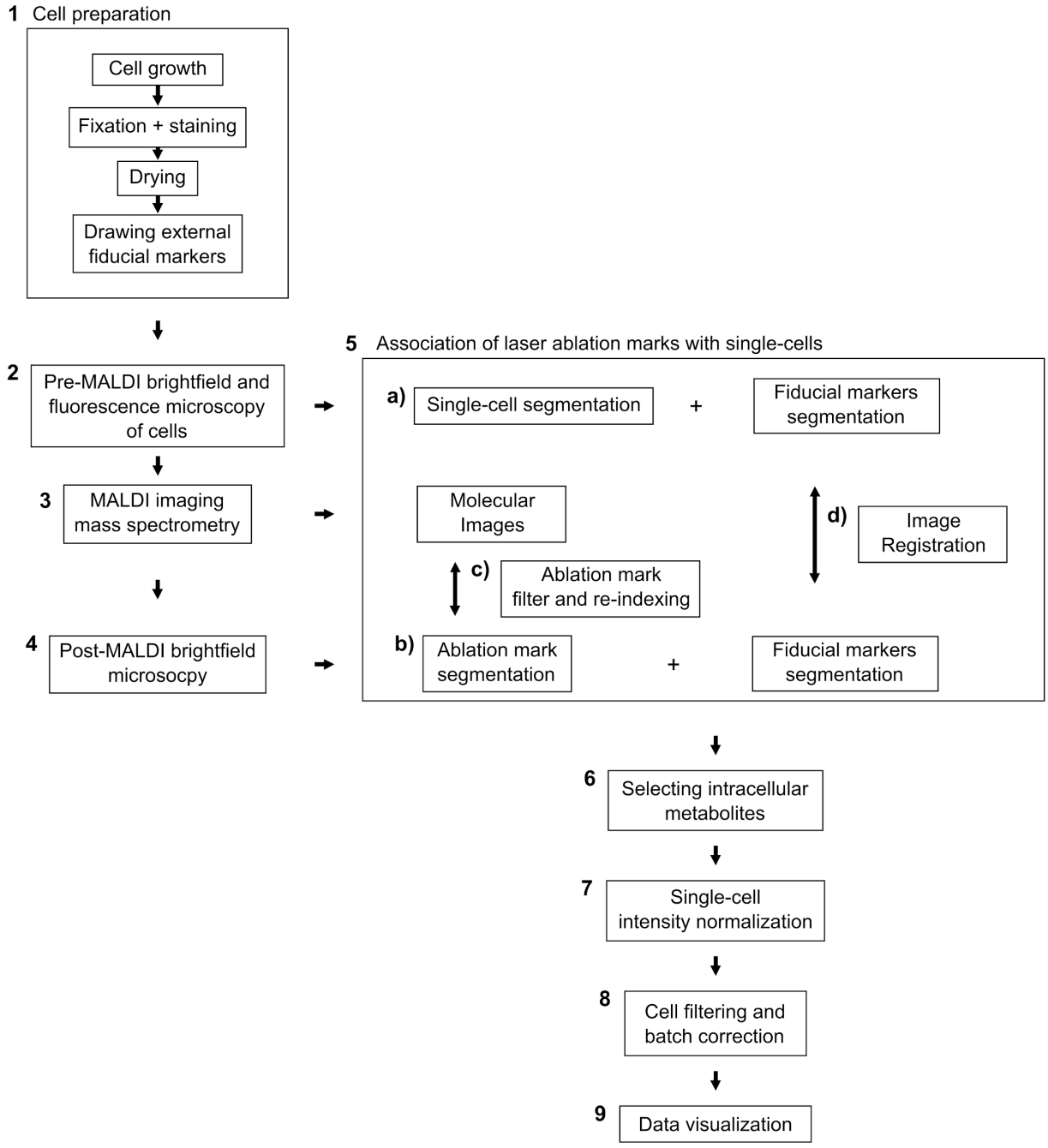
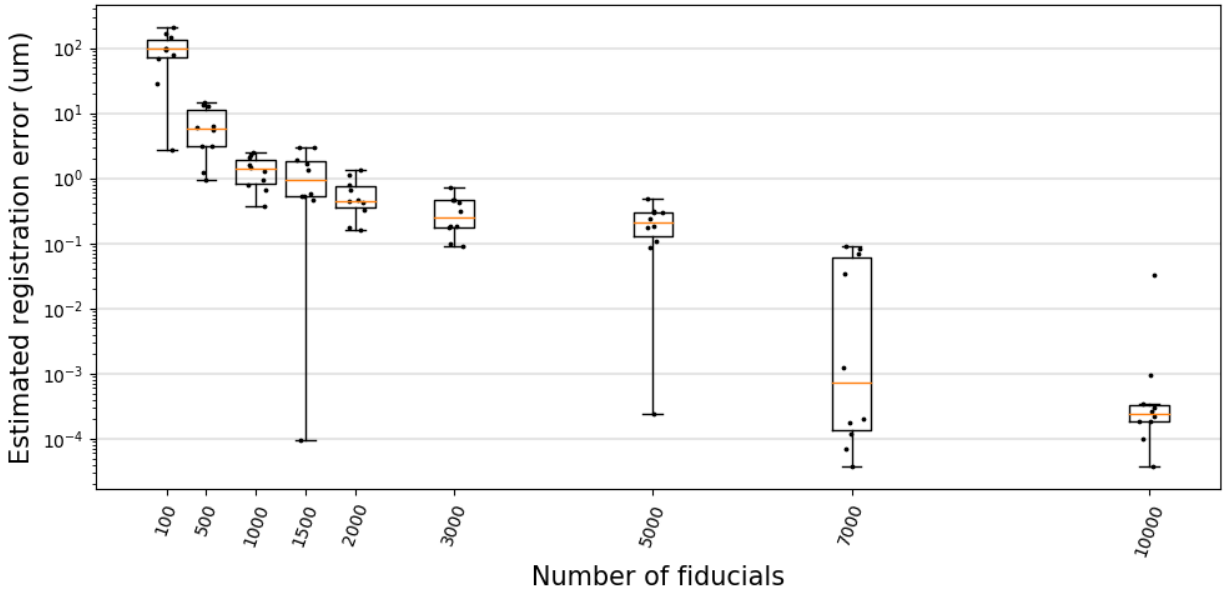
Supplementary Information

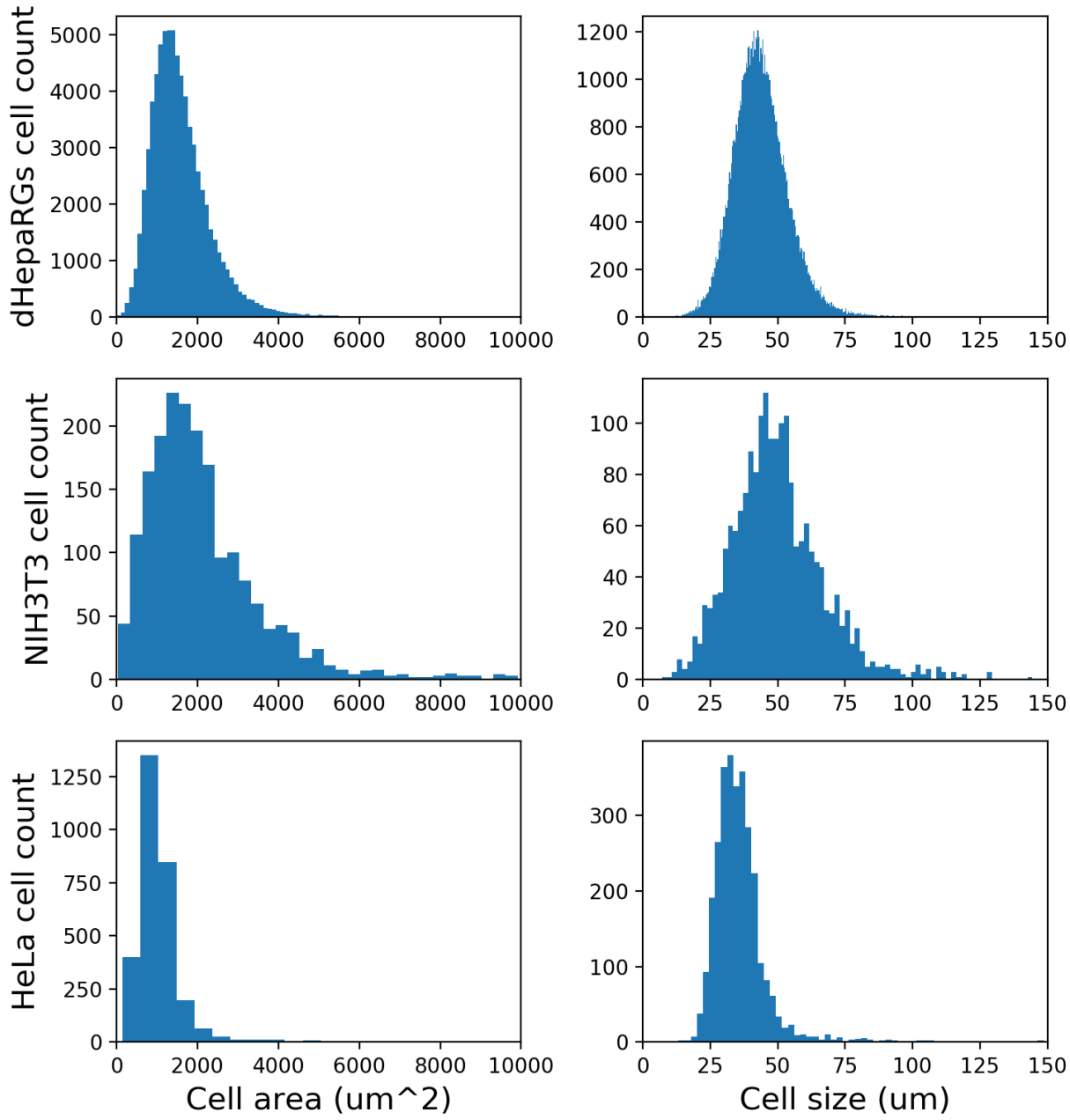


Supplementary Figure 1: Detailed workflow of the SpaceM method, supplementing a brief presentation in **Figure 1**.



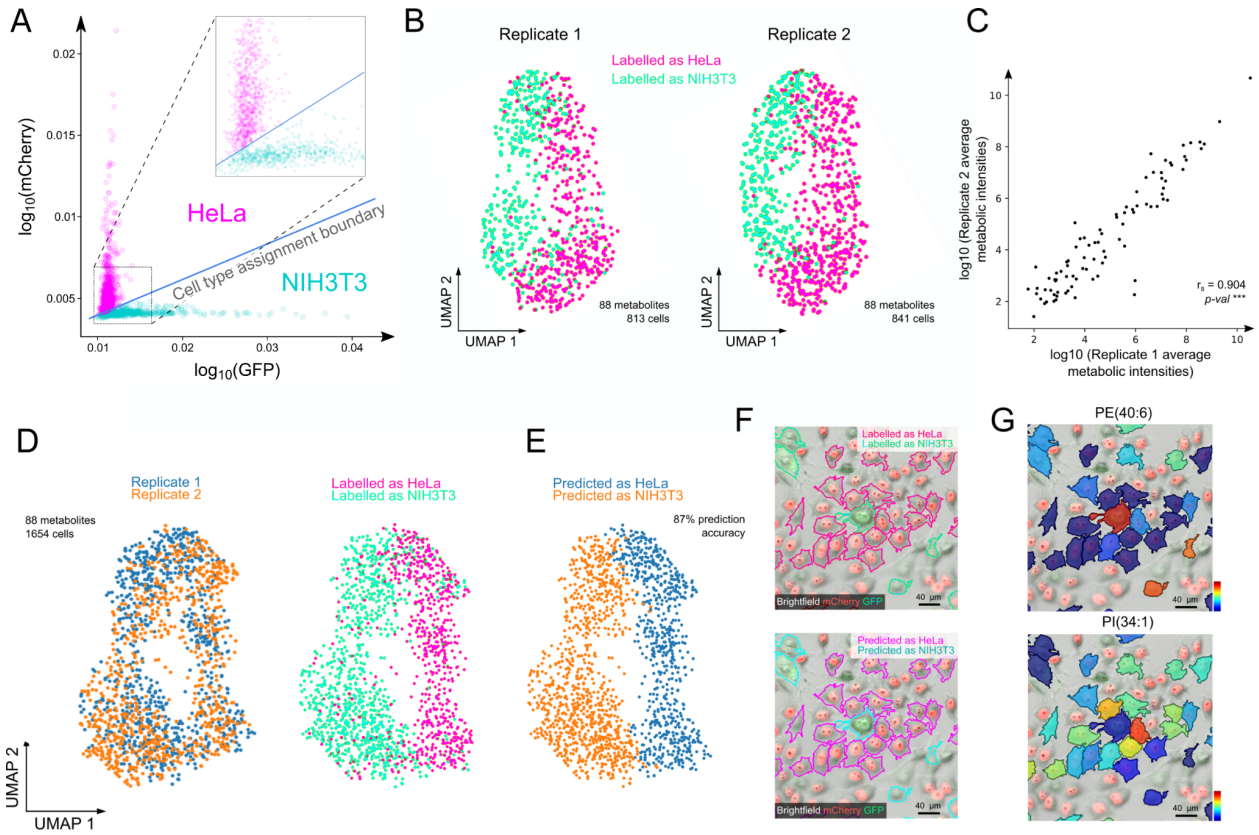
Supplementary Figure 2: **Estimated microscopy registration accuracy and its dependency on the number of fiducials.** As the MALDI matrix prevents detecting cells in the post-MALDI microscopy images, black pen marks are drawn on the glass slide providing features detectable before and after matrix application and fiducial points are automatically detected at the edges of those pen marks (see **Extended Data Figure 2** for a detailed illustration). This provides more than 400.000 individual points from both pre-MALDI and post-MALDI microscopy images which can be used for their registration. Unfortunately, the MALDI matrix also diffracts the light thus distorting post-MALDI microscopy images, preventing a perfect overlap of the fiducials in pre- and post-MALDI microscopy images. As diffraction is omnidirectional, it should theoretically not affect the registration accuracy. On the other hand, diffraction also prevents accurate feature correspondence between datasets, thus not allowing the computation of the registration accuracy. For this reason, we set out to compute the theoretical registration accuracy using the following approach. A coordinate transform was applied to fiducials from the

pre-MALDI image. Random subsets of both the transformed and the non-transformed fiducials were co-registered. The registration error was estimated as the Euclidean distance between the corresponding features between the non-transformed and the registered transformed fiducials. Our estimation shows that at least 3000 fiducials are required in order to systematically obtain sub-micron accuracies. The estimation process was repeated 10 times with different random samplings (Tukey box plots, center line: median; box limits: upper and lower quartiles; whiskers: 1.5x interquartile range). Following this recommendation, in the analyses presented in the paper, 5000 fiducials have been used to register the pre-MALDI and post-MALDI microscopy images.



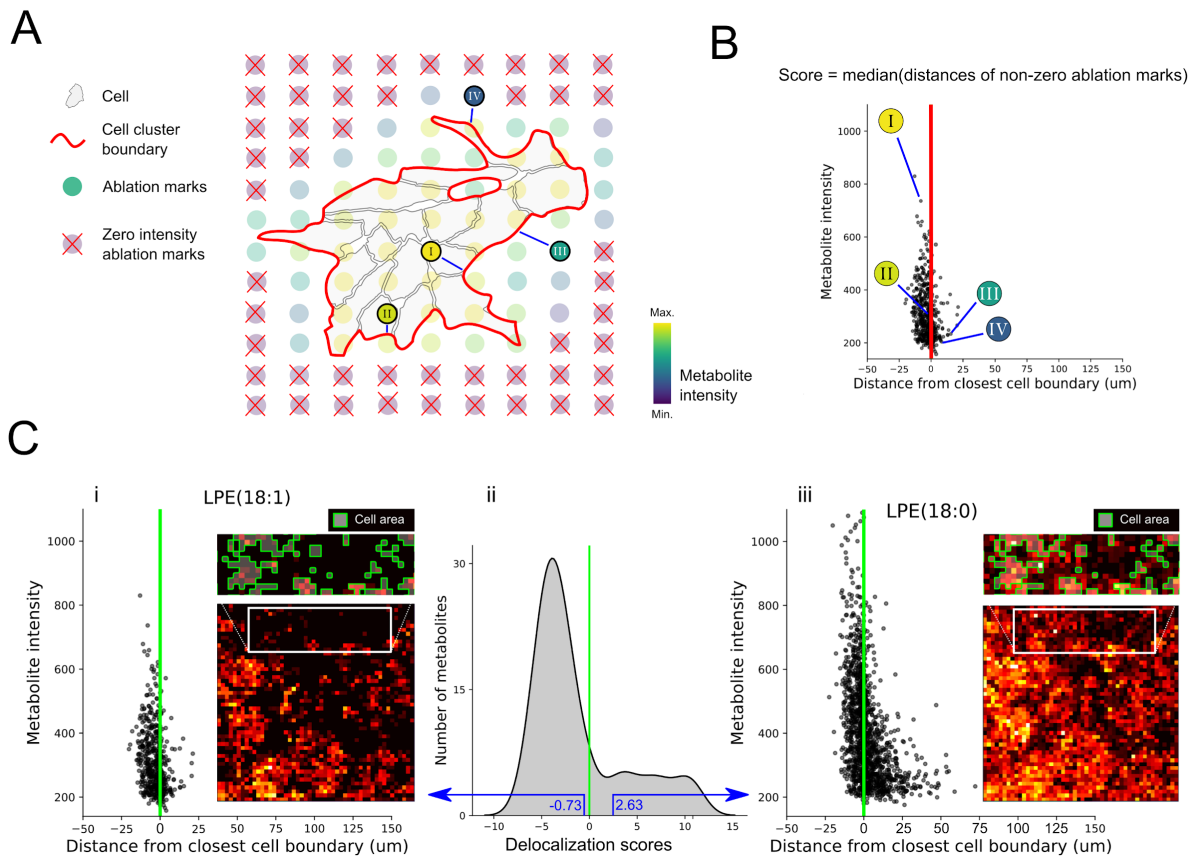
Supplementary Figure 3: Cellular areas and sizes of the different cell types analyzed by SpaceM. Cell area (first column) is defined as the number of pixels enclosed by the cell boundary estimated from cell segmentation and corrected by the pixel size (0.7um). Cell size (second column) is defined as the diameter of the equivalent circular area. The histograms show

data collected from dHepaRGs hepatocytes, NIH3T3 and HeLa cells at the first, second and third rows, respectively.



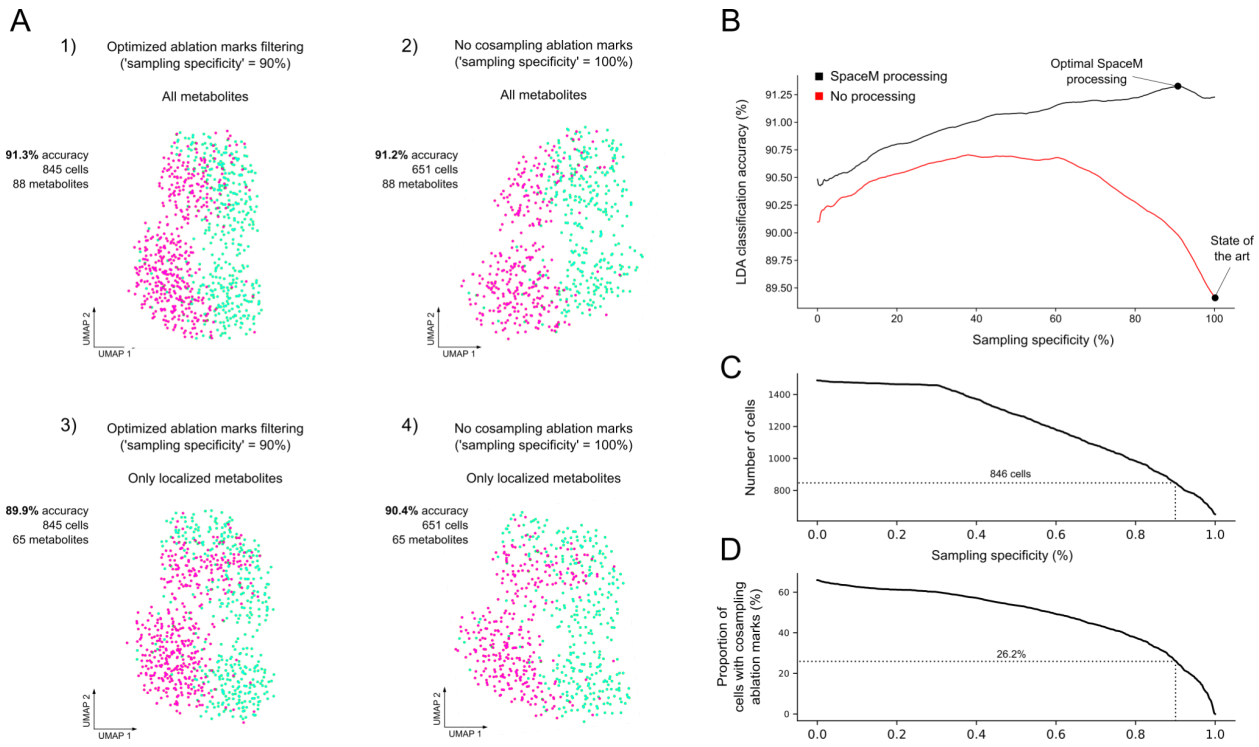
Supplementary Figure 4: **Overview of the SpaceM validation approach**. **A**: Assignment of the cell type for co-cultured HeLa and NIH3T3 cells was done by using a linear separating boundary between mCherry (for HeLa, shown in red) and GFP (for NIH3T3, shown in cyan) fluorescence intensities. **B**: UMAP visualization computed from 88 metabolite intensities of single cells from two independent replicates of co-cultured HeLa (magenta) and NIH3T3 (cyan) cells. The color coding is set as defined in A. The fact that both cell types are segregated from each other in the UMAP of both replicates demonstrates that SpaceM reproducibly captured the

extent of the biological variability present within each sample. **C:** Evaluation of the reproducibility of metabolite intensities across replicates providing a quantitative assessment of the analytical variability of the method. The averaged metabolite intensities across single cells from the two replicates have a Spearman correlation of 0.904 (two-tailed test *p*-value=2e-33, ***). **D:** Both replicates have been integrated using the batch balanced k nearest neighbours (BBkNN) method. UMAP visualization computed on the integrated replicates highlights their overlap and the conserved separation between cell types. **E**: The linear discriminant analysis (LDA) applied to the integrated dataset predicted cell types with an accuracy of 91.3%. The classification accuracy has been estimated by performing a 10 times-repeated stratified 10-fold cross validation with different randomization in each repetition (100 repetitions in total). Cells are color coded by their predicted cell types. **F**: SpaceM demonstrates single-cell resolution by correctly predicting the cell type of an NIH3T3 cell surrounded by HeLa cells. Top, the sampled cells are outlined according to their assigned cell type. Bottom, the same cells color coded according to their predicted cell type. Representative image from 2 independent experiments. **G**: Cells are color-coded by intensities of two metabolite markers of each cell type identified by SpaceM: top, PE(40:6), metabolite marker of NIH3T3 cells; bottom, PI(34:1), marker of HeLa cells. Representative image from 2 independent experiments.



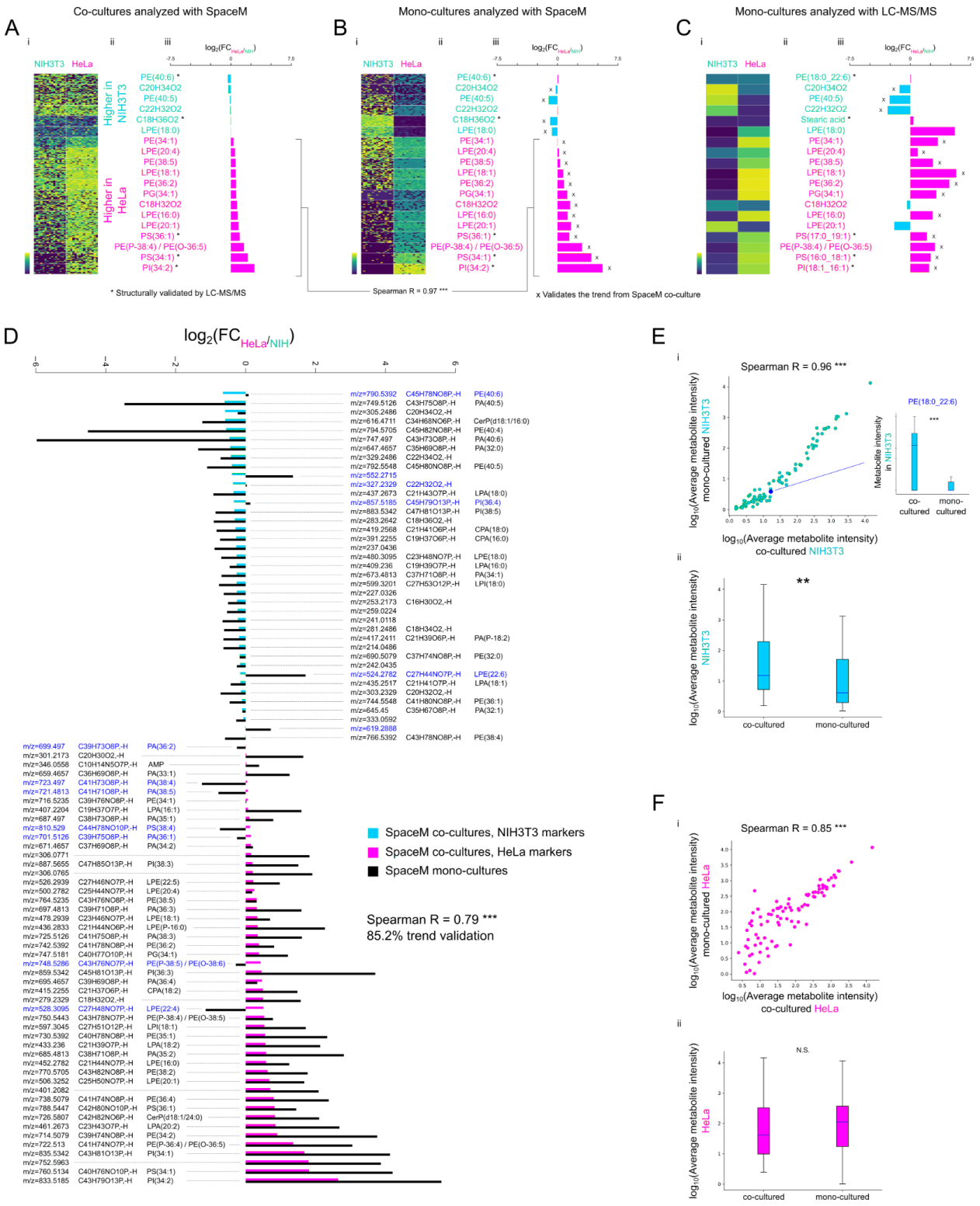
Supplementary Figure 5: **Constructing a diffusion profile for a metabolite and filtering out diffused metabolites. A:** Illustration of the strategy to construct a diffusion profile for a metabolite. Considering ablation marks in both the intracellular and the extracellular space, we estimate how far each metabolite diffuses from the nearest cell boundaries. The distance of ablation marks to the nearest cell boundary (highlighted in red) is set positive if the ablation mark is extracellular and negative if intracellular. Zero-intensity ablation marks are discarded from the analysis. **B:** The diffusion score is defined as the median of these distances. If negative, the metabolite is considered as well-localized, if positive, it is considered as diffused. **C:** The scatterplots for ablation marks for lysophosphoethanolamines LPE(18:1) (**i**) and LPE(18:0) (**iii**) exemplify metabolites with different diffusion scores. LPE(18:0) has been labelled as a diffused metabolite (diffusion score of 2.63) as multiple non-zero metabolite intensities were observed at

extracellular ablation marks. On the other hand, LPE(18:1) exemplifies a well-localized metabolite detected predominantly in the intracellular ablation marks only (diffusion score of -0.73). **ii**) A histogram of the diffusion scores for all metabolites showing that most of the metabolites are well-localized (65 out of 88 have delocalization scores below 0).



Supplementary Figure 6: **Evaluating how metabolite diffusion and co-sampling neighboring cells affects the cell type classification.** Both co-sampling neighboring cells and metabolite diffusion can, in principle, affect the single-cell resolution of SpaceM. To investigate this, we evaluated the impact of these two effects on the cell type classification of co-cultured human HeLa and mouse NIH3T3 cells (**Figure 1D**, **Supplementary Fig. 4**). **A: 1)** Cell type classification of metabolite profiles obtained from the optimal SpaceM processing (co-sampling ablation marks which have a sampling specificity of at least 90% are conserved) without filtering out diffused metabolites resulted in a classification accuracy of 91.3% for 846 cells. **2)** Removing all co-sampling ablation marks (sampling specificity set to a minimum of 100%) worsened the accuracy by 0.1% (91.2%) and reduced the number of cells to 651. **3)** Discarding diffused metabolites did not result in a more accurate classification (89.9%). **4)** Discarding both co-sampling ablation marks and diffused metabolites did not improve the classification (90.4%). **B:** Estimation of the optimal sampling specificity threshold. The cell type classification accuracy

is reported as a function of the sampling specificity for metabolite profiles that have been processed by SpaceM as detailed in **Extended Data Fig. 4** (black line) or for the integrated metabolite profiles (the metabolite intensity of a cell is the sum of intensities from all touching ablation marks, magenta line). The metabolite profiles from the SpaceM processing reached an optimal classification accuracy of 91.3% at a sampling specificity of 90% and consistently outperformed the classifications from the integrated profiles. Additionally, the integrated metabolic profiles with no cosampling ablation marks (sampling specificity of 100%) which is the state of the art in the field [39] provided the poorest accuracy of 89.26%, thus demonstrating the advantage of the SpaceM processing. **C:** Number of cells as a function of the sampling specificity (846 cells at a sampling specificity of 90%). **D:** Proportion of cells with at least one co-sampling ablation mark as a function of the sampling specificity (26.2% at a sampling specificity of 90%).

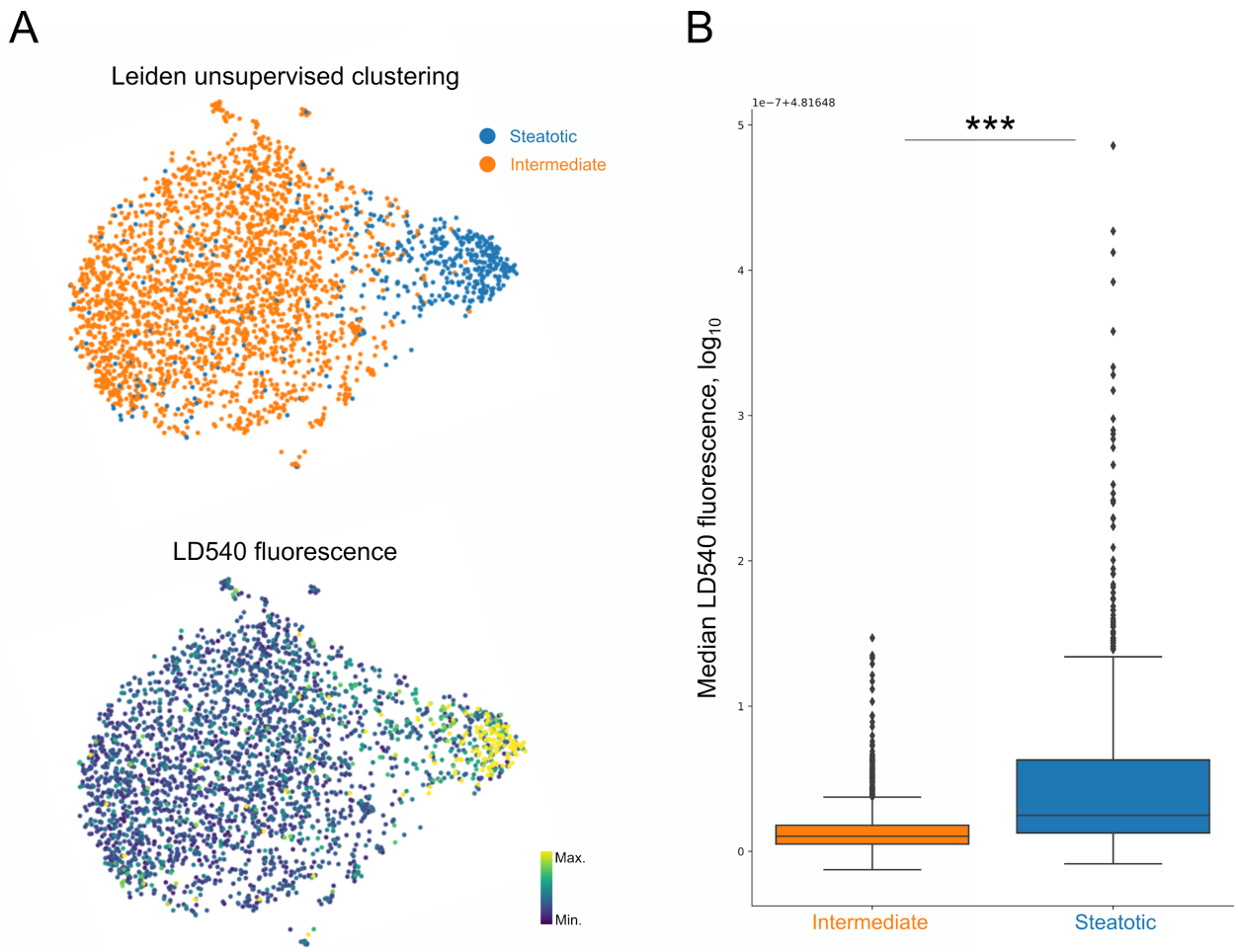


Supplementary Figure 7: Validation of the cell-type metabolite markers identified in the co-cultured HeLa and NIH3T3 cells by using individual mono-cultures analyzed either with

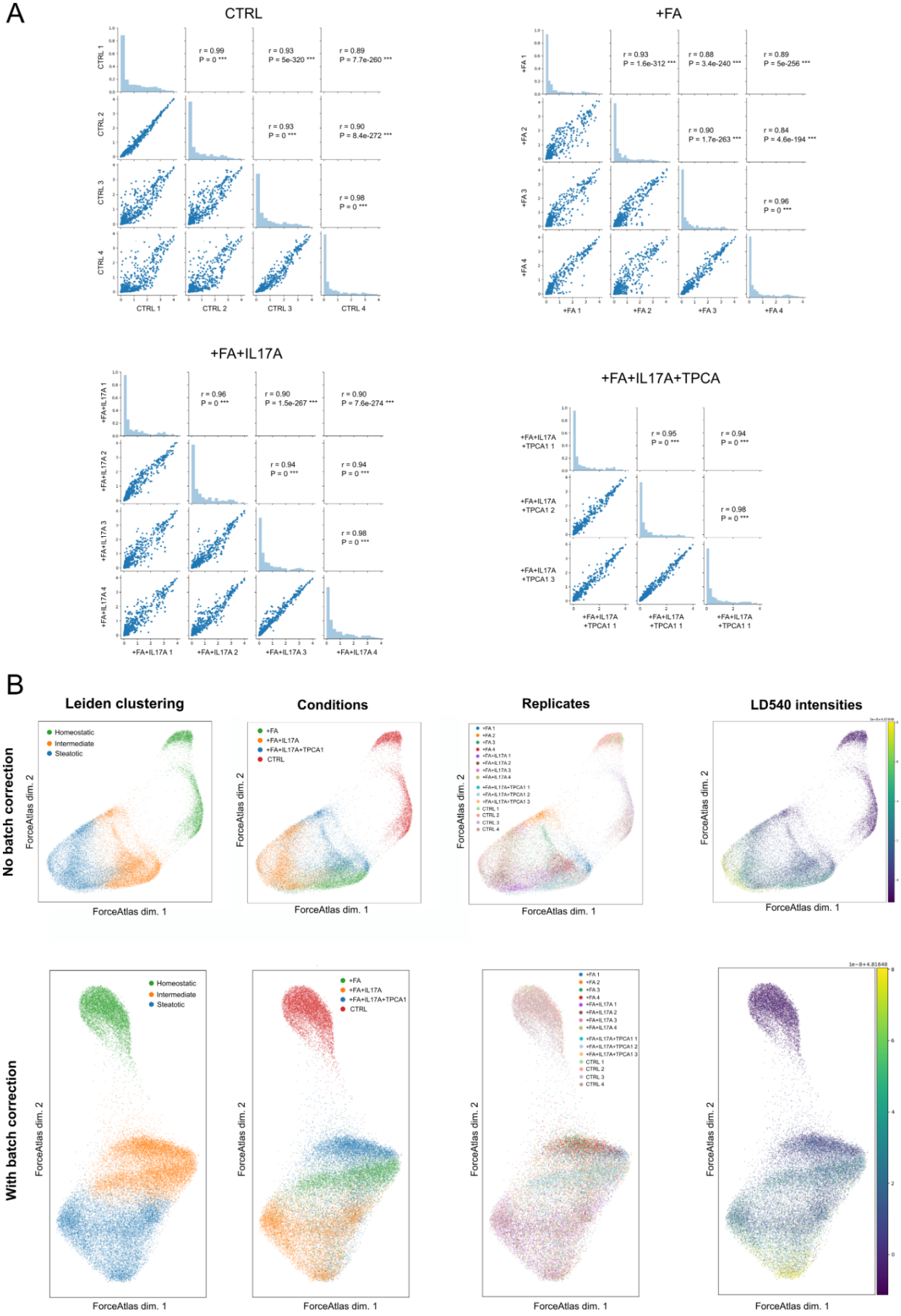
SpaceM or with LC-MS/MS. The markers presented in **A-C** have been identified by both single-cell analysis (SpaceM) and bulk lipidomics (LC-MS/MS). **A: i)** For each marker, we show intensities of 225 randomly selected single cells from each cell type. **ii)** The cell-type metabolic markers from unsupervised clustering on the co-cultured NIH3T3 and HeLa cells discovered using SpaceM (see details in **Supplementary Fig. 6**). The markers are ranked by their fold changes for the estimated cell types: from NIH3T3 top markers (top, cyan) to HeLa markers (bottom, red). Metabolites marked with a star (*) have been validated structurally with bulk LC-MS/MS in monocultures (NIH3T3 and HeLa cells cultured independently); see Supplementary Data S1 for more details. **iii)** The $\log_2$ fold change ($\log_2$(FC)) of the ratio between the average intensities in HeLa and the NIH3T3 cells for each metabolite marker, color coded as in **ii)**. For each cell type, single-cell metabolite intensities were normalized by dividing by the sum intensity of the metabolic profile. **B:** Same analysis as presented in **A** performed on independent mono-cultures of both cell types analyzed by SpaceM. The Spearman correlation between the HeLa marker fold changes obtained from co-cultures and mono-cultures was 0.97 (two-tailed test *p*-value=0, ***). This indicates an almost perfect match of the relative intensities of these metabolites in both cell types from the co-cultures and mono-cultures. In **iii)** the metabolites marked with a cross (x) share the same trend, defined as the log2(FC) sign, as in the co-cultures and therefore validate the specificity of these metabolite markers for their cell type. From the 19 identified metabolite markers across all modalities (SpaceM and LC-MS/MS), 17 markers showed the same trend. **C:** Same approach as **B** but with the bulk LC-MS/MS marker intensities . **i)** The markers intensities from LC-MS/MS (n=5 per cell type) are illustrated for each mono-culture. **ii)** For every marker discovered by SpaceM denoted with a star, the structural validation was performed by using LC-MS/MS and the validated name is displayed. **iii)** From the

19 identified metabolite markers, 15 were validated using LC-MS/MS. **D**: Comparison of the fold changes for all metabolites obtained with SpaceM (88 metabolites) when comparing the cell types in the co-cultures (cyan bars for NIH3T3 markers, magenta bars for HeLa markers) and in the mono-cultures (black bars). From 88 metabolites, 75 show the same trends (85.2%) with a Spearman correlation between the co-culture and mono-culture fold changes values of 0.79 (two-tailed test *p*-value=4.7e-20, ***). This demonstrates that the relative metabolite intensities obtained between cell types are highly consistent between the co-cultures and the mono-cultures. The putative MS1 annotations of the metabolic profile are displayed and are reported in detail in Supplementary Data S2. The metabolites for which the fold change trend across cell types has not been validated between the mono-and co-cultures are highlighted in blue. In **E i)**, we obtained a Spearman correlation of 0.96 (two-tailed test *p*-value=2.2e-48, ***) between the average metabolite intensities across the NIH3T3 cells in both culture conditions, further supporting the relative quantification consistency of SpaceM regardless of the culturing condition. The absolute levels of the marker PE(18:0_22:6) are highlighted (blue) in both culture conditions, showing a significantly higher intensity in the co-cultures (two-tailed independent *t*-test *p*-value=2.7e-194, ***). This explains why PE(18:0_22:6) was selected as a NIH3T3 marker in the co-culture but failed to be validated as such in the mono-cultures by both SpaceM and LC-MS/MS. This difference could be explained by a metabolic adaptation of NIH3T3 cells to the co-culturing with the HeLa cells, where this metabolite would be significantly more accumulated in the NIH3T3 cells. **ii)** The comparison of the average metabolite intensities of NIH3T3 cells in both culture conditions (1218 and 692 cells in mono- and co-cultures, respectively) shows that not only PE(18:0_22:6) but the overall absolute metabolic intensities were significantly higher in co-cultures than in mono-cultures (two-tailed independent *t*-test

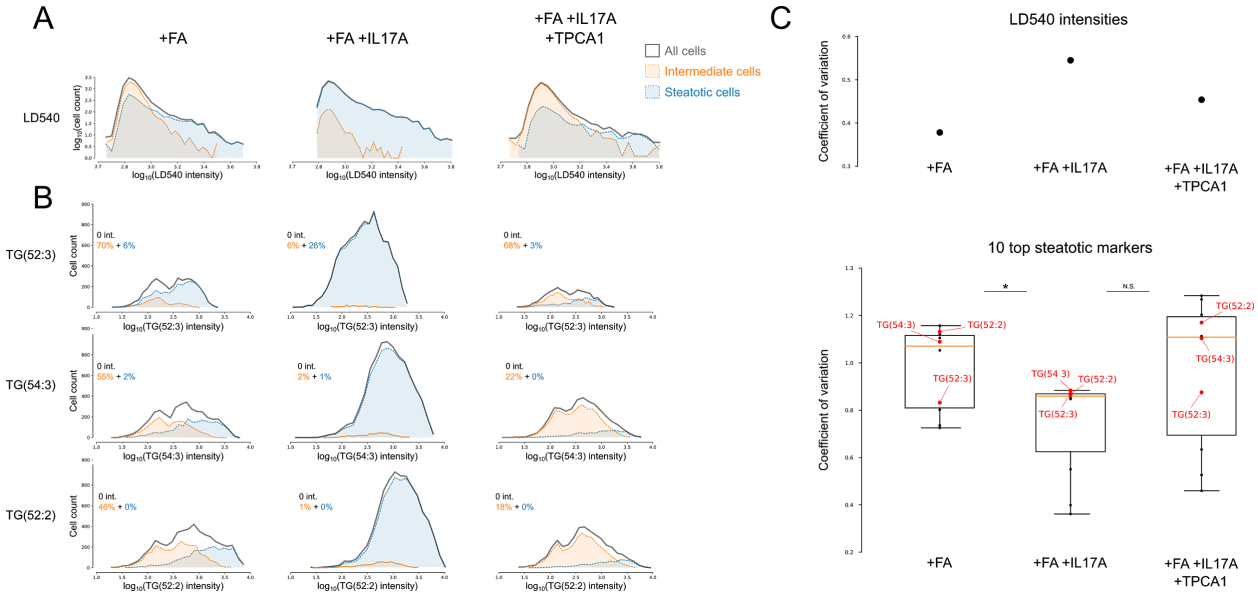
p-value=0.005, **), expanding the metabolic adaptation hypothesis to a larger fraction of the detected metabolites. Tukey box plots with center line: median; box limits: upper and lower quartiles; whiskers: 1.5x interquartile range. **F**: **i)** The relative reproducibility of the average metabolite intensities across cell culture conditions is lower for the HeLa cells than the NIH3T3 with a Spearman correlation of 0.85 (two-tailed test *p*-value=9e-26, ***). **ii)** The overall metabolite intensities of HeLa cells were not significantly different between the mono- and co-cultures (1999 and 962 cells, respectively, two-tailed independent *t*-test *p*-value=0.29, n.s.), although the median metabolite profile intensity is lower in the co-cultures. Tukey box plots with center line: median; box limits: upper and lower quartiles; whiskers: 1.5x interquartile range. This finally suggests the potential metabolic adaptation of HeLa cells to the exposition to NIH3T3 cells and this adaption is more pronounced in NIH3T3 cells.



Supplementary Figure 8: Identification of the subpopulation of steatotic hepatocytes among dHepaRG cells stimulated with palmitic and oleic acids. A: Unsupervised Leiden clustering of single-cell intensities of 740 detected metabolites resolved two subpopulation (76% and 24% of the cell population) characterized by statistically different amounts of lipid droplets (LD) measured with the single-cell median LD540 fluorescence. **B**: Box plots for the LD540 fluorescence intensities for each subpopulation (2394 and 446 cells in the intermediate and the steatotic states, respectively; two-tailed independent *t*-test *p*-value=*4.8e-146,* ***). Tukey box plots with center line: median; box limits: upper and lower quartiles; whiskers: 1.5x interquartile range.

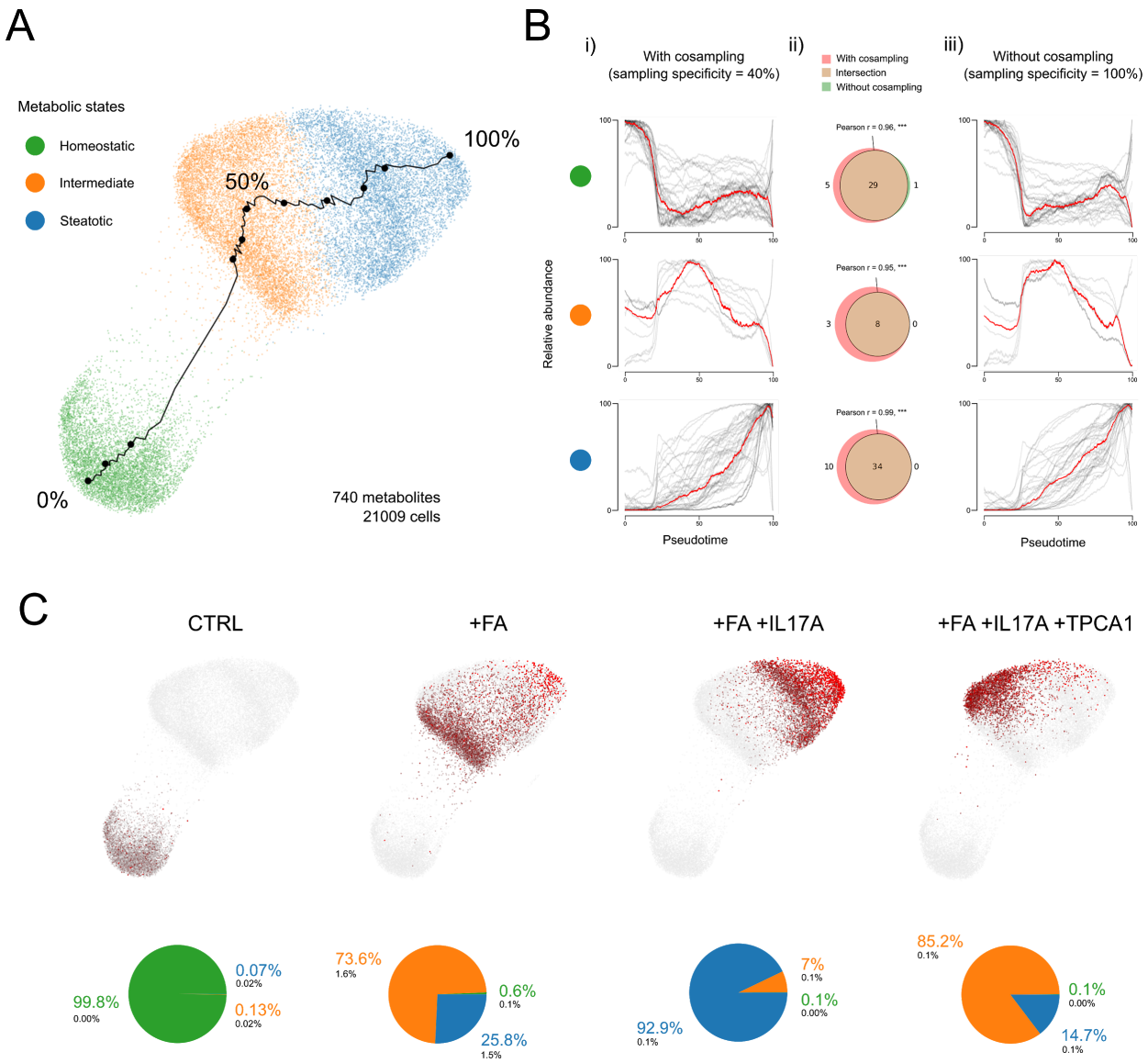

A
CTRL
CTRL 1
CTRL 2
CTRL 3
CTRL 4
r = 0.99
P = 0 ***
r = 0.93
P = 5e-320 ***
r = 0.89
P = 7.7e-260 ***
r = 0.93
P = 0 ***
r = 0.90
P = 8.4e-272 ***
r = 0.98
P = 0 ***
+FA
+FA 1
+FA 2
+FA 3
+FA 4
r = 0.93
P = 1.6e-312 ***
r = 0.88
P = 3.4e-240 ***
r = 0.89
P = 5e-256 ***
r = 0.90
P = 1.7e-263 ***
r = 0.84
P = 4.6e-194 ***
r = 0.96
P = 0 ***
+FA+IL17A
+FA+IL17A 1
+FA+IL17A 2
+FA+IL17A 3
+FA+IL17A 4
r = 0.96
P = 0 ***
r = 0.90
P = 1.5e-267 ***
r = 0.90
P = 7.6e-274 ***
r = 0.94
P = 0 ***
r = 0.94
P = 0 ***
r = 0.98
P = 0 ***
+FA+IL17A+TPCA
+FA+IL17A +TPCA1 1
+FA+IL17A +TPCA1 2
+FA+IL17A +TPCA1 3
r = 0.95
P = 0 ***
r = 0.94
P = 0 ***
r = 0.98
P = 0 ***
B
Leiden clustering
Conditions
Replicates
LD540 intensities
No batch correction
With batch correction
Homeostatic
Intermediate
Steatotic
+FA
+FA+IL17A
+FA+IL17A+TPCA1
CTRL
ForceAtlas dim. 1
ForceAtlas dim. 2

Supplementary Figure 9: Analysis of the sample-sample reproducibility for stimulated dHepaRG hepatocytes and the effect of the batch correction. A: For each condition, we show a correlogram showing reproducibility of single-cell metabolite intensities across replicates. One dot represents one metabolite, with its x-y position within every plot corresponding to the mean intensities across all cells for two selected replicates. Spearman correlations between replicates is equal to (mean +/- std): 0.94 +/- 0.04 for CTRL, 0.9 +/- 0.04 for +FA, 0.94 +/- 0.03 for +FA+IL17A, and 0.96 +/- 0.02 for +FA+IL17A+TPCA1, all correlations are significant with two-tailed test *p*-values < 0.001, ***. Exact *p*-values are noted in the figure. **B**: Top row, manifold visualization on samples with no batch correction applied to the replicates. Bottom row: The same analysis but after the batch correction with the *combat* batch correction method from Scanpy (v0.7.18).



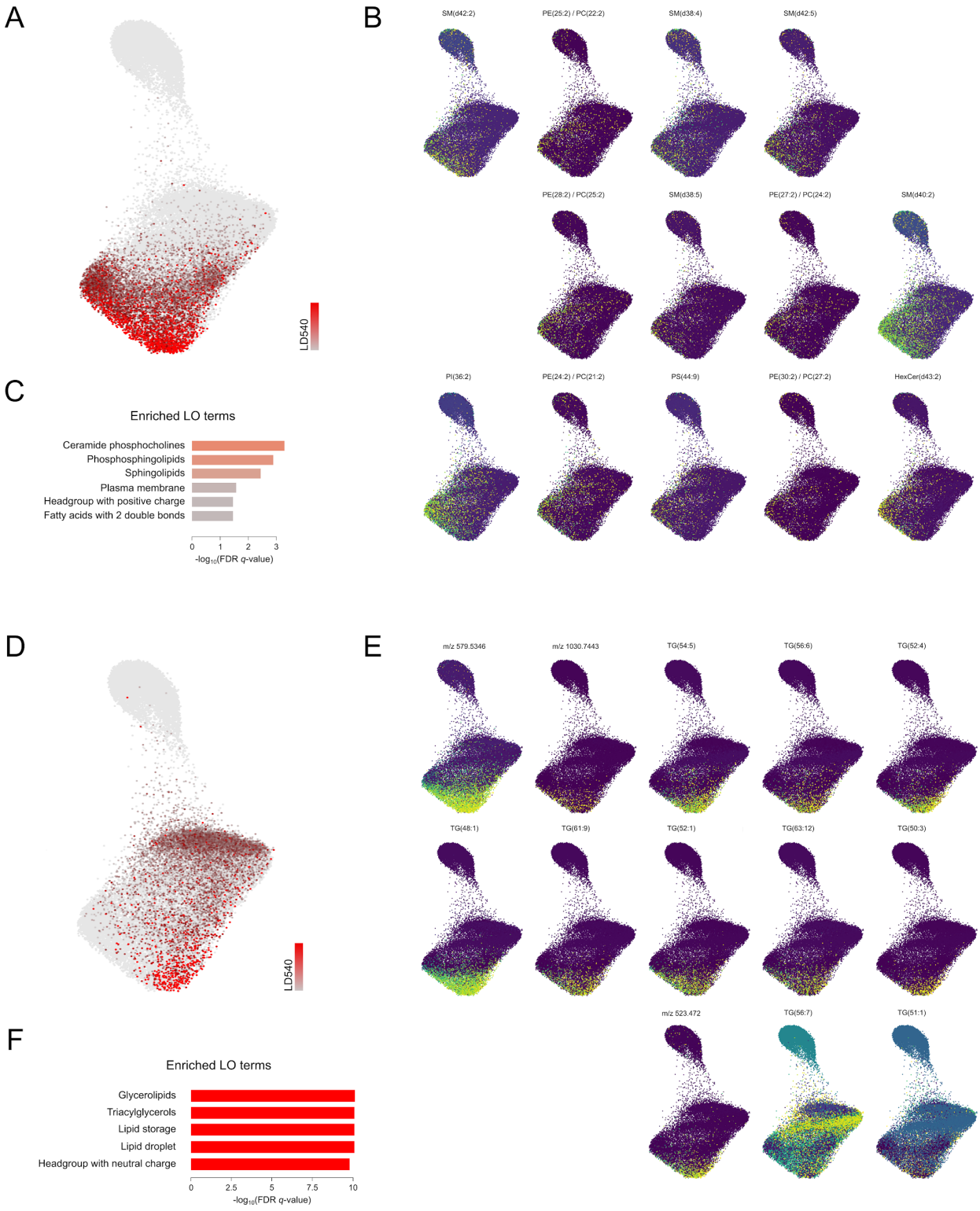
Supplementary Figure 10: Inflammatory signalling increases hepatocyte heterogeneity in lipid droplet accumulation but reduces heterogeneity in steatotic markers. Distribution of LD540 cell intensities (**A**) and representative steatotic markers: TG(52:2), TG(54:3) and TG(52:2) (**B**) for each condition: +FA, +FA+IL17A, +FA+IL17A+TPCA1 (first, second and third columns, respectively). The intermediate (orange) and steatotic (blue) states as well as all hepatocytes (bold dark gray) distributions are shown separately. The cells with 0 intensities are not shown for clarity, instead, their proportion to the whole population size is displayed for both metabolic states, with the same color-coding. We show in **Figure 2** that stimulating hepatocytes with fatty acids profoundly affected their metabolism, which was manifested by the emergence of a steatotic subpopulation with significantly higher levels of lipid droplets and neutral lipids. Such heterogeneity is not reflected in the LD540 intensities (**A**) but is strongly reflected in individual steatotic markers by their bimodal profile (**B**, first column). In **Figure 3**, we show that inducing inflammation with IL-17A abolishes the intermediate sub-population, driving almost all cells to the steatotic metabolic state, a phenomenon which is visible in the unimodality of the neutral lipid profiles (**B**, second column). Conversely, inhibiting NF-κB signalling with TPCA-1 did not result in a uniform response to IL-17A, but rather showed a clear bimodal distribution, as in the +FA condition (**B**, third column). The bimodality observed in the +FA and +FA+IL17A+TPCA1 conditions is partly explained by the intermediate and steatotic metabolic states which systematically reside at lower and higher lipid intensities, respectively. Although this bimodality is also reflected in the intermediate states, it could be explained in two ways. A miss-classification of steatotic cells as intermediate by the unsupervised clustering algorithm could explain why the intermediate subpopulation contains cells with such high neutral lipid levels. Alternatively, the intermediate state bimodality can be due to an inherent variability

where some cells would accumulate a large amount of lipid droplets without being steatotic, thus reflecting a potential variability in both lipid droplet amount and composition. **C:** The coefficient of variation (C.V., defined as the ratio of the biased standard deviation to the mean, used to quantify cell-cell heterogeneity) for LD540 (top) and top-ten steatotic markers for each condition. Cells with zero lipid intensities have been excluded from the computation of the coefficient of variation to avoid biasing the results (9308, 8859 and 4955 cells for the +FA, +FA+IL17A and +FA+IL17A+TPCA1 conditions, respectively). Representative neutral lipids shown in **B** are highlighted in red. The heterogeneity of top-ten markers is different between the conditions (two-tailed independent *t*-test *p*-values of CV values for top-ten markers are: 0.01 ** for +FA vs +FA+IL17A, 0.07 $^{N.S.}$ for +FA+IL17A vs +FA+IL17A+TPCA1). These results indicate that IL-17A-driven inflammation decreases the cell-cell heterogeneity of steatotic markers (neutral lipids specifically) in a process mediated by NF-κB signalling. Surprisingly, lipid droplet levels (as quantified by LD540) show the opposite trend, where the IL-17A-stimulated cells have the highest cell-cell heterogeneity. This can be explained by an increase of hepatocytes with higher amounts of lipid droplets in the +FA+IL17A condition (shown in **A**). Consistently with our previous hypothesis proposing that the intermediate state is heterogenous in lipid droplet composition, we stipulate that the steatotic hepatocytes, on the contrary, have a more homogeneous lipid droplet composition, despite presenting a wider range of abundance of lipid droplets. Tukey box plots with center line: median; box limits: upper and lower quartiles; whiskers: 1.5x interquartile range.



Supplementary Figure 11: Impact of excluding co-sampling ablation marks on the discovered metabolic states of hepatocytes. A: Removing co-sampling ablation marks reduced the number of cells to 21,009 cells (compared to 29,738 cells before this filtration as reported in **Figure 3**). The PAGA manifold reproduced the general organization observed in **Figure 3B** and the unsupervised clustering reproduced the same three metabolic states. The depicted trajectory that spans across the manifold (black line) has been reconstructed on the new metabolic profiles.

B: Comparison of the metabolic state markers identified from the metabolic profiles with and without co-sampling ablation marks. **i)** The metabolic markers from metabolic profiles computed with co-sampling ablation marks (sampling specificity threshold of 40%, **i)** same as **Figure 3D**) or without (**iii**) in function of the pseudotime for each metabolic state (green, homeostatic; orange, intermediate; blue, steatotic). **ii)** Overlap of metabolic markers identified from metabolic profiles computed with and without co-sampling ablation marks. A near-perfect Pearson correlation in each metabolic state was obtained between markers that have been identified from both approaches (r=0.96, two-tailed test *p-value*=2.5e-195 ***;, r=0.95, *p*-value=0 ***;, r=0.99, *p*-value=2.3e-311 *** for the homeostatic, intermediate and steatotic metabolic states, respectively). **C)** The distribution of each condition across the manifold as well as ratios of metabolic states present in each condition are almost identical to the ones identified from metabolic profiles that included co-sampling ablation marks (**Figure 3C**). The difference in metabolic states ratios between the two approaches are depicted in black under each ratio in the pie charts and has a mean value of 0.3%. Altogether, the almost perfect correlations of metabolic state marker trends as a function of the manifold pseudotime and the nearly identical metabolic state ratios within each hepatocytes culture condition highlight the minimal impact of systematically excluding co-sampling ablation mark from the SpaceM processing. Instead, setting a sampling specificity threshold (40% in this case) yields the same results at the benefit of delivering more cells thus providing a higher statistical power.


A
LD540
B
SM(d42:2)
PE(25:2) / PC(22:2)
SM(d38:4)
SM(d42:5)
PE(28:2) / PC(25:2)
SM(d38:5)
PE(27:2) / PC(24:2)
SM(d40:2)
PI(36:2)
PE(24:2) / PC(21:2)
PS(44:9)
PE(30:2) / PC(27:2)
HexCer(d43:2)
C
Enriched LO terms
Ceramide phosphocholines
Phosphosphingolipids
Sphingolipids
Plasma membrane
Headgroup with positive charge
Fatty acids with 2 double bonds
0
1
2
3
-log10(FDR q-value)
D
LD540
E
m/z 579.5346
m/z 1030.7443
TG(54:5)
TG(56:6)
TG(52:4)
TG(48:1)
TG(61:9)
TG(52:1)
TG(63:12)
TG(50:3)
m/z 523.472
TG(56:7)
TG(51:1)
F
Enriched LO terms
Glycerolipids
Triacylglycerols
Lipid storage
Lipid droplet
Headgroup with neutral charge
0
2.5
5
7.5
10
-log10(FDR q-value)

Supplementary Figure 12: Difference between the "inflamed" and "benign" steatotic hepatocytes. We compared the "inflamed" steatotic hepatocytes (+FA+IL17A cells from the steatotic state) with "benign" steatotic hepatocytes (+FA cells from the steatotic state). **A:** "Inflammatory" steatotic cells highlighted on the PAGA plot and color coded by their LD540 intensities. **B:** Metabolic markers of the "inflamed" steatotic cells vs. "benign" steatotic cells. For each marker, its single-cell intensities are shown on the PAGA plot. **C**: The results of the Lipid Ontology enrichment for the markers from (B). **D-F:** The same results as in **A-C** but for the "benign" steatotic cells and their markers.