Cell Genomics, Volume *3*

Supplemental information

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

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

Supplemental Figures

Figure S1: LipocyteProfiler allows to characterize lipid droplet formation and expansion, related to Figure 1.

(a) LipocyteProfiler extracts 3,005 features that map to four channels and use four measurement types to quantify morphological and cellular changes.

(b) MI distribution across features (=nodes) in visceral (top panels) and subcutaneous (bottom panels) AMSCs applied on LipocyteProfiler total feature (left figures) and LipocyteProfiler core feature set (right figures).

(c) Sankey diagram of LipocyteProfiler core feature sets for visceral AMSCs (top panel) and subcutaneous AMSCs (bottom panel).

(d) Representative microscope images of actin-cytoskeleton remodeling during adipocyte differentiation. (left to right day 0, 7, 14; red = Phalloidin, actin-cytoskeleton, blue = Hoechst, DNA). scale bar = 52.8μm

(e) *Cytoplasm_RadialDistribution_RadialCV_AGP* measures (ring 1 to 4) are reduced in CRISPR/Cas9-mediated KO of *INSR* and *IRS-1*. Data is shown for two guides targeting *INSR or IRS-1*. (g1 and g2).

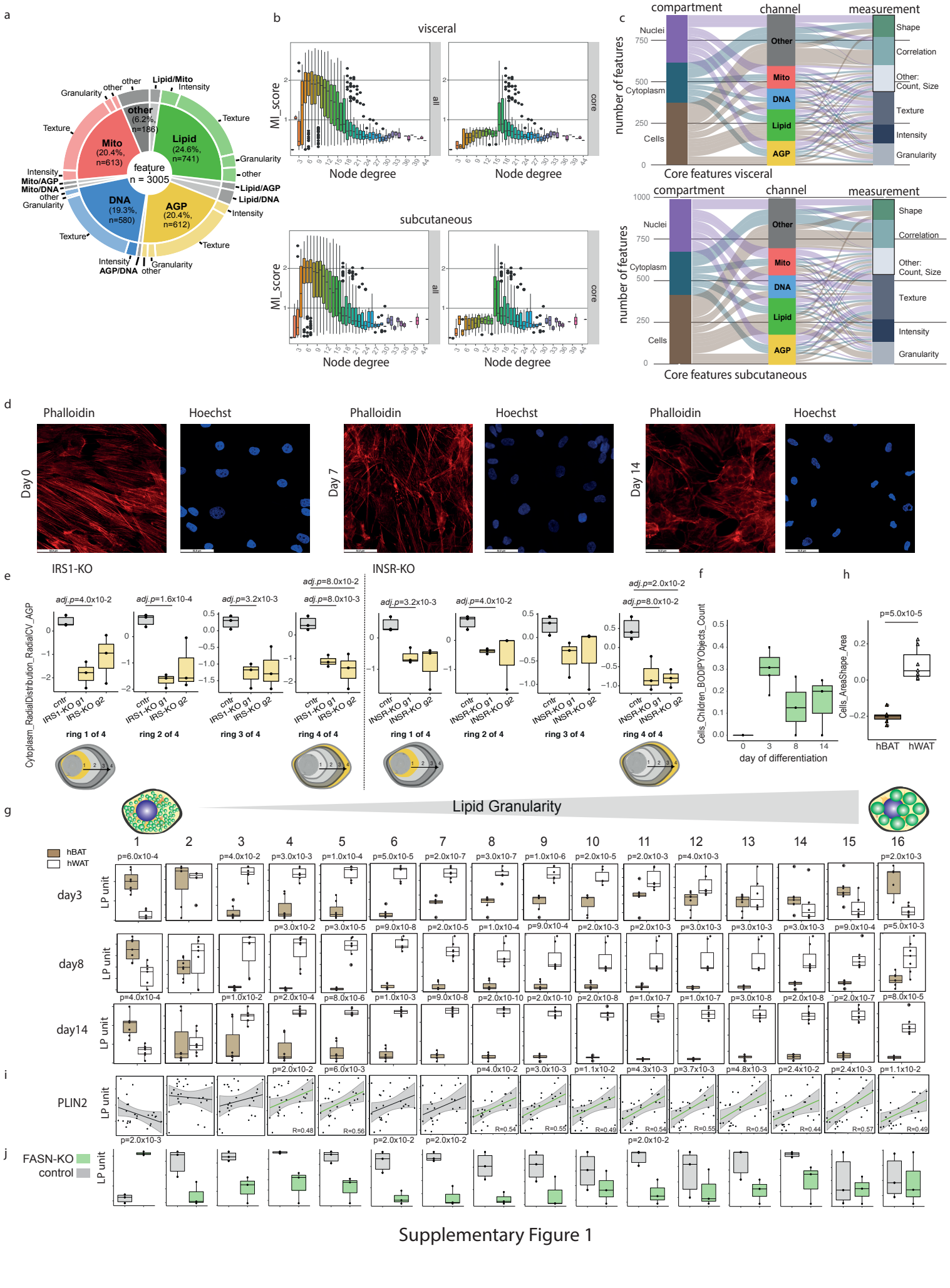
(f) Lipid droplets form in early differentiation and saturate thereafter. Y-axis shows LP units (normalized LP values across eight batches, see STAR Methods).

(g) *Lipid Granularity* measurements, captured by spectra of 16 lipid droplet size measures, show size-specific changes in hWAT and hBAT during differentiation, suggesting hBAT generally accumulate less medium-size and large lipid droplets as seen by lower values across the spectra of granularity. Y-axis shows LP units (normalized LP values across three batches, see STAR Methods).

(h) White (hWAT) are larger than brown (hBAT) adipocytes after 14 days of adipogenic differentiation as cells become lipid-laden, measured by Cells_AreaShape_Area. Y-axis shows LP units (normalized LP values across three batches, see Methods).

(i) Granularity features informative for larger lipid droplets (*Lipid Granularity 10-16*) correlate positively with *PLIN2* gene expression. Y-axis shows LP units (normalized LP values across eight batches, see Methods). X-axis *PLIN2* expression (rpm).

(j) *Lipid Granularity* measures are reduced in CRISPR/Cas9-mediated KO of *FASN* in hWAT at day 14 of differentiation. Y-axis shows LP units (normalized LP values across CRISPR-KO data, see Methods). Data is shown for one guide targeting *FASN*.



Supplementary Figure 1

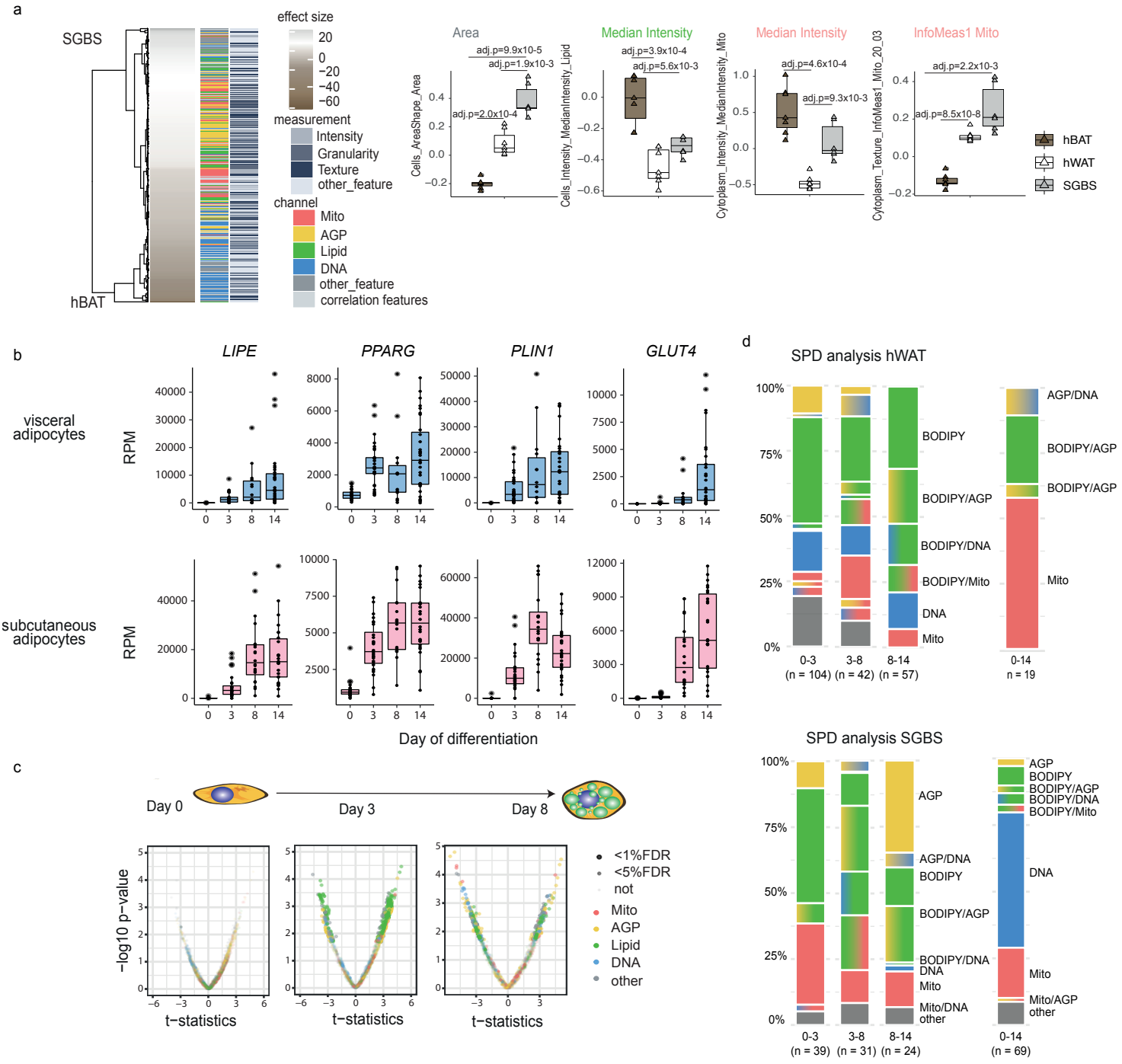
Figure S2: LipocyteProfiler identifies depot- specific cellular signatures in adipocytes, related to Figure 2.

(a) Morphological profiles of white (SGBS) and brown (hBAT) adipocytes at day 14 of differentiation differ significantly across all feature classes (FDR<0.1%). Features are clustered based on effect size (left panel). Representative LipocyteProfiler features comparing characteristics of white adipocytes (hWAT, SGBS) and brown adipocytes (hBAT).

(b) Gene expression from RNAseq of adipogenesis marker genes *LIPE, PPARG, PLIN1* and *GLUT4* in visceral (top) and subcutaneous (bottom) AMSCs throughout differentiation.

(c) Subcutaneous and visceral AMSCs have distinct morphological and cellular profiles with differences that are spread across all channels that become apparent at day 3 of differentiation and are maintained at day 8. See also Figure 2c (day 14).

(d) Sample progression discovery analysis (SPD) on LipocyteProfiler features characterizing white adipocyte differentiation in hWAT (top panel) and SGBS (bottom panel)

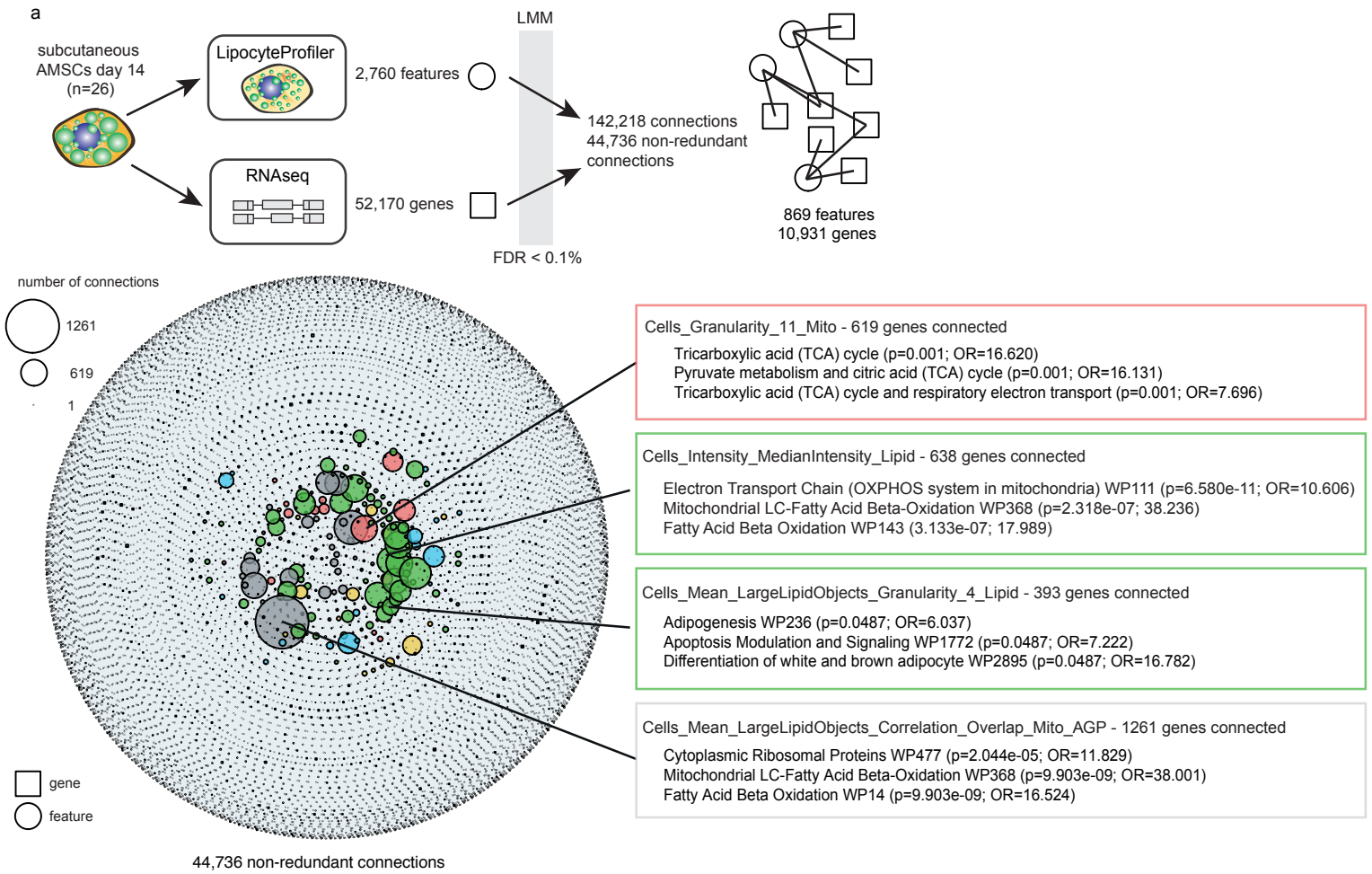


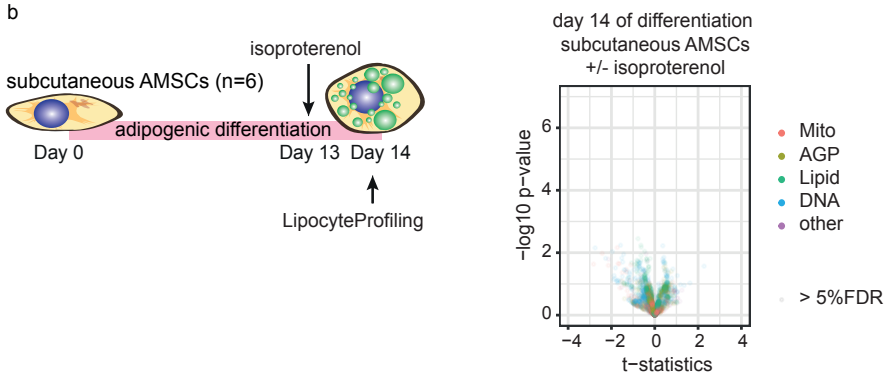
Supplementary Figure 2

Figure S3: LipocyteProfiler reveals depot-specific isoproterenol-induced cellular signatures in subcutaneous adipocytes compared to visceral adipocytes, related to Figure 3 and Figure 4.

(a) Linear mixed model (LMM) was applied to correlate 2,760 morphological features derived from LipocyteProfiler with 52,170 transcripts derived from RNAseq in matched samples of subcutaneous AMSCs at terminal differentiation (day 14). With FDR<0.1%, we discover 44,736 non-redundant connections that map to 869 morphological features and 10,931 genes. See also Figure 3a, b (FDR<0.01%).

(b) Isoproterenol treatment results in no effect on morphological profile in subcutaneous AMSCs at day 14 of differentiation. See also Figure 4b (visceral AMSCs).





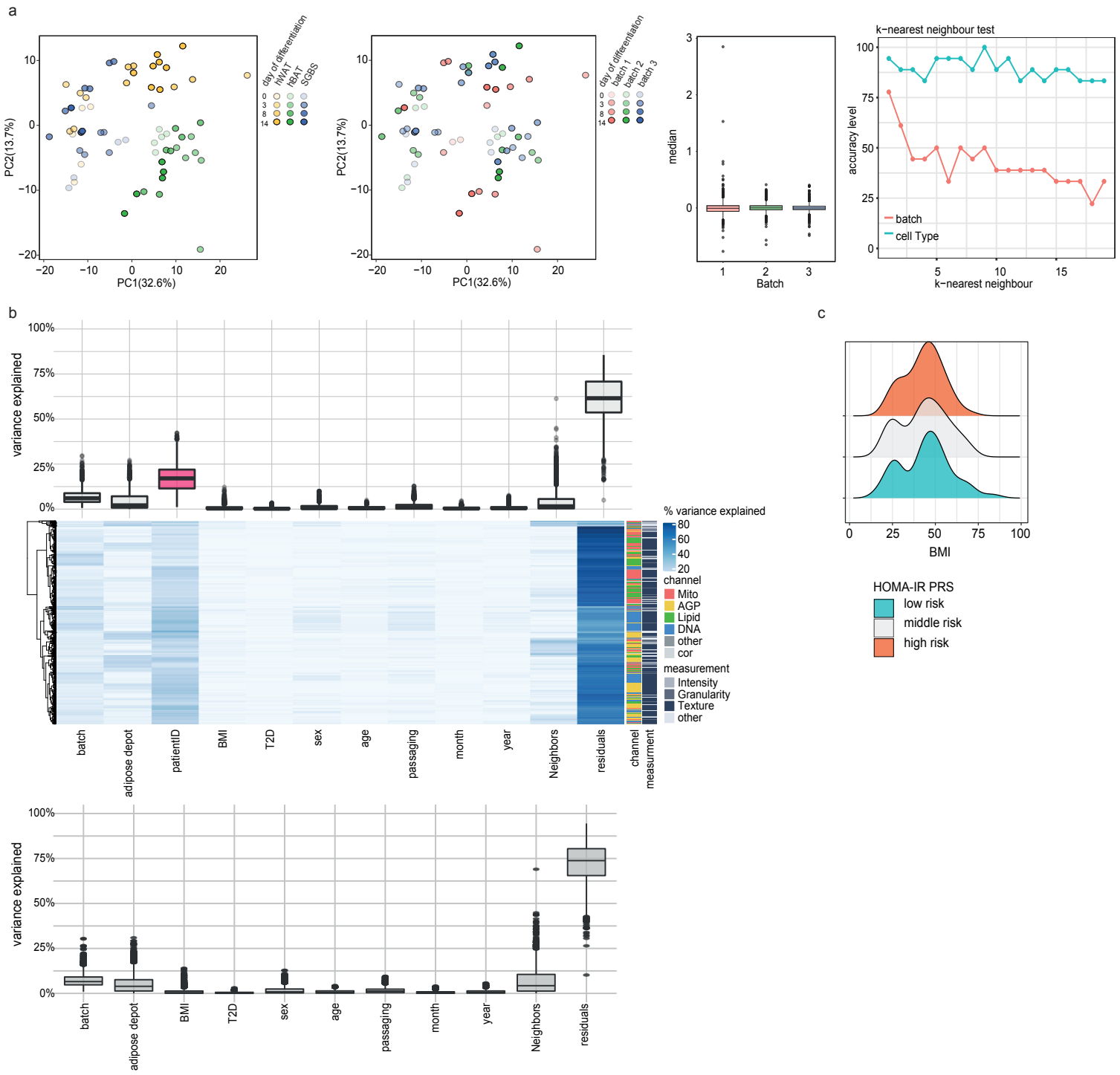
Supplementary Figure 3

Figure S4: LipocyteProfiler detects and distinguishes inter-individual genetic variation from extrinsic confounding factors, related to Figure 5.

(a) Morphological profiles of hBAT, hWAT and SGBS across differentiation cluster according to cell type and show maturation trajectory in PC1 and PC2, though do not cluster by batch (two plots on left). BEclear analysis shows no significant batch effect and accuracy of predicting cell type is higher than predicting batch using a k-nearest neighbor supervised machine learning algorithm (two plots on right).

(b) Variance component analysis across all data to assess contribution of intrinsic genetic variation on adipocyte morphology and cellular traits across 65 donor-derived differentiating AMSCs with (top panel) and without (bottom panel) patientID included in the model. This analysis showed that patientID explains the majority of feature variance compared to contribution of other possible confounding factors such as batch, adipose depot, T2D status, age, sex, BMI, cell density, month/year of sampling and passage numbers.

(c) HOMA-IR PRS is independent of BMI distribution within the AMSC cohort.



Supplementary Figure 4

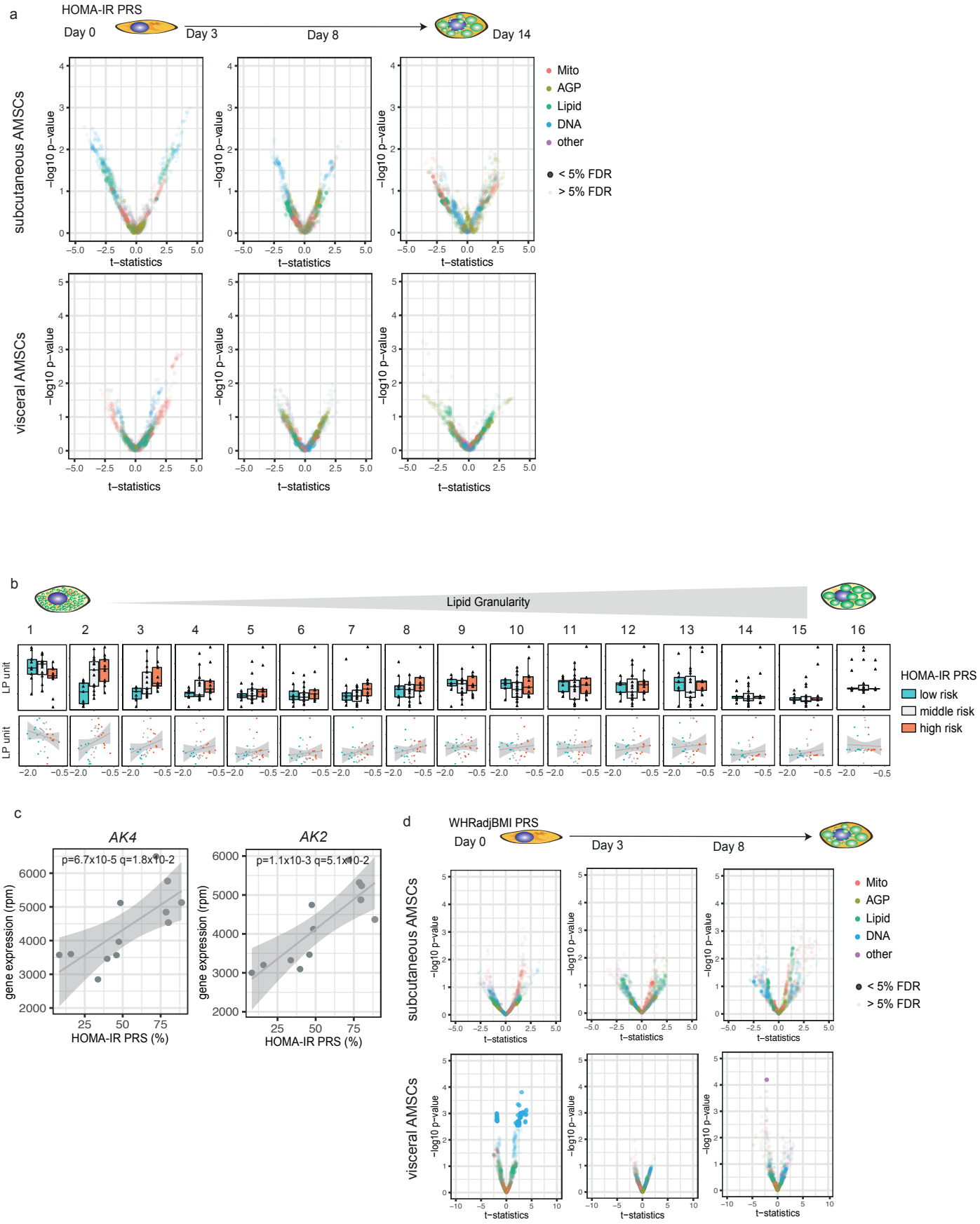
Figure S5: LipocyteProfiler discovers significant polygenic effects on image-based cellular signatures for HOMA-IR and WHRadjBMI, related to Figure 5.

(a) LipocyteProfiler differences between top and bottom 25% of HOMA-IR risk in subcutaneous and visceral AMSCs at day 0, 3, and 8 of adipogenesis. (See also Figure 5b; day 14 visceral)

(b) Mature visceral AMSCs (day 14) derived from individuals with high polygenic risk for insulin resistance show an increased number of small- to medium-sized lipid droplets compared to AMSCs from donors with low polygenic risk. *Cytoplasm_Granularity_BODIPY* measures 1-16. Y-axis shows autoscaled LP units (normalized LP values across eight batches, see methods). Bottom panel; x-axis shows PRS for HOMA-IR.

(c) AK4 and AK2 correlate significantly with HOMA-IR PRS in visceral AMSCs at day 14 of differentiation.

(d) LipocyteProfiler differences between top and bottom 25% of WHRadjBMI risk in subcutaneous and visceral AMSCs at day 0, 3, and 8 of adipogenesis. See also Figure 5b; day 14 subcutaneous).



Supplementary Figure 5

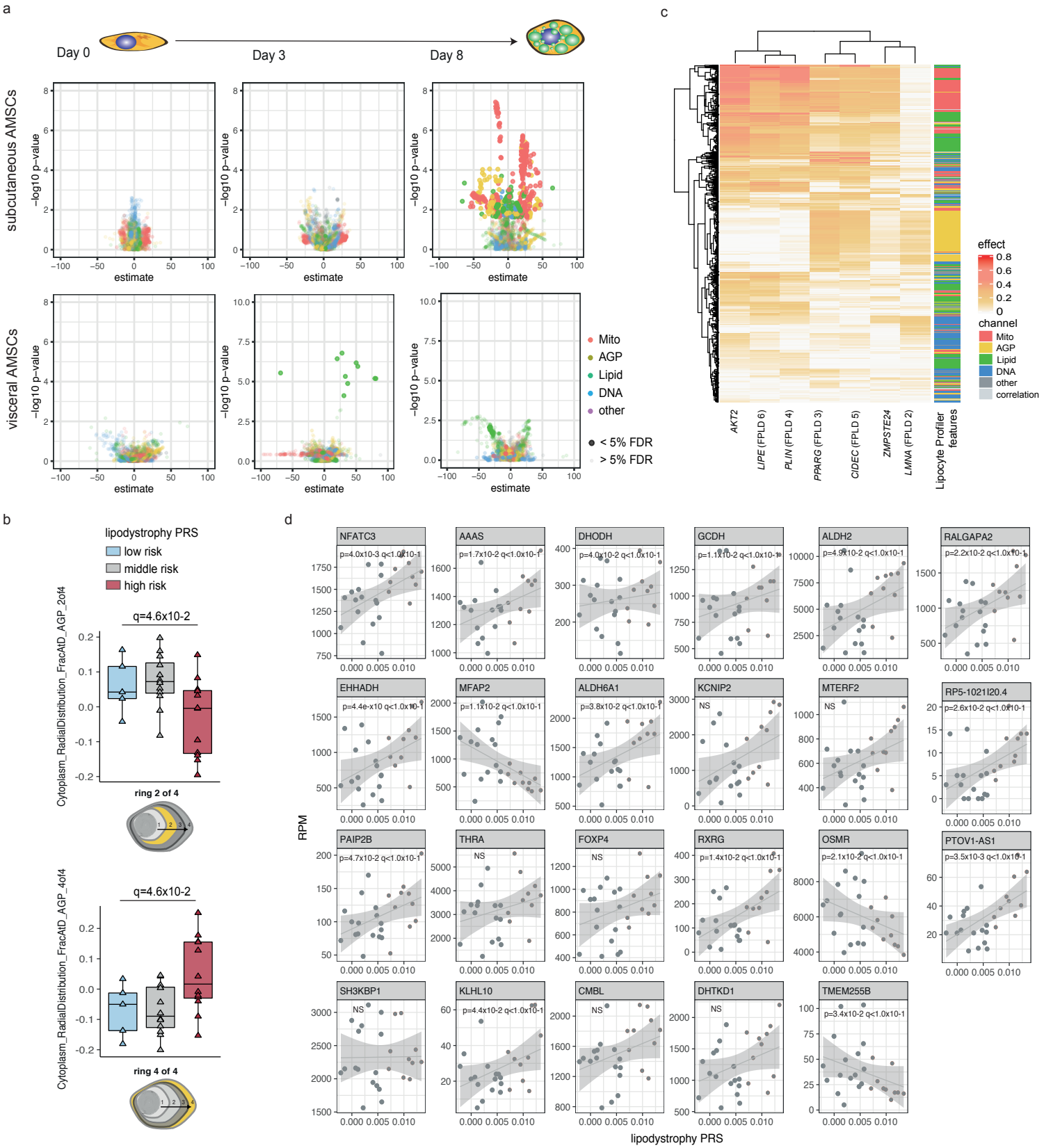
Figure S6: Morphological profile of polygenic risk for lipodystrophy shows similarity to cellular signatures of marker genes of monogenic familial partial lipodystrophy syndromes, related to Figure 6.

(a) LipocyteProfiler differences driven by polygenic risk in subcutaneous and visceral AMSCs at day 0, 3, and 8 of adipogenesis. See also Figure 6b, d; day14.

(b) Polygenic risk for lipodystrophy affects the structure of the actin cytoskeleton; golgi; plasma membrane of differentiated subcutaneous AMSCs (Cytoplasm_RadialDistribution_FraAtD_2of4 and 4of4). Y-axis shows LP units (normalized LP values across eight batches, see Methods).

(c) LipocyteProfiler features correlated with genes underlying monogenic familial partial lipodystrophy syndromes including *PPARG*, *LIPE*, *PLIN1*, *AKT2*, *CIDEC*, *LMNA* and *ZMPSTE24* match LipocyteProfiler morphological and cellular signatures characterizing the lipodystrophy-specific PRS effects, and highlight *Mito* and *AGP* features (See also Figure 6b; day14).

(d) Genes significantly correlating with the lipodystrophy PRS-mediated differential features (LMM, FDR<0.1%) are under polygenic control for lipodystrophy risk.

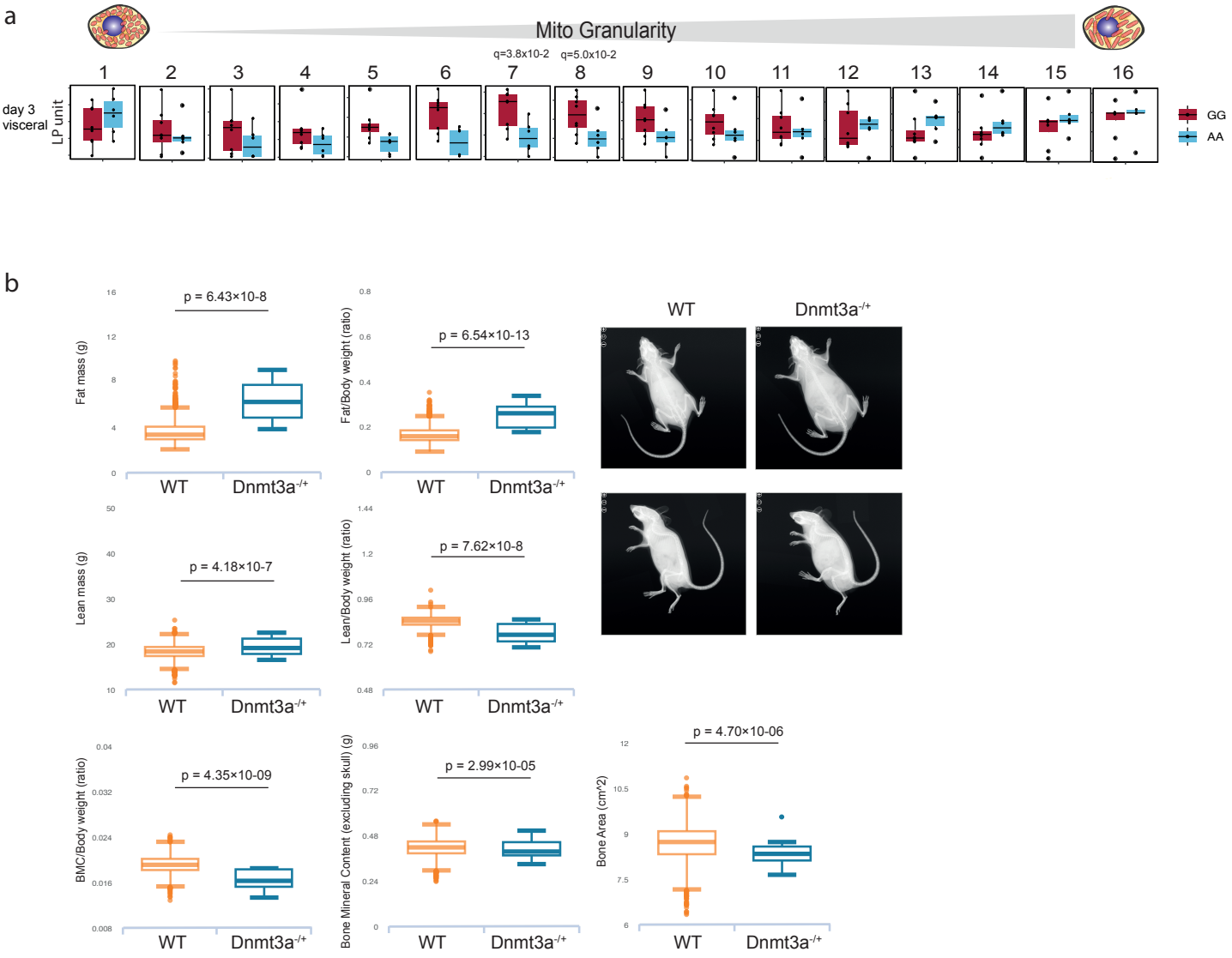


Supplementary Figure 6

Figure S7: Heterozygous knockout mice for DNMT3A have adipose and bone phenotypes, related to Figure 7.

(a) *Mito Granularity* (7-8 size measures) at day 3 of differentiation in visceral AMSCs was increased in risk allele carriers, suggestive of less tubular mitochondria. Y-axis shows autoscaled LP units (normalized LP values across eight batches, see Methods)

(b) Heterozygous knockout mice for DNMT3A show increased body weight due to increased overall fat mass and have reduced bone mineral density. Data retrieved from www.mousephenotype.org (Dickinson *et al.*, 2016).



Supplementary Figure 7

Supplemental Tables

Table S3: Effect of isoproterenol stimulation on morphological and cellular signatures in differentiated visceral adipocytes, related to Figure 4. LipocyteProfiler applied on (a) LipocyteProfiler total feature set and (b) LipocyteProfiler visceral core feature set (n=3 donors each profiled in quadruplicate, ANOVA, significance level FDR<5%). P-value, p-pvalue of ANOVA; q-value, q-value of ANOVA; eta_sq, eta square of ANOVA indicative of effect size; F value of ANOVA; t- t-statistics of t-test.

(a) Changes on morphological and cellular signatures in response to isoproterenol treatment using LipocyteProfiler total feature set (FDR<5%).

visceral adipocytes at day 14	pvalue	q-value	effect size (eta_sq)	F value	t-statistics
Nuclei_Granularity_12_DNA	4.92E-05	1.78E-02	0.08	2.03E+04	-0.75
Cytoplasm_Correlation_K_Lipid_AGP	1.33E-04	1.78E-02	0.11	7.53E+03	0.95
Cytoplasm_Texture_Correlation_Lipid_5_03	1.56E-04	1.78E-02	0.04	6.43E+03	0.49
Cells_Correlation_K_Lipid_AGP	1.64E-04	1.78E-02	0.09	6.08E+03	0.82
Cytoplasm_Texture_DifferenceVariance_Lipid_20_03	1.96E-04	1.78E-02	0.17	5.10E+03	1.18
Cells_Mean_LargeLipidObjects_AreaShape_MaximumRadius	4.16E-04	2.23E-02	0.05	2.40E+03	0.61
Cells_Mean_LargeLipidObjects_AreaShape_MeanRadius	5.27E-04	2.47E-02	0.06	1.90E+03	0.66
Cytoplasm_AreaShape_Zernike_7_3	1.14E-03	3.67E-02	0.19	8.77E+02	1.22
Cells_Mean_LargeLipidObjects_Correlation_K_Lipid_AGP	1.27E-03	3.67E-02	0.15	7.87E+02	-1.10
Cytoplasm_Correlation_K_Lipid_DNA	1.35E-03	3.67E-02	0.27	7.42E+02	1.62
Nuclei_Texture_InfoMeas1_Mito_5_02	1.51E-03	3.67E-02	0.04	6.60E+02	0.45
Nuclei_RadialDistribution_MeanFrac_Lipid_4of4	1.57E-03	3.67E-02	0.04	6.35E+02	-0.51
Nuclei_Granularity_10_DNA	1.57E-03	3.67E-02	0.12	6.34E+02	-0.95
Cytoplasm_Texture_Correlation_DNA_5_03	1.57E-03	3.67E-02	0.02	6.34E+02	-0.35
Cells_Texture_DifferenceVariance_Lipid_5_02	1.67E-03	3.67E-02	0.08	5.99E+02	0.80
Cells_Mean_LargeLipidObjects_Granularity_1_Lipid	1.89E-03	3.83E-02	0.79	5.29E+02	-5.11
Nuclei_Intensity_MassDisplacement_Lipid	1.90E-03	3.83E-02	0.03	5.26E+02	-0.43
Cells_Mean_LargeLipidObjects_Granularity_2_Lipid	1.92E-03	3.83E-02	0.68	5.19E+02	-3.57
Nuclei_RadialDistribution_RadialCV_AGP_3of4	2.17E-03	4.20E-02	0.11	4.59E+02	-0.77
Cells_RadialDistribution_FracAtD_AGP_3of4	2.21E-03	4.20E-02	0.09	4.51E+02	0.74
Nuclei_Texture_InfoMeas2_DNA_5_00	2.43E-03	4.41E-02	0.13	4.10E+02	0.99
Cytoplasm_Texture_InfoMeas1_AGP_5_03	2.50E-03	4.42E-02	0.10	3.98E+02	-0.73
Cytoplasm_Texture_DifferenceVariance_Mito_5_02	2.55E-03	4.42E-02	0.25	3.91E+02	1.48
Cells_Texture_DifferenceVariance_Mito_5_02	2.85E-03	4.71E-02	0.27	3.49E+02	1.56

Nuclei_Texture_InfoMeas2_Lipid_20_00	2.92E-03	4.71E-02	0.01	3.41E+02	-0.18
Nuclei_Texture_Correlation_DNA_5_01	2.95E-03	4.71E-02	0.05	3.38E+02	-0.57
Cells_Mean_LargeLipidObjects_AreaShape_MedianRadius	3.09E-03	4.75E-02	0.06	3.22E+02	0.67
Nuclei_RadialDistribution_FracAtD_AGP_3of4	3.23E-03	4.76E-02	0.17	3.08E+02	1.14
Nuclei_Texture_AngularSecondMoment_DNA_20_02	3.28E-03	4.76E-02	0.16	3.03E+02	-1.15
Nuclei_Granularity_11_DNA	3.32E-03	4.76E-02	0.11	3.00E+02	-0.85
Cells_Mean_LargeLipidObjects_Correlation_K_AGP_Lipid	3.40E-03	4.76E-02	0.13	2.93E+02	1.00
Nuclei_Texture_Correlation_Mito_20_01	3.40E-03	4.76E-02	0.26	2.93E+02	-1.53
Cells_Mean_LargeLipidObjects_Correlation_K_Lipid_Mito	3.48E-03	4.80E-02	0.13	2.86E+02	-1.03
Cells_Texture_DifferenceEntropy_DNA_20_01	3.55E-03	4.81E-02	0.12	2.80E+02	0.88
Cytoplasm_Texture_SumVariance_DNA_5_02	3.65E-03	4.86E-02	0.18	2.72E+02	1.05
Cells_Mean_LargeLipidObjects_AreaShape_Compactness	3.71E-03	4.86E-02	0.85	2.68E+02	-5.72

(b) Changes in morphological and cellular signatures in response to isoproterenol treatment in LipocyteProfiler visceral core feature set (FDR<5%).

visceral adipocytes at day 14	**pvalue**	**q-value**	**effect size (eta_sq)**	**F value**	**t-statistics**
Nuclei_Granularity_12_DNA	4.92E-05	1.26E-02	0.08	2.03E+04	-0.75
Cells_Mean_LargeLipidObjects_AreaShape_MaximumRadius	4.16E-04	4.49E-02	0.05	2.40E+03	0.61
Cells_Mean_LargeLipidObjects_AreaShape_MeanRadius	5.27E-04	4.49E-02	0.06	1.90E+03	0.66
Cytoplasm_AreaShape_Zernike_2_0	1.07E-03	4.72E-02	0.07	9.29E+02	0.60
Cytoplasm_Correlation_K_Lipid_DNA	1.35E-03	4.72E-02	0.27	7.42E+02	1.62
Nuclei_Texture_InfoMeas1_Mito_5_02	1.51E-03	4.72E-02	0.04	6.60E+02	0.45
Nuclei_Granularity_10_DNA	1.57E-03	4.72E-02	0.12	6.34E+02	-0.95
Cells_Mean_LargeLipidObjects_Granularity_1_Lipid	1.89E-03	4.72E-02	0.79	5.29E+02	-5.11
Nuclei_Intensity_MassDisplacement_Lipid	1.90E-03	4.72E-02	0.03	5.26E+02	-0.43
Cells_Mean_LargeLipidObjects_Granularity_2_Lipid	1.92E-03	4.72E-02	0.68	5.19E+02	-3.57
Nuclei_RadialDistribution_RadialCV_AGP_3of4	2.17E-03	4.72E-02	0.11	4.59E+02	-0.77
Cells_RadialDistribution_FracAtD_AGP_3of4	2.21E-03	4.72E-02	0.09	4.51E+02	0.74

Table S4: Effect of oleic acid treatment on morphological and cellular signatures in PHH, related to Figure 4. LipocyteProfiler applied on the total feature set (t-test, significance level FDR<0.1%). The table consists of significant features after feature reduction based on effect size and pearson correlation. P-value of t-test; q-value of t-test; t-statistics of t-test.

LipocyteProfiler features	**pvalue**	**q-value**	**t-statistics**
Cells_Mean_LargeLipidObjects_Correlation_Correlation_DNA_Lipid	1.65E-15	9.50E-13	-37.88

Cytoplasm_Correlation_Correlation_Lipid_AGP	2.28E-15	9.78E-13	-37.01
Cells_Mean_LargeLipidObjects_Granularity_2_Lipid	1.73E-14	5.93E-12	31.98
Nuclei_Intensity_MassDisplacement_Lipid	5.20E-14	1.49E-11	29.52
Cytoplasm_Texture_InfoMeas1_Lipid_5_00	2.81E-13	5.36E-11	-26.12
Cytoplasm_Correlation_K_DNA_Lipid	5.60E-13	6.82E-11	24.84
Cells_Granularity_2_Lipid	7.15E-13	6.82E-11	24.40
Cells_Correlation_Correlation_Lipid_AGP	9.47E-13	7.74E-11	-23.90
Cytoplasm_Correlation_Correlation_Mito_Lipid	1.38E-12	9.87E-11	-23.26
Nuclei_Texture_Correlation_Lipid_5_03	3.65E-12	1.50E-10	21.65
Nuclei_Correlation_Overlap_DNA_Lipid	5.53E-12	1.73E-10	-21.00
Cells_RadialDistribution_FracAtD_Lipid_2of4	1.03E-11	2.09E-10	-20.07
Cells_Correlation_K_Lipid_DNA	2.24E-11	3.34E-10	-18.94
Nuclei_Correlation_Correlation_DNA_Lipid	3.39E-11	4.73E-10	-18.37
Cells_RadialDistribution_MeanFrac_Lipid_2of4	4.08E-11	5.52E-10	-18.12
Cytoplasm_Correlation_K_Lipid_DNA	1.49E-10	1.04E-09	-16.45
Cells_Texture_SumEntropy_Lipid_20_01	1.54E-10	1.05E-09	16.42
Cells_RadialDistribution_RadialCV_Lipid_3of4	1.65E-10	1.13E-09	16.33
Nuclei_RadialDistribution_FracAtD_Lipid_2of4	7.18E-10	4.00E-09	-14.61
Nuclei_RadialDistribution_MeanFrac_Lipid_2of4	2.04E-09	1.03E-08	-13.50
Cytoplasm_Intensity_MassDisplacement_Lipid	2.70E-09	1.32E-08	13.21
Cytoplasm_Correlation_K_Mito_Lipid	5.11E-09	2.38E-08	12.58
Cells_Correlation_K_Mito_Lipid	8.06E-09	3.67E-08	12.14
Cells_Correlation_K_Lipid_Mito	1.25E-08	5.50E-08	-11.73
Cells_Intensity_MassDisplacement_Lipid	1.80E-08	7.79E-08	11.40
Cells_Correlation_Overlap_Mito_Lipid	2.09E-08	8.95E-08	-11.27
Cytoplasm_Correlation_K_Lipid_Mito	2.22E-08	9.49E-08	-11.22
Nuclei_Intensity_StdIntensity_Lipid	1.62E-07	6.54E-07	9.56
Cytoplasm_RadialDistribution_MeanFrac_Lipid_3of4	1.92E-07	7.72E-07	-9.43
Cytoplasm_Texture_AngularSecondMoment_Lipid_20_01	2.23E-07	8.96E-07	-9.31
Cells_Correlation_K_AGP_Lipid	2.61E-07	1.03E-06	9.19
Cytoplasm_Correlation_Overlap_Mito_Lipid	2.65E-07	1.04E-06	-9.19
Cells_Correlation_K_DNA_Lipid	4.14E-07	1.52E-06	8.85
Cells_Texture_AngularSecondMoment_Lipid_20_01	4.91E-07	1.79E-06	-8.73
Cytoplasm_Correlation_Overlap_Lipid_AGP	7.16E-07	2.55E-06	-8.45
Nuclei_RadialDistribution_RadialCV_DNA_3of4	8.33E-07	2.93E-06	8.35
Nuclei_Correlation_Correlation_Mito_Lipid	1.29E-06	4.32E-06	8.04
Nuclei_Granularity_1_Lipid	2.90E-06	9.30E-06	-7.50
Cells_Correlation_Correlation_Mito_Lipid	5.06E-06	1.58E-05	-7.13
Cells_Granularity_1_AGP	5.66E-06	1.75E-05	-7.06

Cells_Texture_Correlation_DNA_20_03	8.51E-06	2.57E-05	6.81
Nuclei_Correlation_K_DNA_Mito	1.01E-05	3.00E-05	6.70
Cytoplasm_Granularity_1_AGP	1.09E-05	3.21E-05	-6.65
Nuclei_Correlation_K_Mito_DNA	1.14E-05	3.33E-05	-6.63
Cytoplasm_Correlation_K_Mito_DNA	1.27E-05	3.69E-05	-6.56
Cytoplasm_RadialDistribution_FracAtD_AGP_1of4	1.70E-05	4.82E-05	6.38
Nuclei_Texture_SumEntropy_Lipid_10_01	1.74E-05	4.89E-05	6.37
Cytoplasm_RadialDistribution_MeanFrac_AGP_1of4	1.92E-05	5.27E-05	6.31
Cytoplasm_Correlation_K_DNA_Mito	1.94E-05	5.31E-05	6.30
Cells_RadialDistribution_RadialCV_DNA_4of4	3.61E-05	9.22E-05	5.94
Cytoplasm_AreaShape_Zernike_7_1	4.25E-05	1.07E-04	-5.85
Nuclei_AreaShape_Zernike_3_3	4.40E-05	1.10E-04	5.83
Nuclei_Correlation_K_Lipid_AGP	4.56E-05	1.14E-04	-5.81
Cells_Intensity_MassDisplacement_DNA	5.28E-05	1.29E-04	5.72
Nuclei_Granularity_10_Lipid	8.26E-05	1.96E-04	5.47
Cells_Texture_Correlation_Mito_20_00	1.09E-04	2.55E-04	5.32
Cells_Correlation_K_DNA_Mito	1.24E-04	2.88E-04	5.24
Cells_Neighbors_SecondClosestDistance_10	3.86E-04	8.36E-04	4.64
Cells_Neighbors_SecondClosestDistance_Adjacent	3.86E-04	8.36E-04	4.64
Cells_Neighbors_NumberOfNeighbors_10	4.42E-04	9.53E-04	-4.56

Table S6: Characteristics of study participants.

	adipose depot of AMSCs		
	sc (N=29)	**vc (N=36)**	**Total (N=65)**
Sex			
male	6 (20.7%)	9 (25.0%)	15 (23.1%)
female	23 (79.3%)	27 (75.0%)	50 (76.9%)
T2D			
not reported	1	2	3
negative	18 (64.3%)	22 (64.7%)	40 (64.5%)
positive	10 (35.7%)	12 (35.3%)	22 (35.5%)
WHR			
not reported	25	30	55
Mean +/- SD	0.87 +/- 0.069	0.86 +/- 0.055	0.86 +/- 0.057
Range	0.77 - 0.92	0.77 - 0.92	0.77 - 0.92
Age			
Mean +/- SD	41.83 +/- 13.17	43.28 +/- 12.13	42.63 +/- 12.46
Range	17 - 72	17 - 72	17 - 72

not reported	1	1	2
Mean +/- SD	50.07 +/- 10.26	48.75 +/- 12.14	49.34 +/- 11.28
Range	22.27 - 69.26	22.27 - 83.26	22.27 - 83.26

Table S7: Significant genome-wide polygenic risk for HOMA-IR effects on LipocyteProfiler features in differentiated visceral AMSCs (day 14 of differentiation), related to Figure 5. ANOVA of (a) total LipocyteProfiler feature set and (b) LipocyteProfiler visceral core feature set comparing individuals with high versus low polygenic risk for HOMA-IR (adjusted for BMI, sex, age, batch; significance level FDR<5%). Linear regression model of (c) HOMA-IR PRSs on total LipocyteProfiler cell feature set (all features but nuclei features) and (d) HOMA-IR adj.BMI PRSs on core visceral LipocyteProfiler feature set (adjusted for sex, age, batch, BMI (only in model (c)); significance level FDR<5%). P-value, p-value of ANOVA or linear regression model; q-value, q-value of ANOVA or linear regression model; eta_sq, eta square of ANOVA or linear regression model indicative of effect size; F value of ANOVA; t-statistics of t-test; estimate, estimate of linear regression model.

(a) Changes of total LipocyteProfiler features in high compared to low HOMA-IR PRS carriers (ANOVA, FDR<5%)

visceral adipocytes at day 14	pvalue	q-value	F value	t-statistics
Cells_Texture_AngularSecondMoment_Lipid_5_02	3.12E-05	3.73E-03	38.67	-3.86
Cytoplasm_Texture_AngularSecondMoment_Lipid_5_02	4.57E-05	3.73E-03	35.80	-3.80
Cytoplasm_Granularity_3_Lipid	4.31E-04	2.00E-02	21.90	3.76
Cytoplasm_Granularity_2_Lipid	4.64E-04	2.07E-02	21.51	4.15
Cells_Texture_DifferenceVariance_Lipid_10_01	1.13E-03	2.70E-02	17.28	-3.19
Nuclei_Correlation_Correlation_Mito_AGP	1.16E-03	2.70E-02	17.16	-3.67
Cells_Texture_InverseDifferenceMoment_Lipid_5_01	1.35E-03	2.70E-02	16.49	-3.54
Cells_Correlation_Overlap_DNA_Lipid	1.37E-03	2.70E-02	16.42	-4.36
Cytoplasm_Texture_InverseDifferenceMoment_Lipid_5_01	1.55E-03	2.70E-02	15.90	-3.54
Nuclei_Intensity_LowerQuartileIntensity_Lipid	1.69E-03	2.70E-02	15.54	3.07
Nuclei_Texture_DifferenceVariance_Lipid_5_00	2.15E-03	2.70E-02	14.55	-2.99
Nuclei_Texture_InfoMeas2_Lipid_10_02	2.15E-03	2.70E-02	14.55	2.60
Cells_AreaShape_Zernike_8_4	2.25E-03	2.70E-02	14.37	2.28
Nuclei_RadialDistribution_MeanFrac_Lipid_1of4	2.59E-03	2.70E-02	13.80	-3.70

Cytoplasm_Texture_DifferenceEntropy_Lipid_5_01	2.62E-03	2.70E-02	13.76	3.28
Cells_Granularity_2_Lipid	2.69E-03	2.70E-02	13.67	3.44
Cytoplasm_Texture_SumEntropy_Lipid_5_01	2.74E-03	2.70E-02	13.60	3.15
Cells_Texture_SumEntropy_Lipid_5_01	2.79E-03	2.70E-02	13.52	3.11
Cells_Texture_DifferenceEntropy_Lipid_5_00	2.97E-03	2.70E-02	13.29	3.27
Cells_Texture_Entropy_Lipid_5_01	3.01E-03	2.70E-02	13.23	3.22
Nuclei_Intensity_MinIntensity_Lipid	3.07E-03	2.70E-02	13.16	3.50
Cytoplasm_Texture_Entropy_Lipid_10_00	3.19E-03	2.70E-02	13.01	3.22
Nuclei_Intensity_MinIntensityEdge_Lipid	3.28E-03	2.70E-02	12.90	3.50
Nuclei_Texture_AngularSecondMoment_Lipid_10_03	3.31E-03	2.70E-02	12.87	-2.73
Cells_Mean_LargeLipidObjects_Granularity_2_Lipid	3.44E-03	2.70E-02	12.73	3.58
Cells_Mean_LargeLipidObjects_Granularity_1_Lipid	3.57E-03	2.76E-02	12.59	3.66
Cytoplasm_Texture_DifferenceVariance_Lipid_10_03	3.68E-03	2.77E-02	12.47	-2.98
Nuclei_Correlation_Correlation_DNA_Lipid	4.05E-03	2.90E-02	12.12	-3.45
Nuclei_RadialDistribution_RadialCV_Lipid_4of4	4.52E-03	3.08E-02	11.73	3.20
Nuclei_Correlation_Overlap_DNA_Lipid	4.81E-03	3.21E-02	11.51	-3.35
Cells_Texture_InfoMeas1_Lipid_5_00	5.22E-03	3.45E-02	11.23	3.15
Cells_Granularity_3_Lipid	5.26E-03	3.46E-02	11.19	2.63
Cytoplasm_RadialDistribution_RadialCV_Lipid_3of4	5.31E-03	3.46E-02	11.16	2.99
Cells_Correlation_K_Lipid_DNA	5.58E-03	3.59E-02	10.99	-2.91
Cells_RadialDistribution_RadialCV_Lipid_2of4	5.98E-03	3.78E-02	10.75	2.87
Cells_RadialDistribution_FracAtD_DNA_3of4	6.44E-03	3.95E-02	10.50	-3.69
Nuclei_Texture_SumEntropy_Lipid_10_01	7.32E-03	4.44E-02	10.07	2.87
Cells_Mean_LargeLipidObjects_Granularity_3_Lipid	7.51E-03	4.51E-02	9.99	3.02

(b) Changes of LipocyteProfiler visceral core features in high compared to low HOMA-IR PRS carriers (ANOVA, FDR<5%)

visceral adipocytes at day 14	**pvalue**	**q-value**	**F value**	**t-statistics**
Nuclei_Correlation_Correlation_Mito_AGP	1.16E-03	2.70E-02	17.16	-3.67
Cells_Correlation_Overlap_DNA_Lipid	1.37E-03	2.70E-02	16.42	-4.36
Cells_AreaShape_Zernike_8_4	2.25E-03	2.70E-02	14.37	2.28
Cells_Granularity_2_Lipid	2.69E-03	2.70E-02	13.67	3.44
Cells_Mean_LargeLipidObjects_Granularity_2_Lipid	3.44E-03	2.70E-02	12.73	3.58
Cytoplasm_Texture_Entropy_Lipid_10_03	3.47E-03	2.70E-02	12.70	3.18
Cells_Mean_LargeLipidObjects_Granularity_1_Lipid	3.57E-03	2.76E-02	12.59	3.66
Nuclei_Correlation_Correlation_DNA_Lipid	4.05E-03	2.90E-02	12.12	-3.45
Nuclei_RadialDistribution_RadialCV_Lipid_4of4	4.52E-03	3.08E-02	11.73	3.20
Cells_Texture_InfoMeas1_Lipid_5_00	5.22E-03	3.45E-02	11.23	3.15

Cytoplasm_RadialDistribution_RadialCV_Lipid_3of4	5.31E-03	3.46E-02	11.16	2.99
Cells_Correlation_K_Lipid_DNA	5.58E-03	3.59E-02	10.99	-2.91
Cells_RadialDistribution_RadialCV_Lipid_2of4	5.98E-03	3.78E-02	10.75	2.87
Cells_Texture_InfoMeas1_Lipid_5_02	6.00E-03	3.78E-02	10.74	3.13
Cytoplasm_RadialDistribution_RadialCV_Lipid_2of4	6.06E-03	3.79E-02	10.71	2.83
Cells_RadialDistribution_FracAtD_DNA_3of4	6.44E-03	3.95E-02	10.50	-3.69
Cells_Mean_LargeLipidObjects_Granularity_3_Lipid	7.51E-03	4.51E-02	9.99	3.02

(c) Associations between HOMA-IR PRSs and total LipocyteProfiler feature set (linear regression model, FDR<5%)

visceral adipocytes at day 14	pvalue	q-value	estimate
Cells_Mean_LargeLipidObjects_Granularity_7_Lipid	2.79E-05	2.67E-02	0.23
Cells_Mean_LargeLipidObjects_Granularity_8_Lipid	7.22E-05	4.59E-02	0.23

(d) Associations between HOMA-IR adj. BMI PRSs and LipocyteProfiler visceral core feature set (linear regression model, FDR<5%)

visceral adipocytes at day 8	pvalue	q-value	estimate
Cytoplasm_Correlation_K_Mito_Lipid	7.24E-04	3.73E-02	0.22
Cytoplasm_Intensity_MaxIntensity_Lipid	8.30E-04	3.73E-02	0.24

Table S9: List of significant genes of linear regression model of HOMA-IR PRSs on gene expression data of 512 genes known to be involved in adipocyte function, related to Figure 5. Adjusted for BMI, sex, age batch, significance level FDR<10%. Gene ID, Ensembl gene identification number; Gene name, gene name; pvalue, p-value of linear regression; q-value, q-value of linear regression.

Gene ID	Gene name	pvalue	q-value
ENSG00000162433	AK4	6.7E-05	1.8E-02
ENSG00000150593	PDCD4	8.4E-04	5.1E-02
ENSG00000165092	ALDH1A1	9.1E-04	5.1E-02
ENSG00000116171	SCP2	9.4E-04	5.1E-02
ENSG00000084234	APLP2	1.1E-03	5.1E-02
ENSG00000004455	AK2	1.1E-03	5.1E-02
ENSG00000141526	SLC16A3	2.2E-03	7.8E-02
ENSG00000130304	SLC27A1	2.5E-03	7.8E-02
ENSG00000148175	STOM	3.6E-03	7.8E-02

ENSG00000211445	GPX3	3.7E-03	7.8E-02
ENSG00000159231	CBR3	3.9E-03	7.8E-02
ENSG00000111275	ALDH2	4.2E-03	7.8E-02
ENSG00000074800	ENO1	4.4E-03	7.8E-02
ENSG00000151640	DPYSL4	4.4E-03	7.8E-02
ENSG00000151552	QDPR	4.7E-03	7.8E-02
ENSG00000122644	ARL4A	5.2E-03	7.8E-02
ENSG00000104812	GYS1	5.5E-03	7.8E-02
ENSG00000164237	CMBL	5.5E-03	7.8E-02
ENSG00000141232	TOB1	5.6E-03	7.8E-02
ENSG00000188994	ZNF292	6.5E-03	8.7E-02
ENSG00000067057	PFKP	7.3E-03	9.2E-02
ENSG00000163516	ANKZF1	8.0E-03	9.2E-02
ENSG00000105976	MET	8.0E-03	9.2E-02
ENSG00000111669	TPI1	8.3E-03	9.2E-02
ENSG00000213619	NDUFS3	9.2E-03	9.5E-02
ENSG00000156709	AIFM1	9.3E-03	9.5E-02
ENSG00000146242	TPBG	9.8E-03	9.5E-02
ENSG00000152952	PLOD2	1.1E-02	9.5E-02
ENSG00000112715	VEGFA	1.1E-02	9.5E-02
ENSG00000128039	SRD5A3	1.1E-02	9.5E-02
ENSG00000159228	CBR1	1.2E-02	9.5E-02
ENSG00000152137	HSPB8	1.2E-02	9.5E-02
ENSG00000111897	SERINC1	1.2E-02	9.5E-02
ENSG00000205726	ITSN1	1.2E-02	9.5E-02
ENSG00000104267	CA2	1.3E-02	9.5E-02
ENSG00000152583	SPARCL1	1.3E-02	9.5E-02
ENSG00000143198	MGST3	1.4E-02	9.5E-02
ENSG00000143590	EFNA3	1.5E-02	9.5E-02
ENSG00000248144	ADH1C	1.5E-02	9.5E-02
ENSG00000167772	ANGPTL4	1.6E-02	9.5E-02
ENSG00000060971	ACAA1	1.6E-02	9.5E-02
ENSG00000100823	APEX1	1.6E-02	9.5E-02
ENSG00000078070	MCCC1	1.7E-02	9.5E-02
ENSG00000131724	IL13RA1	1.7E-02	9.5E-02
ENSG00000102144	PGK 1.00	1.7E-02	9.5E-02
ENSG00000143847	PPFIA4	1.7E-02	9.5E-02
ENSG00000147852	VLDLR	1.7E-02	9.5E-02
ENSG00000127083	OMD	1.7E-02	9.5E-02

ENSG00000015532	XYLT2	1.7E-02	9.5E-02
ENSG00000117620	SLC35A3	1.8E-02	9.5E-02
ENSG00000130203	APOE	1.8E-02	9.5E-02

Table S10: Lists of KEGG pathways enriched (FDR<5%) among associated genes with HOMA-IR PRSs at FDR<10%, related to Figure 5. Term, which pathway; overlap, number of genes that overlap and total genes of pathway; pvalue, enrichment p-value; adj. pvalue, q-value; OR, odds ratio of enrichment; CS, combined score, an approximation of overall association (-log10(P) * log(Odds)), genes, genes of the pathway which are associated with HOMA-IR PRS.

term	overlap	pvalue	adj. pvalue	OR	CS	genes
Glycolysis / Gluconeogenesis	5/68	8.1E-07	7.3E-05	34.31	481.07	TPI1;ALDH2; ADH1C;ENO1; PFKP
Metabolism of xenobiotics by cytochrome P450	4/74	3.8E-05	1.1E-03	24.17	246.10	CBR1;ADH1C; MGST3;CBR3
PPAR signaling pathway	4/74	3.8E-05	1.1E-03	24.17	246.10	SLC27A1;SCP2; ANGPTL4;ACAA1
Fatty acid degradation	3/44	1.9E-04	4.3E-03	30.35	259.69	ADH1C;ALDH2; ACAA1
Valine, leucine and isoleucine degradation	3/48	2.5E-04	4.5E-03	27.64	229.38	ALDH2;MCCC1; ACAA1
Arachidonic acid metabolism	3/63	5.6E-04	7.4E-03	20.72	155.24	CBR1;GPX3;CBR3
Central carbon metabolism in cancer	3/65	6.1E-04	7.4E-03	20.05	148.38	SLC16A3;MET; PFKP
Thiamine metabolism	2/15	6.6E-04	7.4E-03	62.59	458.83	AK2;AK4
RNA degradation	3/79	1.1E-03	1.1E-02	16.34	111.67	ENO1;TOB1;PFKP
Chemical carcinogenesis	3/82	1.2E-03	1.1E-02	15.72	105.72	CBR1;ADH1C; MGST3
Folate biosynthesis	2/26	2.0E-03	1.6E-02	33.89	210.72	QDPR;CBR1
Biosynthesis of unsaturated fatty acids	2/27	2.1E-03	1.6E-02	32.53	199.83	SCP2;ACAA1
Fructose and mannose metabolism	2/33	3.2E-03	2.2E-02	26.23	150.65	TPI1;PFKP
MicroRNAs in cancer	4/299	7.0E-03	4.3E-02	5.67	28.10	EFNA3;PDCD4; MET;VEGFA
Cholesterol metabolism	2/50	7.2E-03	4.3E-02	16.92	83.44	ANGPTL4;APOE
Glutathione metabolism	2/56	9.0E-03	5.1E-02	15.04	70.85	GPX3;MGST3
Lysine degradation	2/59	9.9E-03	5.3E-02	14.24	65.68	ALDH2;PLOD2

PI3K-Akt signaling pathway	4/354	1.3E-02	6.0E-02	4.77	20.87	GYS1;EFNA3;MET; VEGFA
Retinol metabolism	2/67	1.3E-02	6.0E-02	12.49	54.53	ADH1C;ALDH1A1
Renal cell carcinoma	2/69	1.3E-02	6.0E-02	12.11	52.22	MET;VEGFA
Proteoglycans in cancer	3/201	1.5E-02	6.3E-02	6.23	26.34	PDCD4;MET; VEGFA
Rap1 signaling pathway	3/206	1.6E-02	6.4E-02	6.08	25.29	EFNA3;MET; VEGFA
Peroxisome	2/83	1.9E-02	7.4E-02	10.01	39.67	SCP2;ACAA1
Ras signaling pathway	3/232	2.1E-02	8.0E-02	5.38	20.71	EFNA3;MET; VEGFA
HIF-1 signaling pathway	2/100	2.7E-02	9.7E-02	8.27	29.89	ENO1;VEGFA
Insulin resistance	2/108	3.1E-02	1.0E-01	7.64	26.54	GYS1;SLC27A1
Drug metabolism	2/108	3.1E-02	1.0E-01	7.64	26.54	ADH1C;MGST3
AMPK signaling pathway	2/120	3.8E-02	1.2E-01	6.86	22.50	GYS1;PFKP
MAPK signaling pathway	3/295	3.9E-02	1.2E-01	4.21	13.61	EFNA3;MET; VEGFA
Nitrogen metabolism	1/17	4.2E-02	1.2E-01	24.92	78.70	CA2
Primary bile acid biosynthesis	1/17	4.2E-02	1.2E-01	24.92	78.70	SCP2
Purine metabolism	2/129	4.3E-02	1.2E-01	6.37	20.07	AK2;AK4
Pathways in cancer	4/530	4.6E-02	1.2E-01	3.14	9.69	MGST3;MET; IL13RA1;VEGFA
Fluid shear stress and atherosclerosis	2/139	4.9E-02	1.3E-01	5.90	17.80	MGST3;VEGFA
Proximal tubule bicarbonate reclamation	1/23	5.7E-02	1.4E-01	18.12	51.88	CA2
Histidine metabolism	1/23	5.7E-02	1.4E-01	18.12	51.88	ALDH2
alpha-Linolenic acid metabolism	1/25	6.2E-02	1.5E-01	16.60	46.20	ACAA1
Ascorbate and aldarate metabolism	1/27	6.7E-02	1.5E-01	15.33	41.50	ALDH2
Collecting duct acid secretion	1/27	6.7E-02	1.5E-01	15.33	41.50	CA2
Hepatocellular carcinoma	2/168	6.8E-02	1.5E-01	4.86	13.04	MGST3;MET
Alzheimer disease	2/171	7.1E-02	1.5E-01	4.78	12.66	NDUFS3;APOE
Pentose phosphate pathway	1/30	7.4E-02	1.6E-01	13.74	35.81	PFKP
beta-Alanine metabolism	1/31	7.6E-02	1.6E-01	13.28	34.19	ALDH2
Galactose metabolism	1/31	7.6E-02	1.6E-01	13.28	34.19	PFKP
Axon guidance	2/181	7.8E-02	1.6E-01	4.51	11.51	EFNA3;MET
Base excision repair	1/33	8.1E-02	1.6E-01	12.45	31.31	APEX1
Starch and sucrose metabolism	1/36	8.8E-02	1.6E-01	11.38	27.67	GYS1
Tyrosine metabolism	1/36	8.8E-02	1.6E-01	11.38	27.67	ADH1C
Focal adhesion	2/199	9.2E-02	1.7E-01	4.09	9.79	MET;VEGFA

Pyruvate metabolism	1/39	9.5E-02	1.7E-01	10.48	24.68	ALDH2
Bladder cancer	1/41	9.9E-02	1.8E-01	9.95	22.97	VEGFA
Tryptophan metabolism	1/42	1.0E-01	1.8E-01	9.71	22.19	ALDH2
Arginine and proline metabolism	1/49	1.2E-01	2.0E-01	8.29	17.74	ALDH2
Malaria	1/49	1.2E-01	2.0E-01	8.29	17.74	MET
N-Glycan biosynthesis	1/50	1.2E-01	2.0E-01	8.12	17.22	SRD5A3
Glycosaminoglycan biosynthesis	1/53	1.3E-01	2.0E-01	7.65	15.81	XYLT2
VEGF signaling pathway	1/59	1.4E-01	2.2E-01	6.86	13.48	VEGFA
Steroid hormone biosynthesis	1/60	1.4E-01	2.2E-01	6.74	13.15	SRD5A3
Glycerolipid metabolism	1/61	1.4E-01	2.2E-01	6.63	12.83	ALDH2
Epithelial cell signaling in Helicobacter pylori infection	1/68	1.6E-01	2.3E-01	5.93	10.89	MET
Adherens junction	1/72	1.7E-01	2.3E-01	5.60	9.98	MET
Melanoma	1/72	1.7E-01	2.3E-01	5.60	9.98	MET
Bile secretion	1/72	1.7E-01	2.3E-01	5.60	9.98	CA2
Inositol phosphate metabolism	1/74	1.7E-01	2.3E-01	5.45	9.57	TPI1
Thyroid hormone synthesis	1/74	1.7E-01	2.3E-01	5.45	9.57	GPX3
Bacterial invasion of epithelial cells	1/74	1.7E-01	2.3E-01	5.45	9.57	MET
Pancreatic cancer	1/75	1.7E-01	2.3E-01	5.37	9.38	VEGFA
Gastric acid secretion	1/75	1.7E-01	2.3E-01	5.37	9.38	CA2
Rheumatoid arthritis	1/91	2.1E-01	2.7E-01	4.41	6.93	VEGFA
Pancreatic secretion	1/98	2.2E-01	2.9E-01	4.09	6.16	CA2
AGE-RAGE signaling pathway in diabetic complications	1/100	2.3E-01	2.9E-01	4.01	5.97	VEGFA
Glucagon signaling pathway	1/103	2.3E-01	2.9E-01	3.89	5.69	GYS1
Thyroid hormone signaling pathway	1/116	2.6E-01	3.2E-01	3.45	4.69	PFKP
Relaxin signaling pathway	1/130	2.8E-01	3.4E-01	3.07	3.88	VEGFA
Oxidative phosphorylation	1/133	2.9E-01	3.5E-01	3.00	3.73	NDUFS3
Insulin signaling pathway	1/137	3.0E-01	3.5E-01	2.91	3.55	GYS1
Parkinson disease	1/142	3.1E-01	3.5E-01	2.81	3.34	NDUFS3
Apoptosis	1/143	3.1E-01	3.5E-01	2.79	3.30	AIFM1
Retrograde endocannabinoid signaling	1/148	3.2E-01	3.5E-01	2.69	3.11	NDUFS3
Non-alcoholic fatty liver disease (NAFLD)	1/149	3.2E-01	3.5E-01	2.68	3.07	NDUFS3
Gastric cancer	1/149	3.2E-01	3.5E-01	2.68	3.07	MET
JAK-STAT signaling pathway	1/162	3.4E-01	3.7E-01	2.46	2.65	IL13RA1
Necroptosis	1/162	3.4E-01	3.7E-01	2.46	2.65	AIFM1
Kaposi sarcoma-associated herpesvirus infection	1/186	3.8E-01	4.0E-01	2.14	2.07	VEGFA
Transcriptional misregulation in cancer	1/186	3.8E-01	4.0E-01	2.14	2.07	MET
Huntington disease	1/193	3.9E-01	4.1E-01	2.06	1.94	NDUFS3
Human cytomegalovirus infection	1/225	4.4E-01	4.5E-01	1.76	1.45	VEGFA

Thermogenesis	1/231	4.5E-01	4.6E-01	1.71	1.38	NDUFS3
Cytokine-cytokine receptor interaction	1/294	5.3E-01	5.4E-01	1.34	0.85	IL13RA1
Human papillomavirus infection	1/330	5.7E-01	5.7E-01	1.19	0.67	VEGFA

Method S1: Step-by-Step Protocol - LipocyteProfiler tool for image-based profiling of lipid accumulating cellular model systems, related to Figure 1.

Summary

LipocyteProfiler is an image-based profiling tool for the quantification of morphological and cellular phenotypic differences among a variety of lipocyte cell populations. The LipocyteProfiler protocol below describes the specific steps for high-throughput multiplexed fluorescence staining, automated microscopic imaging and the automated computational pipeline to generate rich cellular and morphological profiles in lipid-accumulating cells (Figure 1).

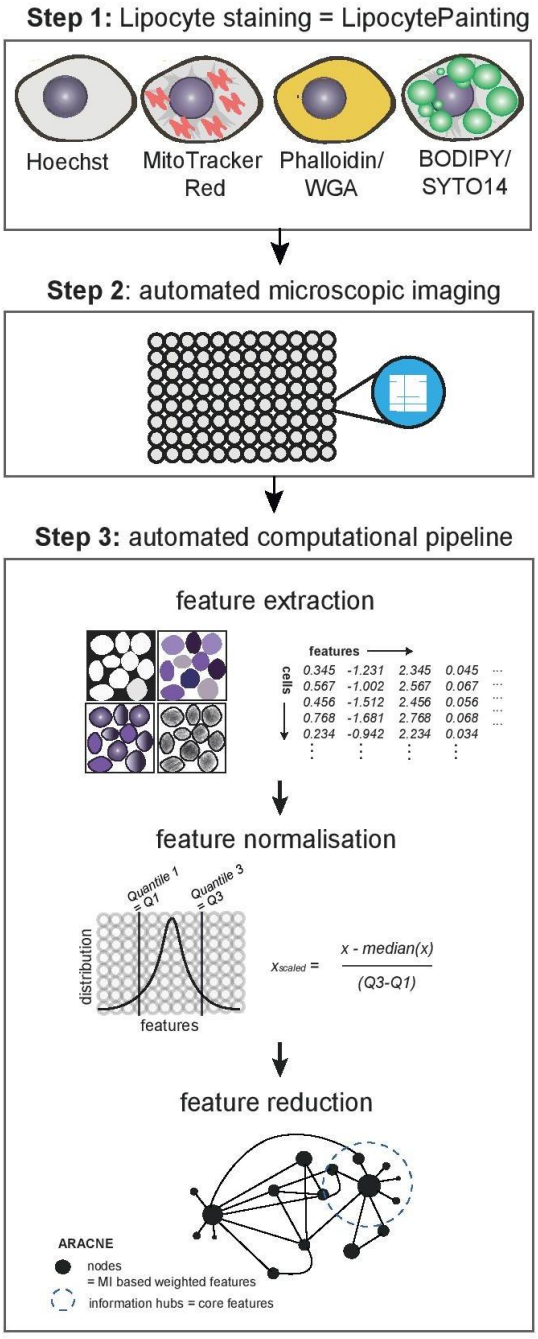


Figure 1: Schematic of the LipocyteProfiler tool

Before you begin

Before you begin you should consider the following steps:

1. Culture cells in black 96-well plates appropriate for high-content fluorescent image-based screening. We recommend the use of PhenoPlate 96-well Perkin Elmer Cell Carrier Black plates.
2. Pre-warm the media in a water bath of 37°C for 1h.
3. Prepare all stock solutions of fluorescence stains. One hour before starting, bring dyes to room temperature (RT) while kept in dark.
4. Buffers need to be prepared freshly and brought to RT.

5. For differentiating primary human adipocytes the following specifications apply: adipose-derived mesenchymal stem cells (AMSCs) are seeded at 10K cells/well in a 100ul/well proliferation medium solution. After 4 days AMSCs are induced for differentiation with induction medium and on day 3 of differentiation adipocytes induction medium needs to be replaced with differentiation medium. Differentiation medium needs to be changed every 4 days.

Key resources table

See separate file

Materials and equipment

1x HBSS

Reagent	**Final concentration**	**Amount**
HBSS	10X	20 ml
ddH2O		180 mL
Total	**1x**	**200 mL**

*prepare fresh

1% BSA in 1x HBSS

Reagent	**Final concentration**	**Amount**
BSA		100mg
1x HBSS		10 mL
Total	**1x**	**10ml**

*prepare fresh

0.1% Triton X-100

Reagent	**Final concentration**	**Amount**
Triton	100X	10 ul
HBSS	1x	10 ml
Total	**0.1x**	**10 ml**

*prepare fresh and prewarm at 37°C

MitoTracker Red stock solution

Reagent	**Final concentration**	**Amount**
MitoTracker Red		50ug
DMSO		91ul
Total	**1mM**	**91ul**

*store at -20C in dark for max. 1 month

BODIPY stock solution

Reagent	**Final concentration**	**Amount**
BODIPY 505/515		10mg
Ethanol		
Total	**3.8mM**	**50 mL**

*store at -20C in dark

Phalloidin stock solution

Reagent	**Final concentration**	**Amount**
Phalloidin/Alexa Fluor 568		300 units
Methanol		1.5 mL
Total	**200units/ml**	**1.5 mL**

*store at -20C in dark for up to 1 year

WGA stock solution

Reagent	**Final concentration**	**Amount**
WGA/Alexa Fluor 555		5 mg
ddH2O		5 mL
Total	**1mg/ml**	**5ml**

*store at -20C in dark for up to 1 year

**centrifuge before use to minimize aggregate proteins in solution (non-specific background staining)

Cell culture medium

Use the culture and differentiation medium optimized for the cellular model system of interest.

Specification for differentiating primary human AMSC model system**:**

Basic

Reagent	**Final concentration**
DMEM-F12	
Penicillin - Streptomycin	1%
Biotin/Pantothenate	33 µM/17 µM

Proliferation

Reagent	**Final concentration**
Basic	
Insulin	0.13 µM
Epidermal growth factor	0.01 µg/ml
Fibroblast growth factor	0.001 µg/ml
Fetal calf serum	2.5%

Induction of differentiation

Reagent	**Final concentration**
Basic	
Insulin	0.861 µM
Triiodothyronine	1 nM
Cortisol	0.1 µM
Transferrin	0.01 mg/ml
Rosiglitazone	1 µM
Dexamethasone	25 nM
3-Isobutyl-1-methylxanthin	2.5 nM

*prepare fresh each time

Differentiation

Reagent	**Final concentration**
Basic	
Insulin	0.861 µM
Triiodothyronine	1 nM
Cortisol	0.1 µM
Transferrin	0.01 mg/ml

*store at 4C for max 1 week

Specification for differentiating primary human hTERT immortalized white adipocyte (hWAT) and brown adipocyte (hBAT) model systems**:**

Basic medium and proliferation medium

Reagent	**Final concentration**
DMEM GlutaMAX	
Bovine Serum	10 %
Penicillin - Streptomycin	1%

Induction and differentiation

Reagent	**Final concentration**
Basic	
Human Insulin	0.5 µM
Triiodothyronine	0.002 µM
Biotin	33 µM
Pantothenate	17 µM
Rosiglitazone	1 µM
Dexamethasone	0.1 µM
3-Isobutyl-1-methylxanthin	500 µM
Indomethacin	30 µM

Specification for differentiating the SGBS adipocyte model system**:**

Basic

Reagent	**Final concentration**
DMEM/F12	
Biotin/Pantothenate	33 µM/17 µM
Penicillin - Streptomycin	1%

Proliferation

Reagent	**Final concentration**
Basic	
Fetal calf serum	10 %

Induction

Reagent	**Final concentration**
Basic	
Insulin	0.02 µM
Cortisol	0.1 µM
Triiodothyronine	0.0002 µM
Dexamethasone	0.025 µM
3-Isobutyl-1-methylxanthin	250 µM
Rosiglitazone	2 µM

Differentiation

Reagent	**Final concentration**
Basic	
Insulin	0.02 µM
Cortisol	0.1 µM
Triiodothyronine	0.0002 µM
Dexamethasone	0.025 µM

Specification for the primary human hepatocyte model system**:**

Medium	**Order No.**
INVITROGRO CP Medium	BioIVT Z99029
TORPEDO Antibiotic Mix	BioIVT Z99000

Additional reagents

Reagent
Hoechst 33342 stock solution 10mg/ml *store at 4C in dark for up to 6 months
SYTO 14 *store at -20C in dark for up to 1 year
16% Methanol-free PFA
DPBS

Equipment

Cell culture equipment
PhenoPlate 96-well, black, optically clear flat-bottom, tissue-culture treated
Multi-channel pipettes
Cell culture reservoirs
Imaging equipment
Opera Phenix® High-Content Screening System, Perkin Elmer (confocal mode) 20x water-immersion objective

Step-by-step method details

A. LipocyteProfiler Staining (= LipocytePainting)

Volumes in this protocol are suitable for staining of one 96-well plate.

(!) Always carefully add and remove reagents without touching the bottom of the well in order to avoid detachment of cells from the plastic.

1. Live cell staining
    a. MitoTracker Red staining
        i. Prepare MitoTracker staining solution by thoroughly mixing 7ml of cell culture medium and 3.5ul MitoTracker stock solution to a final concentration of 0.5uM

           Specification for differentiating adipocytes: Use cell-stage specific media to prepare the solution.
        ii. Remove culture medium carefully from each well and add 60ul/well of MitoTracker staining solution.
        iii. Incubate at 37°C in dark for 30 min
    b. BODIPY staining
        i. Prepare BODIPY staining solution by thoroughly mixing 7ml of DPBS and 5.53ul BODIPY stock solution to a final concentration of 3.0uM
        ii. Remove MitoTracker staining solution carefully from the 96-well plate and wash twice with 100ul/well DPBS.
        iii. Add 60ul/well of BODIPY staining solution to each well.
        iv. Incubate at 37°C in dark for 15 min

Timing: 10 min hands-on, 45 min incubation

2. Fixation and permeabilization of cells
    a. Fixation of cells
        i. Add 15ul/well 16% Methanol-free PFA directly to wells with BODIPY staining solution
        ii. Incubate at RT in dark for 20 min
    b. Permeabilization of cells
        i. Remove 16% Methanol-free PFA and BODIPY staining solution carefully
        ii. Wash once with 100ul/well 1x HBSS

iii. Add 60ul/well of prewarmed 0.1% Triton X-100 solution carefully to the wells

iv. Incubate at RT in dark for 10 min

Timing: 10 min hands-on, 30 min incubation

3. Multi-staining of fixed cells
    a. Prepare the multi-stain solution by thoroughly mixing 50ul Phalloidin stock solution (final concentration 1 unit/ml), 10ul Hoechst 33342 (final concentration 10ug/ml), 15ul WGA stock solution (final concentration 1,5ug/ml), and 6ul SYTO-14 (final concentration 3uM) to 10ml 1x HBSS and 1% BSA (w/v).
    b. Remove 0.1% Triton X-100 carefully and wash cells twice with 100ul/well 1x HBSS.
    c. Add 60ul/well of multi-stain solution carefully to each well.
    d. Incubate at RT in dark for 30 min
    e. Remove multi-stain solution carefully and wash cells three times with 100ul/well 1x HBSS. Leave 1x HBSS from the last washing step in each well.
    f. Wrap the plate in tin foil and store at 4°C until microscopy imaging (keep plates after staining for max. for 1 week)

Timing: 15 min hands-on, 30 min incubation

Troubleshooting tips:

1. 384-well format: This protocol is readily amenable to scaling to 384-well format. The seeding density varies depending on the cell line/type, *e.g.* immortalized human hWAT cells are seeded at a density of 1,250 cells/well diluted in 40ul and primary human AMSCs are seeded at 2,500 cells/well diluted in 40ul. Use half of the LipocytePainting buffer and dye volumes described in this protocol for a 384-well plate.
2. Using automated reagent dispensation (e.g. MULTIDROP COMBI, Thermo SCIENTIFIC) and microplate washer (e.g. 405 Touch, Biotek) is recommended when performing LipocytePainting in a high-throughput setting (e.g. 384-well plates). The

following is an example workflow (details and incubation times between steps are listed in “A”):

   a. 405 Touch: Aspirate 384; Travel rate = 3 CW
   b. MULTIDROP COMBI: Dispense 40ul MitoTracker staining solution; standard cassette; low speed
   c. 405 Touch: Wash twice with 50ul DPBS; Travel rate = 3 CW; Flow rate = 2 CW; with final aspiration
   d. MULTIDROP COMBI: Dispense 40ul BODIPY staining solution; standard cassette; low speed
   e. MULTIDROP COMBI: Dispense 10ul 16% Methanol-free PFA (c_final = 4%); standard cassette; low speed
   f. 405 Touch: Wash once with 50ul 1x HBSS; Travel rate = 3 CW; Flow rate = 2 CW; with final aspiration
   g. MULTIDROP COMBI: Dispense 40ul 0.1% Triton X-100; standard cassette; low speed
   h. 405 Touch: Wash twice with 50ul 1x HBSS; Travel rate = 3 CW; Flow rate = 2 CW; with final aspiration
   i. MULTIDROP COMBI: Dispense 40ul multi-stain solution; standard cassette; low speed
   j. 405 Touch: Wash three times with 50ul 1x HBSS; Travel rate = 3 CW; Flow rate = 2 CW; without final aspiration.

3. Multiwell plates as used for LipocyteProfiling is known to be subject to edge effects. To avoid such technical artifacts on the morphological profile, leave outer wells of the plate empty.
4. Arrange samples as random as possible.
5. Using batch-to-batch control cells across all experiments improves batch effect correction.

B. LipocyteProfiler Microscopy Imaging

For imaging use a High-Content imaging system. In the LipocyteProfiler manuscript we used the Opera Phenix® High-Content Screening System (PerkinElmer Inc.).

The Opera Phenix® High-Content imaging system has four excitation lasers paired to specific emission filters. The lasers and filters utilized for LipocyteProfiler are:

- Ex405/Em435-480nm (Hoechst 33342, blue)
- Ex488/Em500-550 (BODIPY and SYTO14, green)
- Ex561/Em570-630 (Phalloidin and WGA, orange)
- Ex640/Em650-760 (MitoTracker Deep Red, red)
- Transmitted light/brightfield as an additional channel and suitable for feature extraction

Images of the individual channels should be captured in grayscale at high bit-depth and exported in a 'lossless' format. We recommend using a 16-bit grayscale TIFF format.

All images are acquired in confocal-mode using a 20x water-immersion objective (NA 1.0), 1x binning and a two-peak autofocus. Prior to each set of experiments, the exposure time and Z-offset should be adjusted for each channel.

To achieve a high well coverage optimize the number of fields per well to be imaged. For LipocyteProfiling we used 25 fields per well.

Troubleshooting tips:

1. While imaging of a single plane was sufficient for our purposes, some cell types or applications may require the acquisition of multiple planes (Z-stack).
2. A reduction in the number of imaged fields per well may be considered to decrease the experiment runtime and dataset size, however, at the expense of well coverage and statistical power. We recommend testing the settings on a few manually selected wells across the plate before each run.

C. LipocyteProfiler automated image analysis pipeline

C.1. LipocyteProfiler feature extraction

For this step of the pipeline, we explain protocols for using CellProfiler (Stirling et al. 2021) to segment and measure imaging features using the distinct stainings protocols discussed in the previous section. The step-by-step protocols described here are modified from the Imaging-based Profiling Handbook (Image-based Profiling Handbook). These steps should be followed to setup and execute CellProfiler and to extract the imaging features. The handbook step-by-step instructions focus on running the analysis on the Amazon Web Services (AWS), however, CellProfiler and the downstream steps can also be executed on a local machine or an on-premises computing cluster.

C.1.1. Spin up an AWS virtual machine (Details):

It is recommended to use an m4.xlarge instance with an 8Gb EBS volume on AWS. You will need to be able to mount S3 and EFS.

- Launch an ec2 instance on AWS
- AMI: cytomining/images/hvm-ssd/cytominer-ubuntu-trusty-18.04-amd64-server-1529668435
- Instance Type: m4.xlarge
- Network: vpc-35149752
- Subnet: Default (imaging platform terraform)
- IAM role: s3-imaging-platform-role
- No Tags
- Select Existing Security Group: SSH_HTTP
- Review and Launch
- ssh -i <USER>.pem ubuntu@<Public DNS IPv4>
- tmux new -s sessionname

C.1.2. Environment and Required libraries (Details):

Four software packages and set of scripts are required to conduct the analysis: i. Distributed CellProfiler (McQuin et al. 2018), ii. scripts to convert Phenix metadata to comma separated files called pe2loaddata, iii. tools for organizing measured features called cytominer-database,

and iv. data processing tools, called pycytominer. In terms of environment variables, you will need to define global variables, including your project name, plate/batch ID, number of cores to be used on AWS:

- PROJECT_NAME=2022_06_05LipocyteProfiler
- BATCH_ID=BR00101139__2019-09-23T21_17_31-Measurement1
- BUCKET=imaging-platform
- MAXPROCS=3 # m4.xlarge has 4 cores; this should be # of cores on your instance - 1
- mkdir -p ~/efs/${PROJECT_NAME}/workspace/
- cd ~/efs/${PROJECT_NAME}/workspace/
- mkdir -p log/${BATCH_ID}
- cd ~/efs/${PROJECT_NAME}/workspace/
- mkdir software
- cd software
- git clone https://github.com/broadinstitute/pe2loaddata.git
- git clone https://github.com/CellProfiler/Distributed-CellProfiler.git
- cd ..

C.1.3. Prepare images/files (Details): Before starting the analysis, you will need to generate a list of images that you want to analyze. To prepare this list, you can use the scripts provided in the pe2loaddata suite, or follow alternative approaches. It is recommended to compile a plate-map that can be used for preparing the list and also for the downstream analysis. A plate-map should annotate wells in every plate. For example, which well contains biological samples, and which cell type or day of differentiation a sample represents. An example plate-map is provided on the LipocyteProfiler github page.

Upload Images: You will need to upload your images to the AWS environment. Use Cyberduck (or your prefered file-transferring client) or the AWS command line interface using

- aws s3 sync /local/path s3://BUCKET/PROJECT_NAME/BATCH_ID/images

Organize images: Use the pe2loaddata to create image sets.

- cd ~/efs/${PROJECT_NAME}/workspace/
- mkdir images #Run this only if this is the first batch for this project
- cd images

- ln -s ~/bucket/projects/${PROJECT_NAME}/${BATCH_ID}/images/ ${BATCH_ID}
- cd ..

Create list of plates: Use pe2loaddata to create a text file with one plate ID per line.

- mkdir -p ~/efs/${PROJECT_NAME}/workspace/scratch/${BATCH_ID}/
- PLATES=$(readlink -f ~/efs/${PROJECT_NAME}/workspace/scratch/${BATCH_ID}/plates_to_process.txt)
- FULL_PLATES=$(readlink -f ~/efs/${PROJECT_NAME}/workspace/scratch/${BATCH_ID}/full_plates_to_process.txt)
- ls ~/efs/${PROJECT_NAME}/workspace/images/${BATCH_ID}/ | cut -d '_' -f 1 >> $PLATES
- ls ~/efs/${PROJECT_NAME}/workspace/images/${BATCH_ID}/ >> $FULL_PLATES

Create LoadData CSVs: Use pe2loaddata to create a load_data.csv file. For this step you will need to create a config.yml file as instructed here. Pay special attention to the channel names and assure consistency of the names across different files. Follow the instructions:

- cd ~/efs/${PROJECT_NAME}/workspace/software/pe2loaddata
- git pull
- pyenv shell 3.8.10
- pip3 install -e .
- parallel \

  --link \

  --max-procs 1 \

  --eta \

  --joblog ../../log/${BATCH_ID}/create_csv_from_xml.log \

  --results ../../log/${BATCH_ID}/create_csv_from_xml \

  --files \

  --keep-order \

pe2loaddata config.yml \

~/efs/${PROJECT_NAME}/workspace/load_data_csv/${BATCH_ID}/{1}/load_data.csv \

--index-directory ~/efs/${PROJECT_NAME}/workspace/images/${BATCH_ID}/{2}/Images \

--illum \

--illum-directory /home/ubuntu/bucket/projects/${PROJECT_NAME}/${BATCH_ID}/illum/{1} \

--plate-id {1} \

--illum-output ~/efs/${PROJECT_NAME}/workspace/load_data_csv/${BATCH_ID}/{1}/load_data_with_illum.csv \

--sub-string-out efs/${PROJECT_NAME}/workspace/images/${BATCH_ID} \

--sub-string-in bucket/projects/${PROJECT_NAME}/${BATCH_ID}/images :::: ${PLATES} ${FULL_PLATES}

- aws s3 sync \

  ~/efs/${PROJECT_NAME}/workspace/load_data_csv/${BATCH_ID}/ \

  s3://${BUCKET}/projects/${PROJECT_NAME}/workspace/load_data_csv/${BATCH_ID}/

Make sure the files are structures appropriately (see example)

C.1.4. Configure Distributed CellProfiler (Details): You should configure the run_batch_general.py scripts for different steps of the analysis. For the initial step adjust project name and plate names, plate lists, etc:

- topdirname and batchsuffix should match your PROJECT_NAME and BATCH_ID, respectively
- appname is typically the same as topdirname, but if that name is long and cumbersome you can create an abbreviated version here (ie 2015_10_05_DrugRepurposing rather

than 2015_10_05_DrugRepurposing_AravindSubramanian_GolubLab_Broad). This will be used in your config.py file

- rows, columns, and sites should reflect the imaging conditions used
- platelist should contain a list of plates, comma separated, ie ['SQ00015167','SQ00015168']
- If you are using pipeline files with the LoadData module and CSVs, you should make sure that the pipeline names reflect your pipeline names (or adjust if not). Otherwise, you should make sure that the batch file names reflect your batch file names.

Follow the recommended structures and procedures and do not change any of the not project specific sections of the script. You also need to set the AWS_REGION, SSH_KEY_NAME, AWS_BUCKET, and SQS_DEAD_LETTER_QUEUE settings to appropriate settings for your account.

C.1.5. Run CellProfiler (Details):

We recommend running as many as five CellProfiler pipelines i) Z projection, ii) QC, iii) illumination correction, iv) assay development (tuning of the segmentation parameters), v) analysis (segmentation and measurement of the features). Of these, only illumination correction and analysis are mandatory, others are optional depending on your exact workflow. The Z projection step is necessary when images are acquired with multiple planes. The QC step is considered to assure imaging plate quality, identify out-of-focus images and detect images with imaging artifacts. As discussed below, this step should be conducted as outlined by Bray and Carpenter (Bray and Carpenter 2018) and using CellProfiler-Analyst (Stirling et al. 2021). The illumination correction step will create an illumination correction function that is needed for the segmentation analysis. The assay development step is needed to optimize the segmentation parameters without incurring the time or compute costs of performing the feature extraction. Finally, the analysis step performs the final segmentations and measures the features from each image site. The features should be stored in CSV format for the next step of the analysis.

Repeat these steps for each of the five modules as instructed below:

- Configure the config.py file (see below)

- Execute python3 run.py setup
- Uncomment the correct step name in your run_batch_general.py file (and ensure all other steps are commented out)
- Execute python3 run_batch_general.py
- Execute python3 run.py startCluster files/yourFleetFileName.json, where you have set the name of the fleet file previously created or located
- Execute python3 run.py monitor files/APP_NAMESpotFleetRequestId.json, where APP_NAME matches the APP_NAME variable set in the first step.

Tips for setting the parameters in config.py for each step of the analysis:

i) Z projection:

- Your APP_NAME variable should be set to the appname set in run_batch_general.py plus _Zproj, ie 2022_06_05_LipocyteProfiler_Zproj
- Your number of CLUSTER_MACHINES should be medium-large, ie a hundred or few hundred.
- Your SQS_MESSAGE_VISIBILITY should be short, such as 5*60 (5 minutes)

ii) QC:

- Your APP_NAME variable should be set to the appname set in run_batch_general.py plus _QC, ie 2022_06_05_LipocyteProfiler_QC
- Your number of CLUSTER_MACHINES should be medium-large, ie a hundred or few hundred.
- Your SQS_MESSAGE_VISIBILITY should be short, such as 5*60 (5 minutes)

iii) illumination correction:

- Your APP_NAME variable should be set to the appname set in run_batch_general.py plus _Illum, ie 2022_06_05_LipocyteProfiler_Illum
- Your number of CLUSTER_MACHINES should be set to the number of plates you have divided by 4 then rounded up, ie 6 for 22 plates
- Your SQS_MESSAGE_VISIBILITY should be 12 hours 720*60

iv) assay development:

- Your APP_NAME variable should be set to the appname set in run_batch_general.py plus _AssayDev, ie 2022_06_05_LipocyteProfiler_AssayDev
- Your number of CLUSTER_MACHINES should be medium-large, ie a hundred or few hundred.
- Your SQS_MESSAGE_VISIBILITY should be short, such as 5*60 (5 minutes)

v) analysis:

- Your APP_NAME variable should be set to the appname set in run_batch_general.py plus _Analysis, ie 2022_06_05LipocyteProfiler_Analysis
- Your number of CLUSTER_MACHINES should be as many as possible per your account limits, ideally at least a few hundred.
- Your SQS_MESSAGE_VISIBILITY should be 10-20 minutes for images with a binning of 2, longer (30-120 minutes) for unbinned images and/or CellProfiler 2 or 3 runs. This value is the most potentially variable -once you've run a single analysis workflow, you can adjust this value accordingly based on your log files.

If performing the QC step, you may wish to do so before running the final analysis steps in order to throw out bad images ahead of time, otherwise an appropriate QC classifier can be created and applied *post hoc*. Note that in either case, you should ensure the MeasureImageQuality modules in your QC and analysis pipelines are set identically.

C.1.6. Aggregate features in sqlite format (Details): Before aggregating the extracted features, you will need to add an EBS volume to your AWS machine (details). Download the csv files(results of the analysis module) to the EBS volume and follow the instructions to run pycytominer:

- mkdir -p ~/ebs_tmp/${PROJECT_NAME}/workspace/software
- cd ~/ebs_tmp/${PROJECT_NAME}/workspace/software
- if [-d pycytominer]; then rm -rf pycytominer; fi
- git clone https://github.com/cytomining/pycytominer.git
- cd pycytominer
-
- python3 -m pip install -e .
- mkdir -p ../../log/${BATCH_ID}/

- parallel \

  --max-procs ${MAXPROCS} \

  --ungroup \

  --eta \

  --joblog ../../log/${BATCH_ID}/collate.log \

  --results ../../log/${BATCH_ID}/collate \

  --files \

  --keep-order \

  python3 pycytominer/cyto_utils/collate.py ${BATCH_ID}
  pycytominer/cyto_utils/ingest_config.ini {1} \

  --temp ~/ebs_tmp \

  --remote=s3://${BUCKET}/projects/${PROJECT_NAME}/workspace :::: ${PLATES}

C.2. LipocyteProfiler normalization: Use the 'Normalize by Plate nonPACs.ipynb' script from the github page to normalize the sqlite files.

C.3. LipocyteProfiler feature reduction and downstream analysis

The process for identifying the core LipocyteProfiler features uses the ARACNE software package. To install, utilize, and identify the core futures use the following instructions.

C.3.1. Installing prerequisite libraries: Use the Anaconda 4.8.3 to install python 3.9.7 and the prerequisite software packages and libraries ANT 1.10.12, JAVA 1.8.0.41, the networkx 2.6.3, and also download and install the ARACNE package inside the conda environment. Define the environment paths conda_installer_name, path_to_java, path_to_aracne and path_to_ant.

- envname=LP_PostProcessing
- bash ${conda_installer_name}
- conda create -n ${envname} networkx=2.6.3 python=3.9.7 pandas=1.3.4 numpy=1.20.3
- conda activate ${envname}

- conda env config vars set JAVA_HOME=${path_to_java}/java-se-8u41-ri ANT_HOME=${path_to_ant}/apache-ant-1.10.12 PATH=${path_to_ant}/apache-ant-1.10.12/bin:${path_to_java}/java-se-8u41-ri/bin:$PATH
- conda activate ${envname}
- cd ${path_to_aracne}
- ant main

C.3.2. Input files: For every cell type, create a tab separated values (tsv) file containing the complete set of the LipocyteProfiler features in rows and the columns should indicate the feature measurement for every subject. As an example the input file corresponding to the visceral cell type is provided on the LipocyteProfiler github page vc_features_for_core_analysis.tsv. In addition, a text file containing the complete list of feature names is needed.

- input_tsv_file=path/to/vc_features_for_core_analysis.tsv
- complete_feature_set_list=path/to/txt_file

C.3.3 ARACNE threshold calculation: Run ARACNE threshold calculation using p-value = 0.05 and seed = 1 using the input tsv file.

- java -Xmx5G -jar ${path_to_aracne}/dist/aracne.jar -e ${input_tsv_file} -o ${output_file_name} --tfs ${complete_feature_set_list} --pvalue 5E-2 --seed 1 --calculateThreshold

C.3.4 Generate Interactions through bootstrapping: Run ARACNE with 100 bootstraps with p-value = 0.05 and varying seeds from 1 to 100.

- for i in {1..100};
  do
  java -Xmx5G -jar ${path_to_aracne}/dist/aracne.jar -e ${input_tsv_file} -o ${output_file_name} --tfs ${complete_feature_set_list} --pvalue 5E-2 --seed $i
  done;

We note that the number of samples required for an analysis with enough statistical power is about 100. If the power is not enough, the output of this step will be an empty file.

C.3.5. Consolidate bootstrap results: Use the default ARACNE input parameters to aggregate the bootstrap results and to identify the weighted edges (mutual information) between the features.

- aracne_output=path/to/output_consolidated_results.txt
- java -Xmx5G -jar ${path_to_aracne}/dist/aracne.jar -o ${output_file_name} --consolidate

C.3.6. Obtain core features: The entire process for calculating the average mutual information (MI) for every node, identifying hub nodes, and then extracting the list of core features can be run using the LP_core_feature_selection_analysis.py script from the github page. The script first assigns a MI for every node by averaging over the MI of its corresponding edges. Then the script ranks nodes based on their degrees (number of edges) and selects the 75^{th} percentile of top ranked nodes, which are called hub nodes. Then this script ranks and selects the 25^{th} percentile of nodes with lower MI. The outcome of the script, named LipocyteProfiler_core_list.tsv, of the script is the list of core features, resulting from the union of hub and lower MI nodes.

- python LP_core_feature_selection_analysis.py ${aracne_output} 0.75 0.25

.