, this adipocyte capture efficiency was still limited compared with snRNA-seq, where ~48% barcodes were identified as adipocytes. Conversely, at the preadipocyte stage, where cell type recovery is efficient, scRNA-seq recovered equal proportions of the two brown preadipocyte clusters. However, analysis of snRNA-seq data revealed an approximately 1.5-fold enrichment of cluster 1 over cluster 2, suggesting a bias in compositional sampling in snRNA-seq. Therefore, such cell-level sampling biases must be considered when evaluating the composition of complex tissues with snRNA-seq.

Understanding the advantages and drawbacks of using snRNA-seq, a nuclear transcriptome is inherently enriched for nascent transcripts, thereby predominantly reflecting changes in gene expression as a result of differences in transcription rates alone (Gaidatzis et al. 2015). In contrast, a cellular transcriptome is fundamentally enriched for mature transcripts, thereby capturing gene expression changes driven by both transcriptional and

post-transcriptional regulatory processes such as mRNA processing and degradation. A higher relative proportion of nascent to mature transcripts in the nucleus also results in a large fraction of intronic reads in snRNA-seq, which, when considered for count matrix generation, gives rise to detection bias against short genes with few intronic poly(A) stretches. Consequently, for a biological system compatible with both techniques, scRNA-seq may be better for identifying cellular subpopulations. scRNA-seq will also be better for assessing gene expression changes because of post-transcriptional regulation. However, nuclear transcriptome is preferentially enriched for lncRNAs, indicating that functional investigations of these genes will be enhanced by sequencing nuclei. Moreover, some studies of specific nuclear functions may be enhanced by directly accessing nuclei, for example, changes in gene expression profile because of targeted transcriptional activation mediated by epigenetic modifications. Therefore, it is important to evaluate each approach depending on the task at hand. However, for tissues such as the adipose tissue, snRNA-seq may be the only option. In our investigation, lncRNAs regulating adipogenesis were enriched in the nuclear transcriptome. lncRNAs driving differences between cluster 1 and 2 in brown preadipocytes were also better detected in the snRNA-seq data set. However, we also identified poor detection of shorter genes in nuclei, some of which were key to driving heterogeneity between distinct cell types.

Including intronic reads for UMI quantification presents researchers with both advantages and drawbacks. poly(A) stretches are found randomly dispersed along the length of the genome, and introns become the predominant site for the localization of such stretches because of their extensive length (21-fold longer than exons) (Piovesan et al. 2016). These poly(A) stretches present additional priming sites (besides the 3' poly(A) tail) for the poly(T) RT primer, thereby enabling more efficient transcript capture. Conversely, most intronic reads are therefore derived from genes with multiple poly(A) stretches (long genes), thereby introducing technical detection bias. This bias gets further magnified in snRNA-seq libraries that are inherently enriched for nascent transcripts (and hence intronic reads), and filtering such reads would mean reduced gene detection sensitivity, shallower sequencing depth, and underused sequencing cost. Here, we provided a normalization strategy for UMI counts derived from intronic reads that can remove gene length-associated technical biases. Implementation of this normalization strategy removes technical artifacts while retaining true biological features, thereby improving integration and enabling joint analysis of the scRNA-seq and snRNA-seq data sets. In such joint analysis, our normalization strategy would also improve the accuracy of differential expression testing between any technique-specific clusters identified.

Finally, we showed applicability of scVI for integration of scRNA-seq and snRNA-seq data sets. This is critical for the generation of a comprehensive adipose tissue atlas because investigations into the SVF heterogeneity have been performed using scRNA-seq, whereas snRNA-seq is favorable for investigations into the existence of adipocyte subtypes. Therefore, any efforts to identify shared subpopulations across such data sets, and the lineages therein, would demand data integration. However, our findings here are based on an *in vitro* adipogenic model system, with a less heterogeneous cellular composition than primary tissue. Therefore, integration of the scRNA-seq and snRNA-seq data sets with this *in vitro* adipogenic model system is likely more robust than in a primary sample. Indeed, multiple previous reports have shown strong batch effects between scRNA-seq and snRNA-seq data sets derived from the same primary tissue, resulting in subop-

timal integration performance (Slyper et al. 2020; Andrews et al. 2021; Luecken et al. 2022). Such batch effects are likely rooted in technical differences across the two techniques, such as gene length-associated detection biases and high background mRNA levels in nuclear libraries (Fleming et al. 2019; Alvarez et al. 2020). Therefore, to truly generate a comprehensive human cell atlas across varying tissue types, there is a need for evolved algorithms specifically developed for the task of integrating scRNA-seq and snRNA-seq data sets.

Overall, snRNA-seq provides an effective method for characterizing cellular heterogeneity and functionally relevant gene expression profiles within human preadipocytes and adipocytes. We expect that snRNA-seq will be actively adopted by the adipose community for high-throughput transcriptomic profiling of the tissue and will aid in increasing its representation in initiatives such as the Human Cell Atlas. Ultimately, joint analysis of data sets acquired using multiple sequencing techniques will aid in the creation of a comprehensive human adipose tissue atlas, thereby enabling us to dissect its critical role in health and disease.

Methods

Preadipocyte culture and adipogenic differentiation

The detailed protocol for maintenance, cryopreservation, and differentiation of white and brown preadipocytes is outlined in a different study (Shamsi and Tseng 2017). Briefly, preadipocytes were maintained in DMEM (Corning 10-017-CV) supplemented with 10% (v/v) FBS and 1% vol/vol penicillin–streptomycin at 37°C with 5% CO₂. For differentiation, preadipocytes were maintained at 100% confluence for 48 h, after which growth media was replaced with differentiation media every 48 h for the next 20 d (see Supplemental Material).

Harvesting mature adipocyte for scRNA-seq cellenONE X1

Mature adipocytes were washed with PBS and incubated with a monolayer of 0.25% trypsin with 0.1% EDTA (Gibco 25200-056; monolayer obtained by adding and removing 1 mL of trypsin) for 2–3 min in a tissue culture incubator. When adipocytes started to become round and detached from the plate, trypsin was neutralized by adding 1 mL of FBS. Clumps of adipocytes were dislodged using a wide-bore 1-mL pipette tip and filtered using a 70- μ m cell strainer. Concentration of adipocyte suspension was adjusted to approximately 200 cells per microliter using FBS for downstream spotting using the cellenONE X1 machine.

Nuclei isolation from preadipocytes and mature adipocytes for snRNA-seq

Nuclei were isolated from white and brown preadipocytes using an NP-40-based lysis buffer (see Supplemental Material). Cells from a 100-mm dish were scraped onto 500 μ L of chilled lysis buffer, incubated for 5 min on ice, and washed with ice-cold PBS supplemented with 0.2 U/ μ L RNase inhibitor (Protector RNase inhibitor; henceforth called wash buffer) four times by centrifuging at 500 rcf for 5 min at 4°C. Nuclei were finally resuspended in the ice-cold wash buffer, filtered using a 40- μ m cell strainer, and adjusted for a final concentration of approximately 1000 nuclei per microliter. Nuclei were also stained using 0.08% trypan blue dye to assess nuclear membrane integrity under brightfield imaging. For nuclear isolation at the mature adipocyte stage, the same protocol was implemented as mentioned above with the modification of using 1 mL lysis buffer per 100-mm dish.

Table 1. Sequencing chemistries and metric for libraries used in our study

Technique	Sample	Chemistry	Depth (per cell)	No. of cells	Sequencer
scRNA-seq	White and brown preadipocytes	10x Chromium 3' v2 + hashtag antibodies	78,000	8,000	NovaSeq S2
scRNA-seq	White adipocytes	mcSCRb-seq	100,000	200	NextSeq
snRNA-seq	White preadipocytes	10x Chromium 3' v3	174,738	8,000	NovaSeq S2
snRNA-seq	Brown preadipocytes	10x Chromium 3' v3	216,700	7,000	NovaSeq S2
snRNA-seq	White adipocytes	10x Chromium 3' v3	123,700	12,000	NovaSeq S2

For more details, see the [Supplemental Material](#).

scRNA-seq and snRNA-seq library preparation

Sequencing chemistries and metrics for libraries used in our study are listed in Table 1. See [Supplemental Methods](#) for more details. Also see [Supplemental Note S1](#) for cell-hashing strategy used for white and brown preadipocytes scRNA-seq data set (Table 1, row 1).

scRNA-seq and snRNA-seq data analysis

All 10x Genomics libraries were mapped and quantified using Cell Ranger. UMAP visualization for individual libraries was performed using latent space derived from scVI (Lopez et al. 2018) as input. Downstream QC to remove low-quality cells and clusters was performed using Seurat (Butler et al. 2018) in R (R Core Team 2020). Similar steps were performed for analysis of mcSCRb-seq libraries with data processing using zUMIs (Parekh et al. 2018). Vision was used to perform all transcriptional signature analysis. Data integration was performed using scVI-tools (Gayoso et al. 2021) and Seurat (Stuart et al. 2019). For further details, see the [Supplemental Material](#).

RNA smFISH and spot counting analysis

To perform RNA FISH, we followed the protocol of Raj et al. (2008) with minor modifications. We prewashed cells with wash buffer containing 10% formamide and 2× saline-sodium citrate (SSC). We then performed hybridization by adding 1 μL of 6.25 μM probe to 50 μL of hybridization buffer consisting of 10% formamide, 2× SSC, and 10% (w/v) dextran sulfate. The final probe concentration for overnight hybridization was 125 nM. We hybridized the samples overnight in a humidified chamber at 37°C. Following hybridization, we washed the samples twice with wash buffer for 30 min at 37°C. We then washed the samples 2× SSC, antifade buffer. Imaging was performed in antifade buffer supplemented with catalase and glucose oxidase. Quantification of RNA spots per cell was performed using the Find Foci tool (Herbert et al. 2014) in Fiji (see [Supplemental Material](#)).

Data access

The sequence data from this study have been submitted to the NCBI database of Genotypes and Phenotypes (dbGaP; <http://www.ncbi.nlm.nih.gov/dbgap>) under accession number phs002461.v1.p1. The scripts to reproduce the work presented in this study can be found as [Supplemental Code](#).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

This publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM124916. This publication was also supported by grant number CZF2019-002454 from the Chan Zuckerberg Foundation and grant numbers R01DK102898 and K01DK125608 from the National Institutes of Health. A.S. is a Chan-Zuckerberg Biohub Investigator and a Pew Scholar in the Biomedical Sciences, supported by the Pew Charitable Trusts. A.G. is supported by the University of California at Berkeley Lloyd Fellowship in Bioengineering. We thank Joshua Cantlon from Scienion for lending help with the CellenOne instrument, the Gartner laboratory at University of California at San Francisco for providing us with multi-seq probes, and the Groves laboratory at University of California at Berkeley for access to the smFISH microscope.

References

- Acosta JR, Joost S, Karlsson K, Ehlund A, Li X, Aouadi M, Kasper M, Arner P, Rydén M, Laurencikiene J. 2017. Single cell transcriptomics suggest that human adipocyte progenitor cells constitute a homogeneous cell population. *Stem Cell Res Ther* **8**: 250. doi:10.1186/s13287-017-0701-4
- Alvarez M, Rahmani E, Jew B, Garske KM, Miao Z, Benhammou JN, Ye CJ, Pisegna JR, Pietiläinen KH, Halperin E, et al. 2020. Enhancing droplet-based single-nucleus RNA-seq resolution using the semi-supervised machine learning classifier DIEM. *Sci Rep* **10**: 11019. doi:10.1038/s41598-020-67513-5
- Ameur A, Zaghlool A, Halvardson J, Wetterbom A, Gyllensten U, Cavellier L, Feuk L. 2011. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat Struct Mol Biol* **18**: 1435–1440. doi:10.1038/nsmb.2143
- Andrews TS, Atif J, Liu JC, Perciani CT, Ma XZ, Thoeni C, Slyper M, Eraslan G, Segerstolpe A, Manuel J, et al. 2021. Single-cell, single-nucleus, and spatial RNA sequencing of the human liver identifies cholangiocyte and mesenchymal heterogeneity. *Hepatology* doi:10.1002/hep4.1854
- Bagnoli JW, Ziegenhain C, Janjic A, Wange LE, Vieth B, Parekh S, Geuder J, Hellmann I, Enard W. 2018. Sensitive and powerful single-cell RNA sequencing using mcSCRb-seq. *Nat Commun* **9**: 1–8. doi:10.1038/s41467-018-05347-6
- Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN, Aevermann B, Barkan E, Bertagnolli D, Casper T, Dee N, et al. 2018. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* **13**: e0209648. doi:10.1371/journal.pone.0209648
- Berry R, Rodeheffer MS. 2013. Characterization of the adipocyte cellular lineage in vivo. *Nat Cell Biol* **15**: 302–308. doi:10.1038/ncb2696
- Birnbaum KD. 2018. Power in numbers: single-cell RNA-seq strategies to dissect complex tissues. *Annu Rev Genet* **52**: 203–221. doi:10.1146/annurev-genet-120417-031247
- Braga M, Reddy ST, Vergnes L, Pervin S, Grijalva V, Stout D, David J, Li X, Tomasian V, Reid CB, et al. 2014. Follistatin promotes adipocyte differentiation, browning, and energy metabolism. *J Lipid Res* **55**: 375–384. doi:10.1194/jlr.M039719
- Burl RB, Ramseyer VD, Rondini EA, Pique-Regi R, Lee YH, Granneman JG. 2018. Deconstructing adipogenesis induced by β3-adrenergic receptor activation with single-cell expression profiling. *Cell Metab* **28**: 300–309.e4. doi:10.1016/j.cmet.2018.05.025

- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**: 411–420. doi:10.1038/nbt.4096
- Cabili MN, Dumagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A. 2015. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol* **16**: 20. doi:10.1186/s13059-015-0586-4
- Carobbio S, Rodriguez-Cuenca S, Vidal-Puig A. 2011. Origins of metabolic complications in obesity: ectopic fat accumulation, the importance of the qualitative aspect of lipotoxicity. *Curr Opin Clin Nutr Metab Care* **14**: 520–526. doi:10.1097/MCO.0b013e32834ad966
- Cawthorn WP, Scheller EL, MacDougald OA. 2012. Adipose tissue stem cells meet preadipocyte commitment: going back to the future. *J Lipid Res* **53**: 227–246. doi:10.1194/jlr.R021089
- Chen T, van Steensel B. 2017. Comprehensive analysis of nucleocytoplasmic dynamics of mRNA in drosophila cells. *PLoS Genet* **13**: e1006929. doi:10.1371/journal.pgen.1006929
- Chen X, Teichmann SA, Meyer KB. 2018. From tissues to cell types and back: single-cell gene expression analysis of tissue architecture. *Annu Rev Biomed Data Sci* **1**: 29–51. doi:10.1146/annurev-biodatasci-080917-013452
- Cho DS, Lee B, Doles JD. 2019. Refining the adipose progenitor cell landscape in healthy and obese visceral adipose tissue using single-cell gene expression profiling. *Life Sci Alliance* **2**: e201900561. doi:10.26508/lsa.201900561
- de Jong JMA, Larsson O, Cannon B, Nedergaard J. 2015. A stringent validation of mouse adipose tissue identity markers. *Am J Physiol Endocrinol Metab* **308**: E1085–E1105. doi:10.1152/ajpendo.00023.2015
- Delá MJ, Jackson BT, Erard N, Knott SR V, Hannon GJ, Kovacevic T, Vangelisti S, Maravilla EM, Wild SA, Stork EM. 2019. lncRNA *Spehd* regulates hematopoietic stem and progenitor cells and is required for multilineage differentiation. *Cell Rep* **27**: 719–729. doi:10.1016/j.celrep.2019.03.080
- DeTomaso D, Yosef N. 2016. Fastproject: a tool for low-dimensional analysis of single-cell RNA-seq data. *BMC Bioinformatics* **17**: 315. doi:10.1186/s12859-016-1176-5
- DeTomaso D, Jones MG, Subramaniam M, Ashuach T, Ye CJ, Yosef N. 2019. Functional interpretation of single cell similarity maps. *Nat Commun* **10**: 4376. doi:10.1038/s41467-019-12235-0
- Deutsch A, Feng D, Pessin JE, Shinoda K. 2020. The impact of single-cell genomics on adipose tissue research. *Int J Mol Sci* **21**: 4773. doi:10.3390/ijms21134773
- Ding C, Lim YC, Chia SY, Walet ACE, Xu S, Lo KA, Zhao Y, Zhu D, Shan Z, Chen Q, et al. 2018. De novo reconstruction of human adipose transcriptome reveals conserved lncRNAs as regulators of brown adipogenesis. *Nat Commun* **9**: 1329. doi:10.1038/s41467-018-03754-3
- Dupont S, Morsut L, Aragona M, Enzo E, Giulitti S, Cordenonsi M, Zanconato F, Le Digabel J, Forcato M, Bicciato S, et al. 2011. Role of YAP/TAZ in mechanotransduction. *Nature* **474**: 179–183. doi:10.1038/nature10137
- Ferrero R, Rainer P, Deplancke B. 2020. Toward a consensus view of mammalian adipocyte stem and progenitor cell heterogeneity. *Trends Cell Biol* **30**: 937–950. doi:10.1016/j.tcb.2020.09.007
- Fleming S, Marioni J, Babadi M. 2019. Cellbender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. bioRxiv doi:10.1101/791699
- Gaidatzis D, Burger L, Florescu M, Stadler MB. 2015. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nat Biotechnol* **33**: 722–729. doi:10.1038/nbt.3269
- Gao H, Volat F, Sandhow L, Galitzky J, Nguyen T, Esteve D, Åström G, Mejhert N, Ledoux S, Thalamos C, et al. 2017. CD36 is a marker of human adipocyte progenitors with pronounced adipogenic and triglyceride accumulation potential. *Stem Cells* **35**: 1799–1814. doi:10.1002/stem.2635
- Gaublomme JT, Li B, McCabe C, Knecht A, Yang Y, Drokhylyansky E, Van Wittenbergh N, Waldman J, Dionne D, Nguyen L, et al. 2019. Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat Commun* **10**: 2907. doi:10.1038/s41467-019-10756-2
- Gayoso A, Lopez R, Xing G, Boyeau P, Wu K, Jayasuriya M, Melhman E, Langevin M, Liu Y, Samaran J, et al. 2021. scvi-tools: a library for deep probabilistic analysis of single-cell omics data. bioRxiv doi:10.1101/2021.04.28.441833
- Gerin I, Bommer GT, Lidell ME, Cederberg A, Enerback S, MacDougald OA. 2009. On the role of FOX transcription factors in adipocyte differentiation and insulin-stimulated glucose uptake. *J Biol Chem* **284**: 10755–10763. doi:10.1074/jbc.M809115200
- Ghaben AL, Scherer PE. 2019. Adipogenesis and metabolic health. *Nat Rev Mol Cell Biol* **20**: 242–258. doi:10.1038/s41580-018-0093-z
- Giordano A, Coppari R, Castellucci M, Cinti S. 2001. Sema3a is produced by brown adipocytes and its secretion is reduced following cold acclimation. *J Neurocytol* **30**: 5–10. doi:10.1023/a:1011916822633
- Gleason CE, Ning Y, Cominski TP, Gupta R, Kaestner KH, Pintar JE, Birnbaum MJ. 2010. Role of insulin-like growth factor-binding protein 5 (IGFBP5) in organismal and pancreatic β -cell growth. *Mol Endocrinol* **24**: 178–192. doi:10.1210/me.2009-0167
- Gray JM, Harmin DA, Boswell SA, Cloonan N, Mullen TE, Ling JJ, Miller N, Kuersten S, Ma Y-C, McCarroll SA, et al. 2014. SnapShot-seq: a method for extracting genome-wide, in vivo mRNA dynamics from a single total RNA sample. *PLoS One* **9**: e89673. doi:10.1371/journal.pone.0089673
- Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O’Shaughnessy AL, Lambert GM, Araúzo-Bravo MJ, Lee J, Fishman M, Robbins GE, et al. 2013. RNA-sequencing from single nuclei. *Proc Natl Acad Sci* **110**: 19802–19807. doi:10.1073/pnas.1319700110
- Gu W, Nowak WN, Xie Y, Le Bras A, Hu Y, Deng J, Issa Bhaloo S, Lu Y, Yuan H, Fidanis E, et al. 2019. Single-cell RNA-sequencing and metabolomics analyses reveal the contribution of perivascular adipose tissue stem cells to vascular remodeling. *Arterioscler Thromb Vasc Biol* **39**: 2049–2066. doi:10.1161/ATVBAHA.119.312732
- Gubelmann C, Schwalie PC, Raghav SK, Röder E, Delessa T, Kiehlmann E, Waszak SM, Corsinotti A, Udin G, Holcombe W, et al. 2014. Identification of the transcription factor ZEB1 as a central component of the adipogenic gene regulatory network. *eLife* **3**: e03346. doi:10.7554/eLife.03346
- Gupta A, Dorlhiac GF, Streets AM. 2019. Quantitative imaging of lipid droplets in single cells. *Analyst* **144**: 753–765. doi:10.1039/C8AN01525B
- Habib N, Avraham-Davidi I, Basu A, Burks T, Shekhar K, Hofree M, Choudhury SR, Aguet F, Gelfand E, Ardlie K, et al. 2017. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* **14**: 955–958. doi:10.1038/nmeth.4407
- Hagberg CE, Li Q, Kutschke M, Bhowmick D, Kiss E, Shabalina IG, Harms MJ, Shilkova O, Kozina V, Nedergaard J, et al. 2018. Flow cytometry of mouse and human adipocytes for the analysis of browning and cellular heterogeneity. *Cell Rep* **24**: 2746–2756.e5. doi:10.1016/j.celrep.2018.08.006
- Hansen CG, Moroishi T, Guan KL. 2015. YAP and TAZ: a nexus for Hippo signaling and beyond. *Trends Cell Biol* **25**: 499–513. doi:10.1016/j.tcb.2015.05.002
- Hendriks GJ, Gaidatzis D, Aeschmann F, Großhans H. 2014. Extensive oscillatory gene expression during *C. elegans* larval development. *Mol Cell* **53**: 380–392. doi:10.1016/j.molcel.2013.12.013
- Henriques F, Bedard AH, Guilherme A, Kelly M, Chi J, Zhang P, Lifshitz LM, Bellvé K, Rowland LA, Yenilmez B, et al. 2020. Single-cell RNA profiling reveals adipocyte to macrophage signaling sufficient to enhance thermogenesis. *Cell Rep* **32**: 107998. doi:10.1016/j.celrep.2020.107998
- Hepler C, Shan B, Zhang Q, Henry GH, Shao M, Vishvanath L, Ghaben AL, Mobley AB, Strand D, Hon GC, et al. 2018. Identification of functionally distinct fibro-inflammatory and adipogenic stromal subpopulations in visceral adipose tissue of adult mice. *eLife* **7**: e39636. doi:10.7554/eLife.39636
- Herbert AD, Carr AM, Hoffmann E, Lichten M. 2014. Findfoci: a focus detection algorithm with automated parameter training that closely matches human assignments, reduces human inconsistencies and increases speed of analysis. *PLoS One* **9**: e114749. doi:10.1371/journal.pone.0114749
- Hou S, Jiao Y, Yuan Q, Zhai J, Tian T, Sun K, Chen Z, Wu Z, Zhang J. 2018. S100a4 protects mice from high-fat diet-induced obesity and inflammation. *Lab Invest* **98**: 1025–1038. doi:10.1038/s41374-018-0067-y
- Huang Y, Jin C, Zheng Y, Li X, Zhang S, Zhang Y, Jia L, Li W. 2017. Knockdown of lncRNA MIR31HG inhibits adipocyte differentiation of human adipose-derived stem cells via histone modification of FABP4. *Sci Rep* **7**: 8080. doi:10.1038/s41598-017-08131-6
- Jaitin DA, Adlung L, Thaiss CA, Weiner A, Li B, Descamps H, Lundgren P, Blieriot C, Liu Z, Deczkowska A, et al. 2019. Lipid-associated macrophages control metabolic homeostasis in a Trem2-dependent manner. *Cell* **178**: 686–698.e14. doi:10.1016/j.cell.2019.05.054
- Jo J, Gavrilova O, Pack S, Jou W, Mullen S, Sumner AE, Cushman SW, Periwai V. 2009. Hypertrophy and/or hyperplasia: dynamics of adipose tissue growth. *PLoS Comput Biol* **5**: e1000324. doi:10.1371/journal.pcbi.1000324
- Karbiener M, Glantschnig C, Pisani DF, Laurencikiene J, Dahlman I, Herzog S, Amri EZ, Scheideler M. 2015. Mesoderm-specific transcript (MEST) is a negative regulator of human adipocyte differentiation. *Int J Obes* **39**: 1733–1741. doi:10.1038/ijo.2015.121
- Kriszt R, Arai S, Itoh H, Lee MH, Goralczyk AG, Xia XM, Cypess AM, White AP, Shamsi F, Xue R, et al. 2017. Optical visualisation of thermogenesis in stimulated single-cell brown adipocytes. *Sci Rep* **7**: 1383. doi:10.1038/s41598-017-00291-9
- Lake BB, Codeluppi S, Yung YC, Gao D, Chun J, Kharchenko PV, Linnarsson S, Zhang K. 2017. A comparative strategy for single-nucleus and single-

- cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci Rep* **7**: 6031. doi:10.1038/s41598-017-04426-w
- La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastrioti ME, Lönnerberg P, Furlan A, et al. 2018. RNA velocity of single cells. *Nature* **560**: 494–498. doi:10.1038/s41586-018-0414-6
- Lee KY, Luong Q, Sharma R, Dreyfuss JM, Ussar S, Kahn CR. 2019. Developmental and functional heterogeneity of white adipocytes within a single fat depot. *EMBO J* **38**: e99291. doi:10.15252/embj.201899291
- Levelt E, Pavlides M, Banerjee R, Mahmood M, Kelly C, Sellwood J, Ariga R, Thomas S, Francis J, Rodgers C, et al. 2016. Ectopic and visceral fat deposition in lean and obese patients with type 2 diabetes. *J Am Coll Cardiol* **68**: 53–63. doi:10.1016/j.jacc.2016.03.597
- Li SN, Wu JF. 2020. TGF- β /SMAD signaling regulation of mesenchymal stem cells in adipocyte commitment. *Stem Cell Res Ther* **11**: 41. doi:10.1186/s13287-020-1552-y
- Li Z, Jin C, Chen S, Zheng Y, Huang Y, Jia L, Ge W, Zhou Y. 2017. Long noncoding RNA MEG3 inhibits adipogenesis and promotes osteogenesis of human adipose-derived mesenchymal stem cells via miR-140-5p. *Mol Cell Biochem* **433**: 51–60. doi:10.1007/s11010-017-3015-z
- Lidell ME, Betz MJ, Leinhard OD, Heglund M, Elander L, Slawik M, Mussack T, Nilsson D, Romu T, Nuutila P, et al. 2013. Evidence for two types of brown adipose tissue in humans. *Nat Med* **19**: 631–634. doi:10.1038/nm.3017
- Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. 2018. Deep generative modeling for single-cell transcriptomics. *Nat Methods* **15**: 1053–1058. doi:10.1038/s41592-018-0229-2
- Luecken MD, Büttner M, Chaochompou K, Danese A, Interlandi M, Mueller MF, Strobl DC, Zappia L, Dugas M, Colomé-Tatché M, et al. 2022. Benchmarking atlas-level data integration in single-cell genomics. *Nat Methods* **19**: 41–50. doi:10.1038/s41592-021-01336-8
- Luther J, Driessler F, Megges M, Hess A, Herbort B, Mandic V, Zaiss MM, Reichardt A, Zech C, Tuckermann JP, et al. 2011. Elevated Fra-1 expression causes severe lipodystrophy. *J Cell Sci* **124**: 1465–1476. doi:10.1242/jcs.079855
- Luther J, Ubieta K, Hannemann N, Jimenez M, Garcia M, Zech C, Schett G, Wagner EF, Bozec A. 2014. Fra-2/AP-1 controls adipocyte differentiation and survival by regulating PPAR γ and hypoxia. *Cell Death Differ* **21**: 655–664. doi:10.1038/cdd.2013.198
- Majka SM, Miller HL, Helm KM, Acosta AS, Childs CR, Kong R, Klemm DJ. 2014. Analysis and isolation of adipocytes by flow cytometry. *Meth Enzymol* **537**, 281–296. doi:10.1016/B978-0-12-411619-1.00015-X
- Mereu E, Lafzi A, Moutinho C, Ziegenhain C, McCarthy DJ, Álvarez-Varela A, Battle E, Sagar, Grün D, Lau JK, et al. 2020. Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat Biotechnol* **38**: 747–755. doi:10.1038/s41587-020-0469-4
- Merrick D, Sakers A, Irgebay Z, Okada C, Calvert C, Morley MP, Percec I, Seale P. 2019. Identification of a mesenchymal progenitor cell hierarchy in adipose tissue. *Science* **364**: eaav2501. doi:10.1126/science.aav2501
- Mildmay-White A, Khan W. 2017. Cell surface markers on adipose-derived stem cells: a systematic review. *Curr Stem Cell Res Ther* **12**: 484–492. doi:10.2174/1574888X11666160429122133
- Min SY, Kady J, Nam M, Rojas-Rodriguez R, Berkenwald A, Kim JH, Noh HL, Kim JK, Cooper MP, Fitzgibbons T, et al. 2016. Human 'brite/beige' adipocytes develop from capillary networks, and their implantation improves metabolic homeostasis in mice. *Nat Med* **22**: 312–318. doi:10.1038/nm.4031
- Modica S, Wolfrum C. 2017. The dual role of BMP4 in adipogenesis and metabolism. *Adipocyte* **6**: 141–146. doi:10.1080/21623945.2017.1287637
- Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. 2018. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience* **7**: giy059. doi:10.1093/gigascience/giy059
- Patil M, Sharma BK, Satyanarayana A. 2014. Id transcriptional regulators in adipogenesis and adipose tissue metabolism. *Front Biosci (Landmark Ed)* **19**: 1386–1397. doi:10.2741/4289
- Patrick R, Humphreys DT, Janbandhu V, Oshlack A, Ho JW, Harvey RP, Lo KK. 2020. Sierra: discovery of differential transcript usage from polyA-captured single-cell RNA-seq data. *Genome Biol* **21**: 167. doi:10.1186/s13059-020-02071-7
- Pfeifer A, Hoffmann LS. 2015. Brown, beige, and white: the new color code of fat and its pharmacological implications. *Annu Rev Pharmacol Toxicol* **55**: 207–227. doi:10.1146/annurev-pharmtox-010814-124346
- Piovesan A, Caracausi M, Antonaros F, Pelleri MC, Vitale L. 2016. Genebase 1.1: a tool to summarize data from NCBI gene datasets and its application to an update of human gene statistics. *Database* **2016**: baw153. doi:10.1093/database/baw153
- Quinodoz S, Guttman M. 2014. Long noncoding RNAs: an emerging link between gene regulation and nuclear organization. *Trends Cell Biol* **24**: 651–663. doi:10.1016/j.tcb.2014.08.009
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**: 877–879. doi:10.1038/nmeth.1253
- Rajbhandari P, Arneson D, Hart SK, Ahn IS, Diamante G, Santos LC, Zaghari N, Feng AC, Thomas BJ, Vergnes L, et al. 2019. Single cell analysis reveals immune cell-adipocyte crosstalk regulating the transcription of thermogenic adipocytes. *eLife* **8**: e49501. doi:10.7554/eLife.49501
- R Core Team. 2020. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.
- Rondini EA, Granneman JG. 2020. Single cell approaches to address adipose tissue stromal cell heterogeneity. *Biochem J* **477**: 583–600. doi:10.1042/BCJ20190467
- Samata M, Akhtar A. 2018. Dosage compensation of the X chromosome: a complex epigenetic assignment involving chromatin regulators and long noncoding RNAs. *Annu Rev Biochem* **87**: 323–350. doi:10.1146/annurev-biochem-062917-011816
- Sanchez-Gurmaches J, Hung CM, Guertin DA. 2016. Emerging complexities in adipocyte origins and identity. *Trends Cell Biol* **26**: 313–326. doi:10.1016/j.tcb.2016.01.004
- Sárvári AK, Van Hauwaert EL, Markussen LK, Gammelmark E, Marcher AB, Ebbesen MF, Nielsen R, Brewer JR, Madsen JGS, Mandrup S. 2021. Plasticity of epididymal adipose tissue in response to diet-induced obesity at single-nucleus resolution. *Cell Metab* **33**: 437–453.e5. doi:10.1016/j.cmet.2020.12.004
- Satyanarayana A, Klarmann KD, Gavrilova O, Keller JR. 2012. Ablation of the transcriptional regulator *Id1* enhances energy expenditure, increases insulin sensitivity, and protects against age and diet induced insulin resistance, and hepatosteatosis. *FASEB J* **26**: 309–323. doi:10.1096/fj.11-190892
- Schwalie PC, Dong H, Zachara M, Russeil J, Alpern D, Akchiche N, Caprara C, Sun W, Schlaudraff KU, Soldati G, 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
- Selewa A, Dohn R, Eckart H, Lozano S, Xie B, Gauchat E, Elorbany R, Rhodes K, Burnett J, Gilad Y, et al. 2020. Systematic comparison of high-throughput single-cell and single-nucleus transcriptomes during cardiomyocyte differentiation. *Sci Rep* **10**: 1–13. doi:10.1038/s41598-020-58327-6
- Shamsi F, Tseng YH. 2017. Protocols for generation of immortalized human brown and white preadipocyte cell lines. *Methods Mol Biol* **1566**, 77–85. doi:10.1007/978-1-4939-6820-6_8
- Shamsi F, Piper M, Ho LL, Huang TL, Gupta A, Streets A, Lynes MD, Tseng YH. 2021. Vascular smooth muscle-derived Trpv1+ progenitors are a source of cold-induced thermogenic adipocytes. *Nat Metab* **3**: 485–495. doi:10.1038/s42255-021-00373-z
- Sherstyuk VV, Medvedev SP, Zakian SM. 2018. Noncoding RNAs in the regulation of pluripotency and reprogramming. *Stem Cell Rev Reports* **14**: 58–70. doi:10.1007/s12015-017-9782-9
- Shinoda K, Luijten IHN, Hasegawa Y, Hong H, Sonne SB, Kim M, Xue R, Chondronikola M, Cypess AM, Tseng YH, et al. 2015. Genetic and functional characterization of clonally derived adult human brown adipocytes. *Nat Med* **21**: 389–394. doi:10.1038/nm.3819
- Shulman ED, Elkon R. 2019. Cell-type-specific analysis of alternative polyadenylation using single-cell transcriptomics data. *Nucleic Acids Res* **47**: 10027–10039. doi:10.1093/nar/gkz781
- Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhllyansky E, Wakiro I, Smillie C, Smith-Rosario G, Wu J, Dionne D, et al. 2020. A single-cell and single-nucleus RNA-seq toolbox for fresh and frozen human tumors. *Nat Med* **26**: 792–802. doi:10.1038/s41591-020-0844-1
- Song A, Dai W, Jang MJ, Medrano L, Li Z, Zhao H, Shao M, Tan J, Li A, Ning T, et al. 2020. Low- and high-thermogenic brown adipocyte subpopulations coexist in murine adipose tissue. *J Clin Invest* **130**: 247–257. doi:10.1172/JCI129167
- Spaethling JM, Sanchez-Alavez M, Lee JH, Xia FC, Dueck H, Wang W, Fisher SA, Sul JY, Seale P, Kim J, et al. 2016. Single-cell transcriptomics and functional target validation of brown adipocytes show their complex roles in metabolic homeostasis. *FASEB J* **30**: 81–92. doi:10.1096/fj.15-273797
- Strutz F, Okada H, Lo CW, Danoff T, Carone RL, Tomaszewski JE, Neilson EG. 1995. Identification and characterization of a fibroblast marker: FSP1. *J Cell Biol* **130**: 393–405. doi:10.1083/jcb.130.2.393
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive integration of single-cell data. *Cell* **177**: 1888–1902.e21. doi:10.1016/j.cell.2019.05.031
- Sun L, Lin JD. 2019. Function and mechanism of long noncoding RNAs in adipocyte biology. *Diabetes* **68**: 887–896. doi:10.2337/dbi18-0009
- Sun L, Goff LA, Trapnell C, Alexander R, Lo KA, Hacsulleyman E, Sauvageau M, Tazon-Vega B, Kelley DR, Hendrickson DG, et al. 2013. Long

- noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci* **110**: 3387–3392. doi:10.1073/pnas.1222643110
- Sun W, Dong H, Balaz M, Slyper M, Drokhyansky E, Colletuori G, Giordano A, Kovanicova Z, Stefanicka P, Balazova L, et al. 2020. snRNA-seq reveals a subpopulation of adipocytes that regulates thermogenesis. *Nature* **587**: 98–102. doi:10.1038/s41586-020-2856-x
- The Tabula Muris Consortium. 2018. Single-cell transcriptomics of 20 mouse organs creates a *Tabula muris*. *Nature* **562**: 367–372. doi:10.1038/s41586-018-0590-4
- Tews D, Schwar V, Scheithauer M, Weber T, Fromme T, Klingenspor M, Barth TF, Möller P, Holzmann K, Debatin KM, et al. 2014. Comparative gene array analysis of progenitor cells from human paired deep neck and subcutaneous adipose tissue. *Mol Cell Endocrinol* **395**: 41–50. doi:10.1016/j.mce.2014.07.011
- Tharp KM, Kang MS, Timblin GA, Dempersmier J, Dempsey GE, Zushin PJH, Benavides J, Choi C, Li CX, Jha AK, et al. 2018. Actomyosin-mediated tension orchestrates uncoupled respiration in adipose tissues. *Cell Metab* **27**: 602–615.e4. doi:10.1016/j.cmet.2018.02.005
- Thrupp N, Frigerio CS, Wolfs L, Skene NG, Fattorelli N, Poovathingal S, Fourné Y, Matthews PM, Theys T, Mancuso R, et al. 2020. Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. *Cell Rep* **32**: 108189. doi:10.1016/j.celrep.2020.108189
- Trapnell C. 2015. Defining cell types and states with single-cell genomics. *Genome Res* **25**: 1491–1498. doi:10.1101/gr.190595.115
- Vijay J, Gauthier MF, Biswell RL, Louiselle DA, Johnston JJ, Cheung WA, Belden B, Pramatarova A, Biertho L, Gibson M, et al. 2020. Single-cell analysis of human adipose tissue identifies depot- and disease-specific cell types. *Nat Metab* **2**: 97–109. doi:10.1038/s42255-019-0152-6
- Villarroya F, Vidal-Puig A. 2013. Beyond the sympathetic tone: the new brown fat activators. *Cell Metab* **17**: 638–643. doi:10.1016/j.cmet.2013.02.020
- Wei S, Du M, Jiang Z, Hausman GJ, Zhang L, Dodson MV. 2016. Long non-coding RNAs in regulating adipogenesis: new RNAs shed lights on obesity. *Cell Mol Life Sci* **73**: 2079–2087. doi:10.1007/s00018-016-2169-2
- Weiner FR, Shah A, Smith PJ, Rubin CS, Zern MA. 1989. Regulation of collagen gene expression in 3T3-L1 cells. effects of adipocyte differentiation and tumor necrosis factor α . *Biochemistry* **28**: 4094–4099. doi:10.1021/bi00435a070
- Wen X, Gao L, Guo X, Li X, Huang X, Wang Y, Xu H, He R, Jia C, Liang F. 2018. lncSLdb: a resource for long non-coding RNA subcellular localization. *Database* **2018**: 85.
- Wickramasinghe VO, Andrews R, Ellis P, Langford C, Gurdon JB, Stewart M, Venkitaraman AR, Laskey RA. 2014. Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Res* **42**: 5059–5071. doi:10.1093/nar/gku095
- Wu H, Kirita Y, Donnelly EL, Humphreys BD. 2019. Advantages of single-nucleus over single-cell RNA sequencing of adult kidney: rare cell types and novel cell states revealed in fibrosis. *J Am Soc Nephrol* **30**: 23–32. doi:10.1681/ASN.2018090912
- Xue R, Lynes MD, Dreyfuss JM, Shamsi F, Schulz TJ, Zhang H, Huang TL, Townsend KL, 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. doi:10.1038/nm.3881
- Yu FX, Zhao B, Panupinthu N, Jewell JL, Lian I, Wang LH, Zhao J, Yuan H, Tumaneng K, Li H, et al. 2012. Regulation of the Hippo-YAP pathway by G-protein-coupled receptor signaling. *Cell* **150**: 780–791. doi:10.1016/j.cell.2012.06.037
- Zappia L, Phipson B, Oshlack A. 2018. Exploring the single-cell RNA-seq analysis landscape with the scRNA-tools database. *PLoS Comput Biol* **14**: e1006245. doi:10.1371/journal.pcbi.1006245
- Zeng W, Jiang S, Kong X, El-Ali N, Ball AR, Ma CIH, Hashimoto N, Yokomori K, Mortazavi A. 2016. Single-nucleus RNA-seq of differentiating human myoblasts reveals the extent of fate heterogeneity. *Nucleic Acids Res* **44**: e158. doi:10.1093/nar/gkw739
- Zhang R, Gao Y, Zhao X, Gao M, Wu Y, Han Y, Qiao Y, Luo Z, Yang L, Chen J, et al. 2018. FSP1-positive fibroblasts are adipogenic niche and regulate adipose homeostasis. *PLoS Biol* **16**: e2001493. doi:10.1371/journal.pbio.2001493
- Zhang L, Zhang D, Qin ZY, Li J, Shen ZY. 2020. The role and possible mechanism of long noncoding RNA PVT1 in modulating 3T3-L1 preadipocyte proliferation and differentiation. *IUBMB Life* **72**: 1460–1467. doi:10.1002/iub.2269
- Zhou Q, Wan Q, Jiang Y, Liu J, Qiang L, Sun L. 2020. A landscape of murine long non-coding RNAs reveals the leading transcriptome alterations in adipose tissue during aging. *Cell Rep* **31**: 107694. doi:10.1016/j.celrep.2020.107694
- Zimta AA, Tigu AB, Braicu C, Stefan C, Ionescu C, Berindan-Neagoe I. 2020. An emerging class of long non-coding RNA with oncogenic role arises from the snoRNA host genes. *Front Oncol* **10**: 389. doi:10.3389/fonc.2020.00389

Received March 24, 2021; accepted in revised form December 10, 2021.