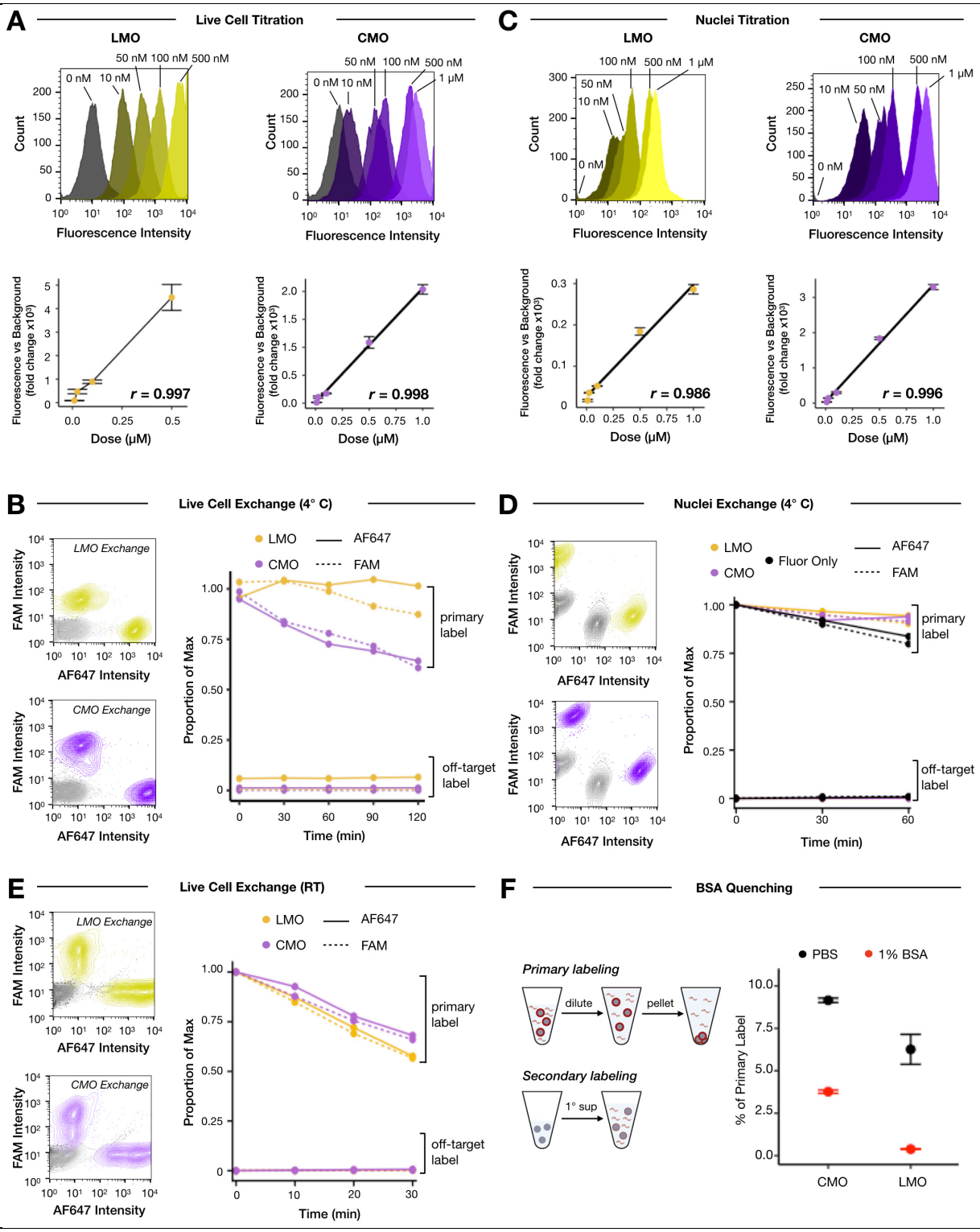



**Supplementary Figure 1**

Flow cytometry using fluorophore-conjugated MULTI-seq barcode oligonucleotide probes demonstrates robust LMO and CMO labeling efficiency on living cells and nuclei, as well as label stability over time and LMO quenching with BSA, related to Figure 1

(A) Live-cell LMO (gold) and CMO (purple) labeling efficiency varies predictably across a titration curve of anchor and co-

anchor LMO/CMO concentrations. Qualitative trends shown with histograms (top) are supported by regression analyses (bottom) demonstrating technical reproducibility and linear relationship between LMO/CMO concentration and fluorescence abundance. n = 10,000 events/sample. Data represented as mean ± SEM over 3 experimental replicates.

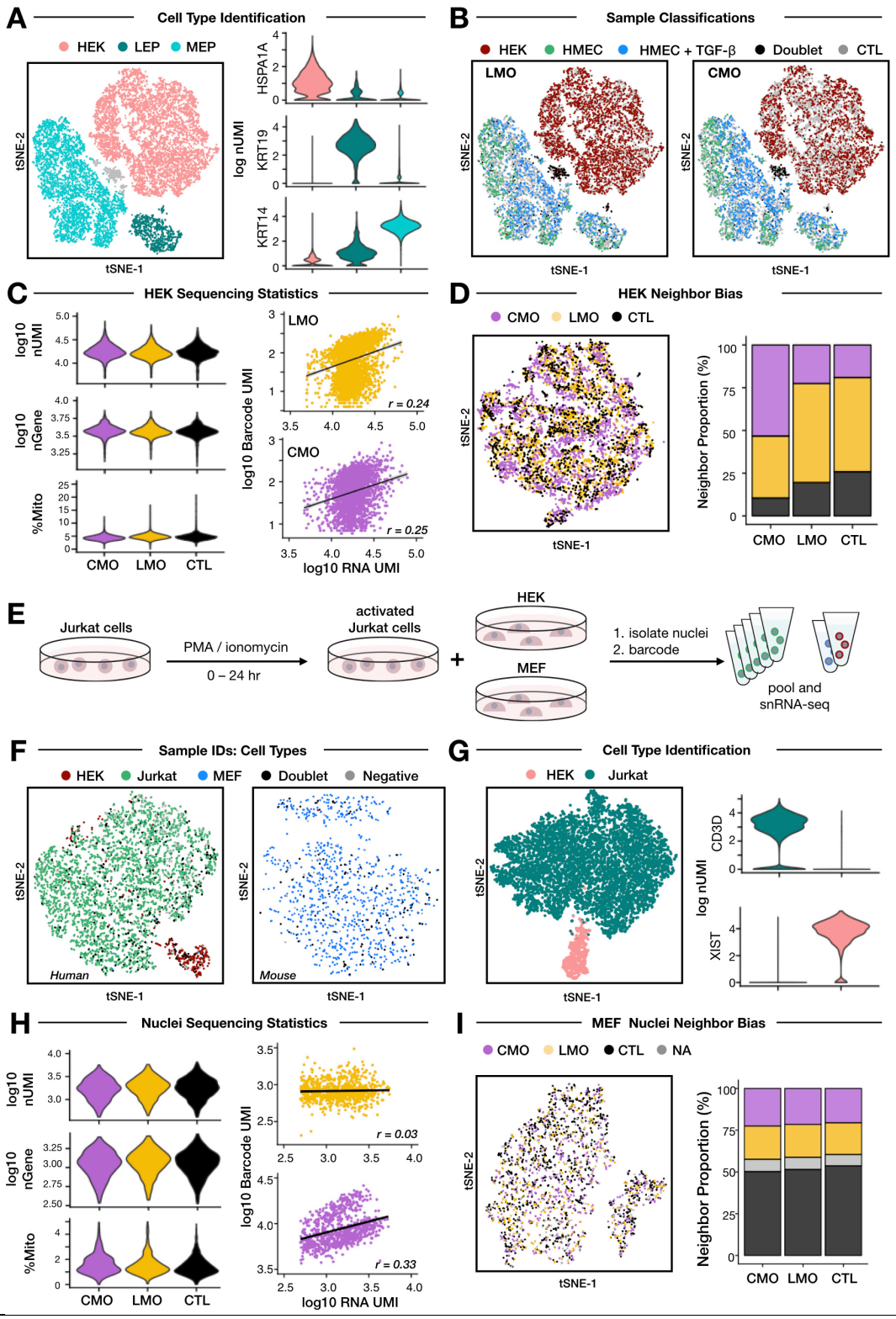
(B) Time-course analysis of LMO and CMO scaffold loss and exchange on ice following mixing of live cell populations labeled with either AF647- or FAM-conjugated barcode probes. Qualitative trends (contour plots, left) document how LMO or CMO labeled cells maintain fluorescence signal over unlabeled control cells (grey) over time. Quantitative analysis (right) illustrates how LMO scaffolds more stably embed in the plasma membrane relative to CMO scaffolds, although sample-to-sample cross-talk is minimal. n = 10,000 events/time-point. Experiment was repeated 3 times with similar results.

(C) Same experiment as described in Fig. S1A, except with nuclei. n = 10,000 events/sample. Data represented as mean ± SEM over 3 experimental replicates.

(D) Same experiment as described in Fig. S1B, except with nuclei. Difference between LMO and CMO membrane residency kinetics does not occur during nuclear membrane labeling. n = 10,000 events/time-point.

(E) Same experiment as described in Fig. S1C, except at room temperature. The LMO advantage in label stability shown at 4 °C is lost at RT as both CMO (purple) and LMO (gold) labels decrease at similar rates. n = 10,000 events/time-point

(F) Live-cells were labeled with LMO or CMO at 200 nM and diluted with either PBS (black) or 1% BSA in PBS (red). The cells were pelleted and the supernatant was transferred to unlabeled cells to determine the labeling efficiency of remaining LMO or CMO label. Dilution with BSA leads to a decrease in supernatant labeling relative to dilution with PBS alone. n = 10,000 events/sample. Data represented as mean ± SEM over 3 experimental replicates.




**Supplementary Figure 2**

MULTI-seq with LMOs preserves endogenous gene expression during single-cell RNA sequencing, while both LMOs and CMOs are non-perturbative during single-nucleus RNA sequencing, related to Figure 1.

(A) Live-cell LMO (gold) and CMO (purple) labeling efficiency varies predictably across a titration curve of anchor and co-anchor LMO/CMO concentrations. Qualitative trends shown with histograms (top) are supported by regression analyses (bottom) demonstrating technical reproducibility and linear relationship between LMO/CMO concentration and fluorescence abundance. n = 10,000 events/sample. Data represented as mean ± SEM over 3 experimental replicates.

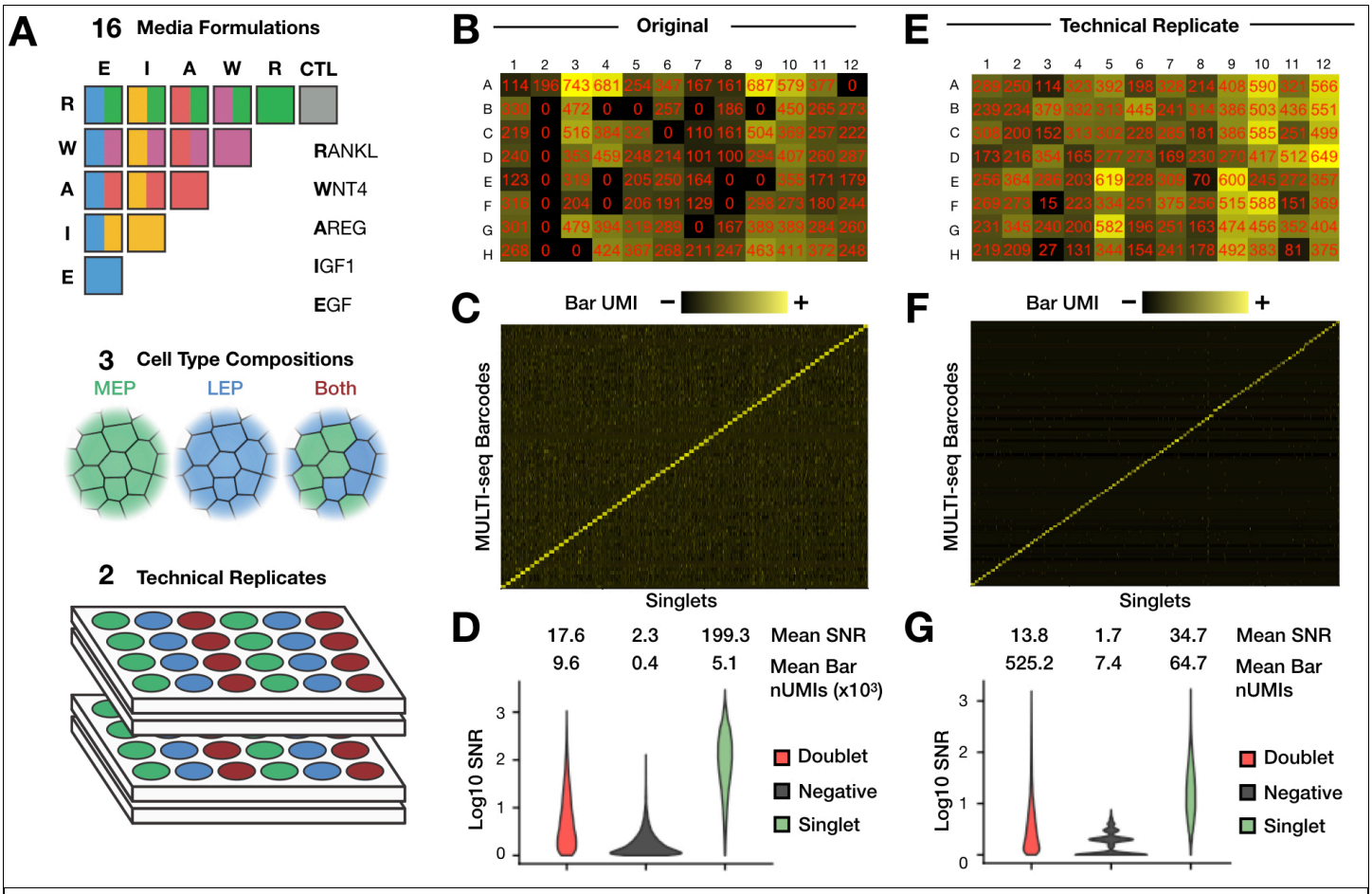
(B) Time-course analysis of LMO and CMO scaffold loss and exchange on ice following mixing of live cell populations labeled with either AF647- or FAM-conjugated barcode probes. Qualitative trends (contour plots, left) document how LMO or CMO labeled cells maintain fluorescence signal over unlabeled control cells (grey) over time. Quantitative analysis (right) illustrates how LMO scaffolds more stably embed in the plasma membrane relative to CMO scaffolds, although sample-to-sample cross-talk is minimal. n = 10,000 events/time-point. Experiment was repeated 3 times with similar results.

(C) Same experiment as described in Fig. S1A, except with nuclei. n = 10,000 events/sample. Data represented as mean ± SEM over 3 experimental replicates.

(D) Same experiment as described in Fig. S1B, except with nuclei. Difference between LMO and CMO membrane residency kinetics does not occur during nuclear membrane labeling. n = 10,000 events/time-point.

(E) Same experiment as described in Fig. S1C, except at room temperature. The LMO advantage in label stability shown at 4 °C is lost at RT as both CMO (purple) and LMO (gold) labels decrease at similar rates. n = 10,000 events/time-point

(F) Live-cells were labeled with LMO or CMO at 200 nM and diluted with either PBS (black) or 1% BSA in PBS (red). The cells were pelleted and the supernatant was transferred to unlabeled cells to determine the labeling efficiency of remaining LMO or CMO label. Dilution with BSA leads to a decrease in supernatant labeling relative to dilution with PBS alone. n = 10,000 events/sample. Data represented as mean ± SEM over 3 experimental replicates.




**Supplementary Figure 3**

HMEC MULTI-seq sample classification results and technical replicate, related to Figure 2.

(A) Schematic overview of 96-plex HMEC scRNA-seq analysis. 96 distinct HMEC cultures consisting of LEPs alone (blue), MEPs alone (green), or both cell types together (dark red) were grown in media supplemented with 15 distinct signaling molecules or signaling molecule combinations and one control.

(B) 96-well plate schematic overlaid with a heat map showing the number of cells assigned to each sample barcode group. Twenty samples — predominantly those arising from column 2 — were not represented in the original large-scale HMEC experiment due to technical error during sample preparation.

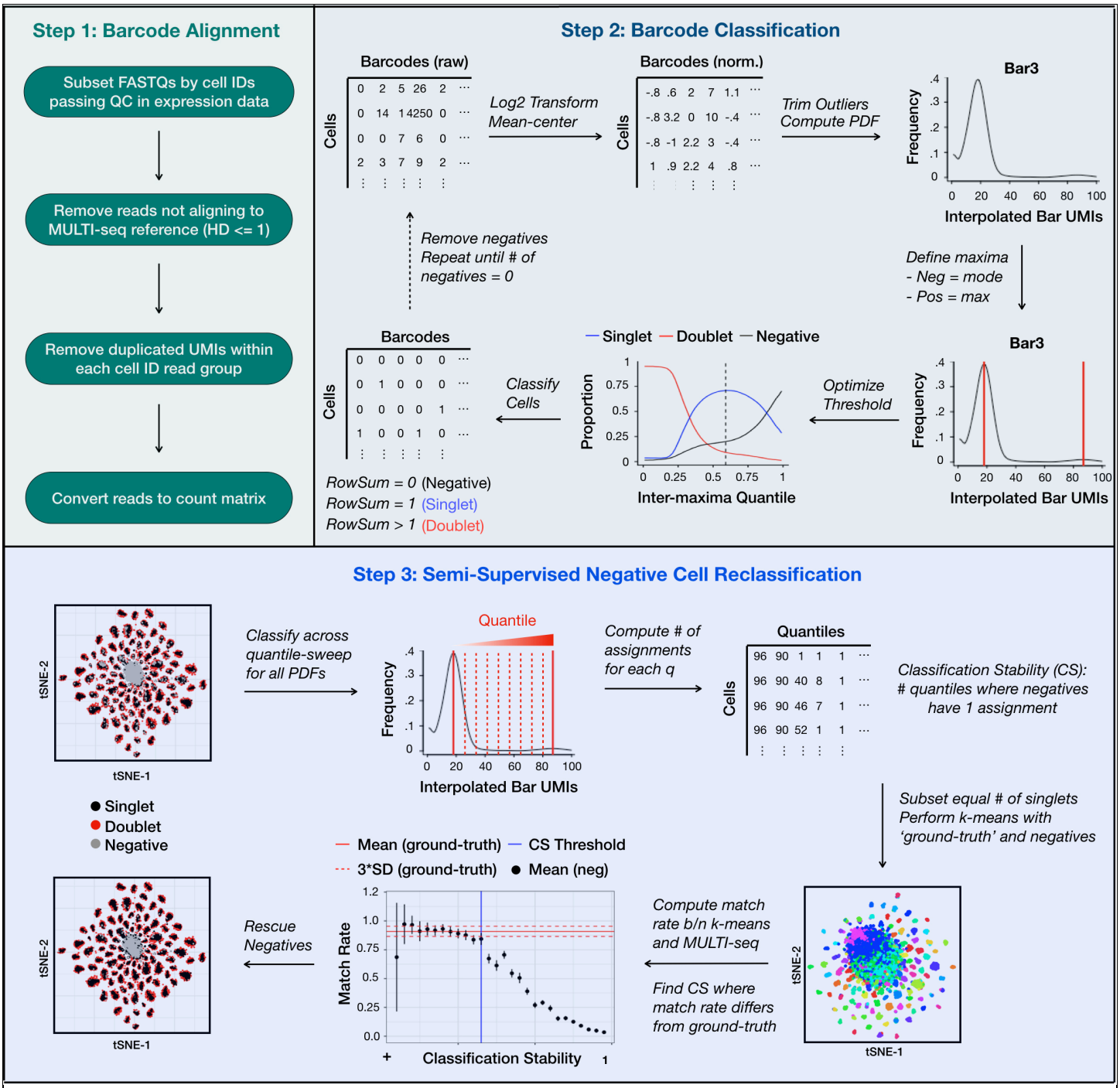
(C) Normalized barcode UMI heat map demonstrating that sample groups are predominantly associated with single MULTI-seq barcodes.

(D) Violin plots describing the barcode UMI SNR for negative cells, doublets, and singlets. n = 40,009 cells.

(E) Same analysis as described in Fig. S4A, except with the 96-plex HMEC technical replicate experiment. All samples were classified in the technical replicate.

(F) Same analysis as described in Fig. S4B, except with the 96-plex HMEC technical replicate experiment.

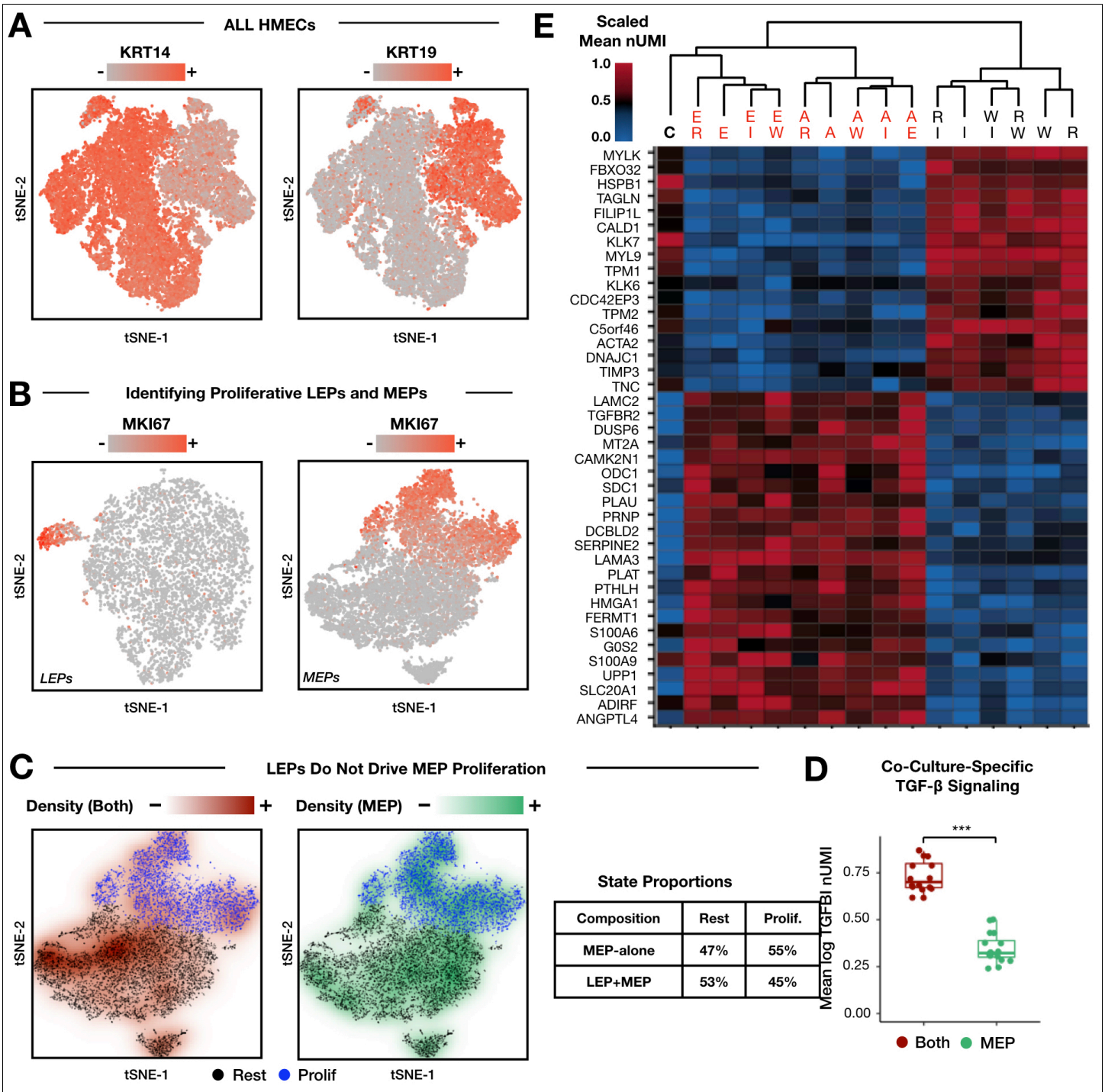
(G) Same analysis as described in Fig. S4C, except with the 96-plex HMEC technical replicate experiment. n = 48,091 cells




**Supplementary Figure 4**

MULTI-seq barcode pre-processing and sample classification workflows, related to Figure 2.

Results from the 96-plex HMEC experiment are used as representative examples for the barcode classification workflow. Results from the 96-plex technical replicate HMEC experiment as used as representative examples for the semi-supervised negative cell reclassification workflow. PDF = probability density function.




**Supplementary Figure 5**

Transcriptional signatures of co-culturing and signaling molecule stimulation in MEPs, related to Figure 2.

(A) Distributions of marker gene expression used to identify MEPs (KRT14) and LEPs (KRT19) in gene expression space. n = 25,166 cells.

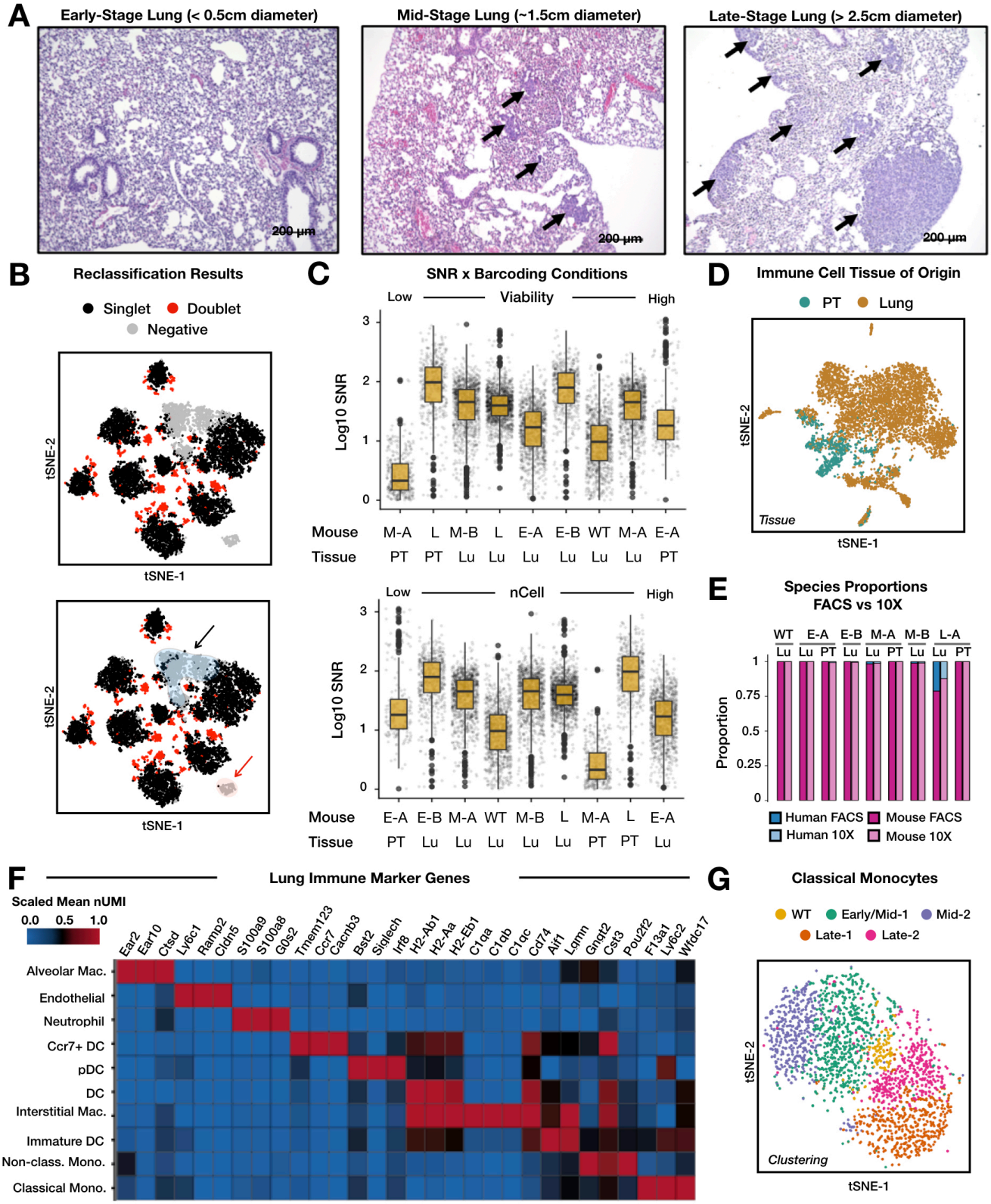
(B) Distributions of MKI67 expression in gene expression space for LEPs (left) and MEPs (right). MKI67 enrichment was used as a proxy for distinguishing proliferative and resting LEPs and MEPs. n = 6,159 (LEP) and 14,428 (MEP) cells.

(C) MEPs co-cultured with LEPs are not induced to proliferate relative to MEPs grown in monoculture. Clusters

corresponding to resting (black) and proliferative (blue) LEPs are identifiable in gene expression space (Fig. S5B). Projecting sample classification densities onto gene expression space for co-cultured MEPs (red, left) and MEPs cultured alone (green, middle) illustrates that both culture compositions are equally proliferative (table, right). n = 14,428 cells.

(D) MEPs co-cultured with LEPs exhibit enriched TGF-β signaling (as measured by TGFBI expression) relative to MEPs grown in monoculture. Each point represents an average of MEPs grouped according signaling molecule treatment. *** = Wilcoxon rank sum test (two-sided), $p = 1.5 \times 10^{-6}$. n = 32 signaling molecule condition groups. Data are represented as mean ± SEM.

(E) Hierarchical clustering and heat map analysis of MEPs grouped by signaling molecule treatment highlights an EGFR signaling transcriptional response specific to EGF and AREG treatment. Dendrogram labels: E = EGF, W = WNT4, A = AREG, I = IGF-1, R = RANKL, C = Control.




**Supplementary Figure 6**

MULTI-seq application to primary PDX samples, related to Figure 3.

(A) Representative histology of lung tissue illustrates metastatic progression in early, mid, and late-stage PDX mice. Individual metastases denoted with black arrows. H&E staining was performed 3 times (early), 4 times (mid), and 10 times (late), yielding the same result.

(B) Negative cell reclassification improves sample classification results. Singlets (black) localize into clusters in barcode space whereas doublets (red) localize between singlet clusters. Negative cells either co-localize with singlet or doublet clusters (blue outline, bottom) or cluster separately (red outline, bottom). Negative cell reclassification is insensitive to the true-negatives that cluster separately, while rescuing a subset of false-negatives that cluster amongst singlets and doublets. n = 12,086 cells.

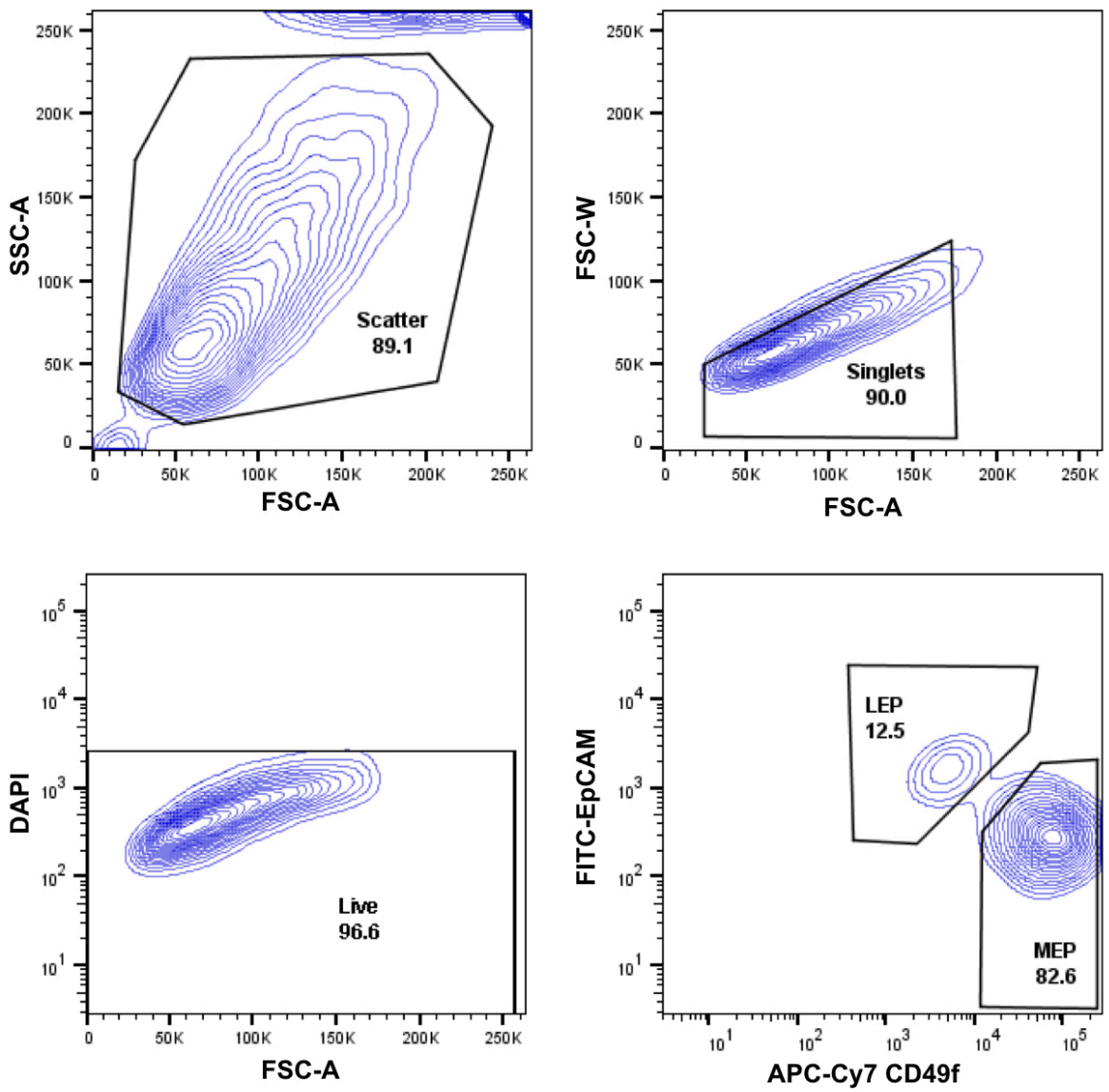
(C) Barcode SNR comparisons between samples ordered according to the viability (top) or total cell number (nCell, bottom) of the MULTI-seq barcoding conditions. See table S3 for details. Data are represented as mean ± SEM. n = 10,427 cells.

(D) Mouse immune cells in gene expression space colored according to tissue of origin. Lung immune cells (brown) cluster separately from primary tumor immune cells (teal). n = 8,420 cells.

(E) Bar plots describing the proportion of mouse (pink) and human (blue) cells detected during FACS enrichment and detected in the final 10X dataset. Classification of human and mouse cells from the L-A lung demonstrates sample classification accuracy and species-independence.

(F) Marker gene heat map describing markers utilized for defining cell type annotations. RNA UMI abundances are scaled from 0-1 for each gene. Values correspond to the average expression within each annotation group. Displayed genes represent the top 3 most statistically-significant genes for each cell type (Likelihood-ratio test for single cell gene expression[57] with Bonferroni multiple comparisons adjustment).

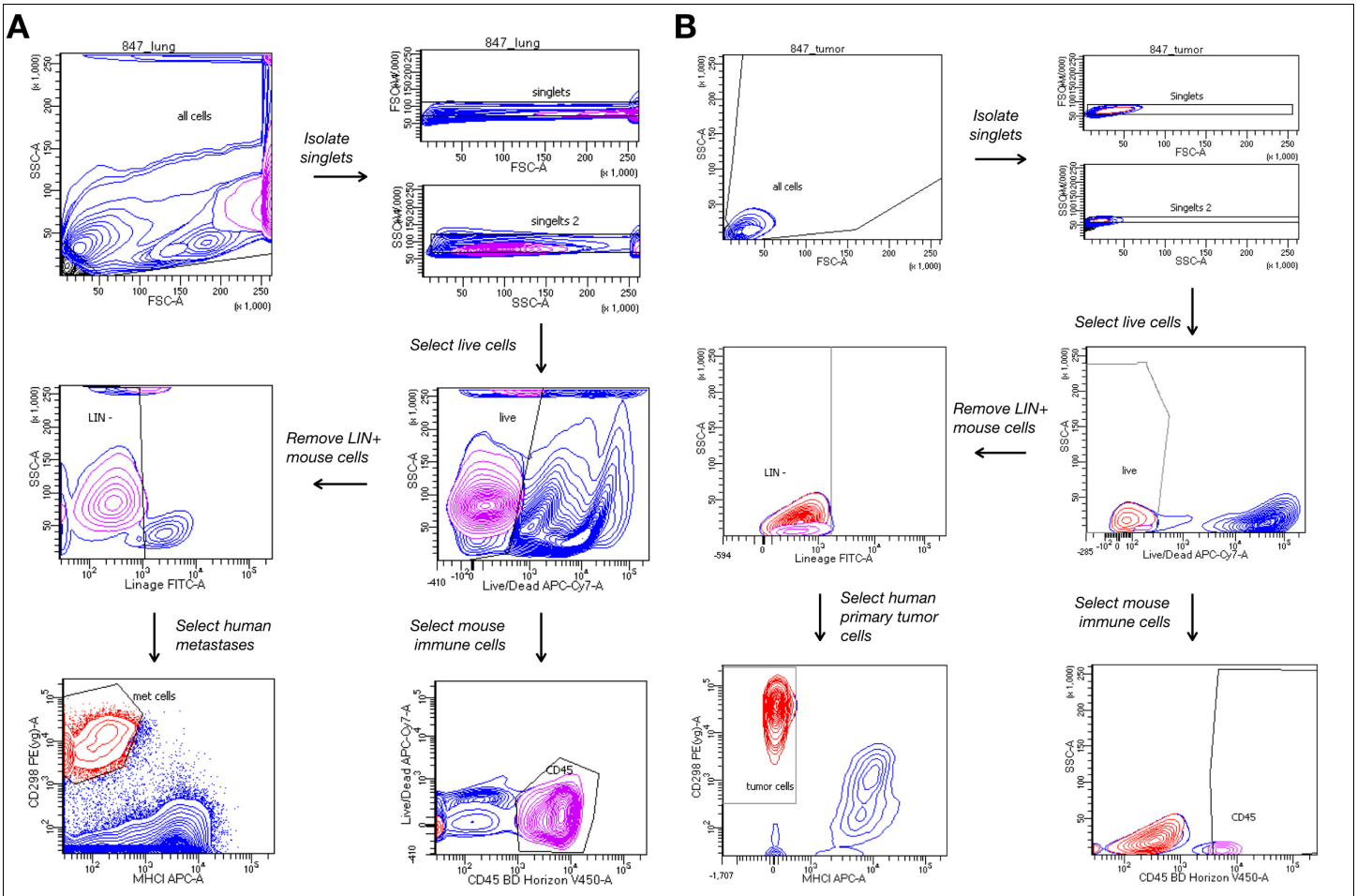
(G) Unsupervised clustering of classical monocytes recapitulates intercellular heterogeneity due to metastatic progression (displayed in Fig. 3D). See Table S6 for differential gene expression analysis results. n = 2,496 cells.




**Supplementary Figure 7**

FACS purification of LEP and MEP cells from bulk HMECs, related to Experimental Methods

Bulk HMECs were labeled with FITC anti-EpCAM and APC-Cy7 anti-CD49f to identify and isolate LEPs and MEPs. LEPs are identified as EpCAM high and CD49f low, while MEPs are CD49f high and EpCAM low. Gating strategy causes minor cell type impurities in final sorted population.

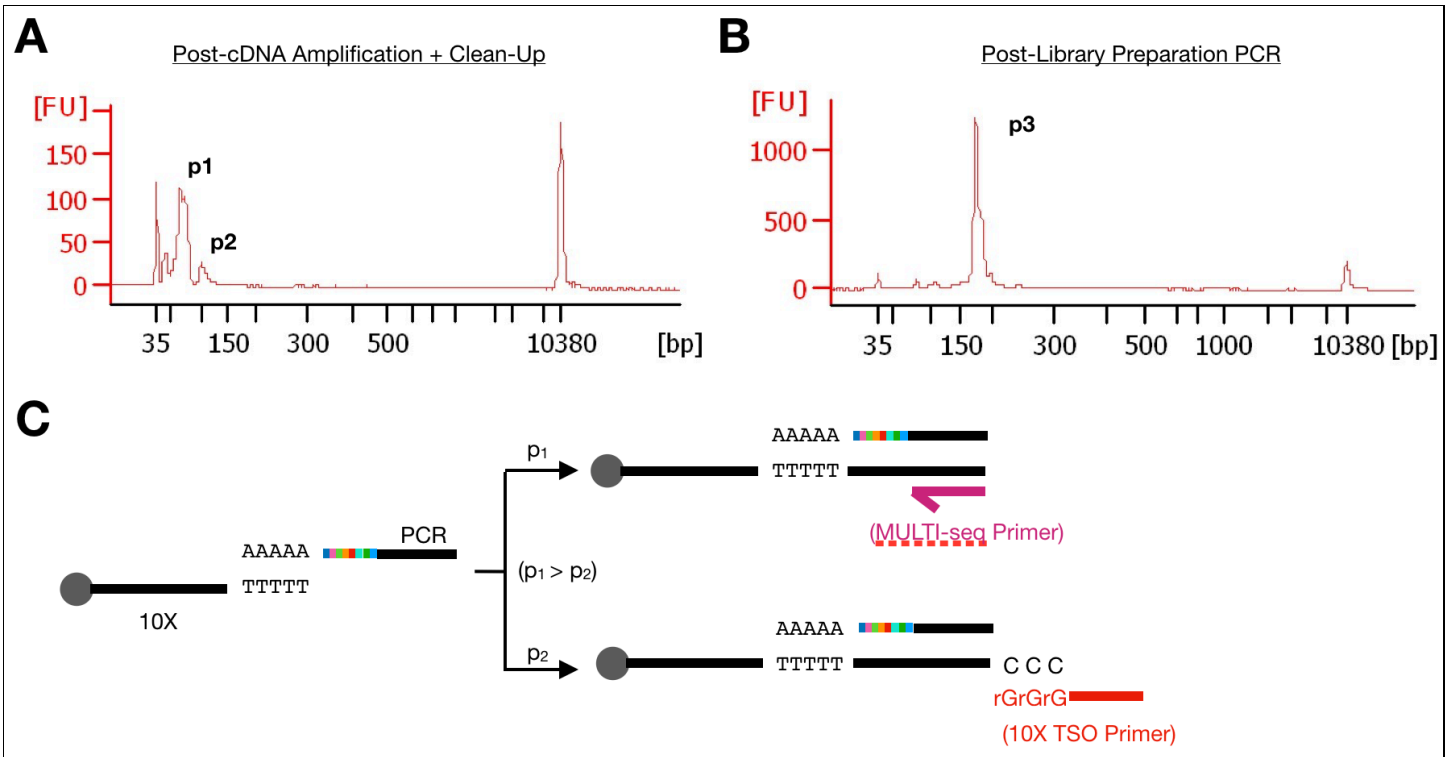



**Supplementary Figure 8**

FACS gating strategy for PDX lung and primary tumor samples, related to Experimental Methods

(A) Dissociated human metastases and mouse immune cells were separated from dissociate PDX mouse lungs using hCD298 and mCD45 following gating for live singlets. Mouse 847 (Sample L-A) is presented here as a representative example.

(B) Dissociated human primary tumor cells and mouse tumor-associated immune cells were separated using hCD298 and mCD45 following gating for live, singlets. Sample A is presented here as a representative example for all other primary tumor samples.




**Supplementary Figure 9**

Bioanalyzer traces of representative MULTI-seq barcode library, related to Experimental Methods

(A) Bioanalyzer traces following cDNA amplification and MULTI-seq barcode enrichment using 3.2X SPRI with 1.8X 100% isopropanol exhibits two distinct peaks. Bioanalyzer traces are representative of all datasets presented in this study (n = 4). The first peak (p1) is an average of 65-70bp in length and likely corresponds to barcodes amplified via the MULTI-seq additive primer. The second peak (p2) is an average of 100bp in length and likely corresponds to barcodes that successfully underwent MMLV-RTase template switching and were subsequently amplified by the standard 10X Genomics Single Cell V2 primer.

(B) Bioanalyzer analysis following library preparation PCR exhibits one distinct peak (p3) with an average length of 173bp, matching expectations. Bioanalyzer traces are representative of all datasets presented in this study (n = 4).

(C) Schematic illustrating the two species of reverse-transcribed MULTI-seq barcodes with and without template switching. Processive reverse-transcription without template switching (p1) is more likely than reverse-transcription with template switching (p2), resulting in relative enrichment of the 65-70bp product following cDNA amplification.

**Table S1:** List of genes with >1.5-fold expression difference between LMO/CMO-labeled and unlabeled HEKs, related to Fig. 1

| Comparison | Markers | FC | Comparison | Markers | FC | Comparison | Markers | FC |
|---|---|---|---|---|---|---|---|---|
| LMO vs UN | MIF | 2.8 | LMO vs CMO | SNORA76 | 1.7 | CMO vs UN | MIF | 1.9 |
| | KRTCAP2 | 1.5 | | NMT1 | 1.5 | | S100A2 | 1.6 |
| | TOMM5 | 1.5 | | | | | MT2A | 1.5 |
| | | | | | | | AP2B1 | -1.5 |
| | | | | | | | TOP2A | -1.5 |
| | | | | | | | MALAT1 | -1.5 |
| | | | | | | | SNORA76 | -1.7 |
| | | | | | | | NMT1 | -1.8 |

LMO = LMO-labeled HEKs, CMO = CMO-labeled HEKs,
UN = Unlabeled HEKs, FC = Fold-change

**Table S2:** MULTI-seq barcode sequencing statistics

| Dataset | % Aligned | % Duplicate UMI | UMIs/cell | SNR | NGS Format |
|---|---|---|---|---|---|
| POC-LMO | 93.8 | 6.25 | 77 | 69 | HiSeq 4000 (33.3%) |
| POC-CMO | 94.4 | 5.74 | 55 | 41 | HiSeq 4000 (33.3%) |
| HMEC-orig | 97.1 | 12.62 | 5095 | 199 | HiSeq 4000 (100%) |
| HMEC-rep | 97.3 | 13.67 | 65 | 34.7 | HiSeq 4000 (50%) |
| PDX | 99.0 | 17.36 | 4453 | 49.6 | NovaSeq (2.5%) |
| Nuclei (LMO) | 98.8 | 28.67 | 731 | 2.3 | NovaSeq (1.25%) |
| Nuclei (CMO) | 98.6 | 28.18 | 8415 | 91.5 | NovaSeq (1.25%) |

% Aligned = Alignment rate, i.e., the proportion of reads aligning to any reference barcode.
SNR = Signal-to-noise ratio, i.e., the primary classification barcode divided by the next most abundant barcode.
All sequencing statistics were computed on the subset of reads associated with classified singlets.
POC = Proof-of-concept. PDX = Patient-Derived Xenograft, HMEC = Human mammary epithelial cell.

**Table S3:** PDX metadata, related to Fig. 3

| Sample | MouseID | #Cells | Viability | pHu-FACS | pHu-10X | pMo-FACS | pMo-10X |
|---|---|---|---|---|---|---|---|
| Early-1-PT | 1 | 280K | 0.647 | 0 | 0 | 1 | 1 |
| Mid-1-PT | 2 | 580K | 0.0153 | 0 | 0 | 1 | 1 |
| Late-PT | 3 | 665K | 0.0845 | 0 | 0 | 1 | 1 |
| Early-1-Lung | 1 | 700K | 0.33 | 0 | 0.01 | 1 | 0.99 |
| Early-2-Lung | 4 | 340K | 0.38 | 0 | 0 | 1 | 1 |
| Mid-1-Lung | 2 | 370K | 0.50 | 0.02 | 0.02 | 0.98 | 0.98 |
| Mid-2-Lung | 5 | 550K | 0.309 | 0.01 | 0.01 | 0.99 | 0.99 |
| Late-Lung | 3 | 550K | 0.314 | 0.22 | 0.15 | 0.78 | 0.85 |
| WT-Lung | 6 | 460K | 0.392 | 0 | 0 | 1 | 1 |

Number of cells and viability refer to the LMO labeling reaction.
pHu-FACS, pMo-FACS = Proportion of human and mouse cells sorted using FACS.
pHu-10X, pMo-10X = Proportion of human and mouse cells present in the final dataset.

**Table S4:** Top 5 marker genes for each low-RNA cluster within classified and unclassified datasets, related to Fig. 3

**Classified**

| Cluster | Markers | FC |
|---|---|---|
| Monocyte | Pou2f2 | 4.2 |
| | Rpsa | 2.6 |
| | Csf1r | 2.6 |
| | Ahnak | 2.7 |
| | Cx3cr1 | 4.0 |
| Interstitial Mac. | C1qa | 11.0 |
| | C1qb | 9.0 |
| | Apoe | 6.1 |
| | C1qc | 7.6 |
| | Ccl8 | 20.4 |
| Alveolar Mac. | Ear2 | 16.0 |
| | Ear10 | 17.7 |
| | Ctsd | 6.4 |
| | Chil3 | 7.5 |
| | Lpl | 7.7 |
| Neutrophil | S100a8 | 61.1 |
| | S100a9 | 57.6 |
| | Il1b | 9.3 |
| | Msrb1 | 5.8 |
| | Srgn | 5.0 |
| Endothelial | Ramp2 | 58.2 |
| | Ly6c1 | 60.2 |
| | Cldn5 | 38.1 |
| | Ctla2a | 34.4 |
| | Aqp1 | 25.3 |

**Unclassified**

| Cluster | Markers | FC |
|---|---|---|
| Broken-1 | Xist | 4.3 |
| | Lgals1 | 1.6 |
| | Rpl25 | 1.6 |
| | Rps20 | 1.8 |
| | Grn | 1.9 |
| Broken-2 | mt-Nd4 | 2.3 |
| | mt-Nd1 | 2.4 |
| | Rpl14 | 1.7 |
| | mt-Co2 | 2.4 |
| | mt-Co3 | 2.5 |
| Neutrophil | S100a9 | 47.4 |
| | S100a8 | 48.6 |
| | Retnlg | 48.3 |
| | Pglyrp1 | 13.1 |
| | S100a11 | 9.2 |

**Table S5:** List of genes with >1.5-fold expression difference between classical monocytes at distinct stages of metastatic progression, related to Fig. 3.

| Cluster | Markers | FC |
|---|---|---|
| WT | Ear2 | 2.3 |
| | Rsrp1 | 2.1 |
| | Plaur | 1.7 |
| | Rgcc | 2.1 |
| | Klf4 | 1.7 |
| | Jund | 1.6 |
| | Wsb1 | 1.6 |
| | Pdlim1 | 1.6 |
| | Tagln2 | 1.6 |
| | Pglyrp1 | 2.0 |
| | Fn1 | 1.5 |
| | Ezr | 1.5 |
| | Tsc22d3 | 1.5 |
| | Cks2 | 1.8 |
| | Hspa1a | 1.6 |
| Early/Mid | Isg15 | 1.8 |
| | Thbs1 | 1.7 |
| Mid-2 | Tppp3 | 2.4 |
| | Adgre5 | 1.7 |
| | Tagln2 | 1.7 |
| | Crip1 | 1.6 |
| | Metrnl | 1.8 |
| | Emp3 | 1.6 |
| | Cd74 | 2.1 |
| | Cd300a | 1.6 |
| | Btg2 | 1.5 |
| | Pou2f2 | 1.7 |
| | Sgk1 | 1.6 |
| | Il1b | 1.8 |
| | Gngt2 | 1.5 |
| | Hist1h1c | 1.6 |

| Cluster | Markers | FC |
|---|---|---|
| Late-1 | Fos | 2.6 |
| | Dusp1 | 2.3 |
| | Jun | 2.6 |
| | Atf3 | 2.2 |
| | Ier3 | 2.0 |
| | Ccl3 | 2.6 |
| | Tsc22d3 | 1.7 |
| | S100a8 | 3.7 |
| | Ccl2 | 2.0 |
| | Saa3 | 2.5 |
| | Fosb | 1.7 |
| | Socs3 | 1.8 |
| | Wfdc21 | 2.4 |
| | Klf6 | 1.5 |
| | S100a9 | 2.2 |
| | Egr1 | 1.5 |
| | Hspa1a | 1.8 |
| | Lcn2 | 1.6 |
| | Lrg1 | 1.6 |
| | Ccl4 | 1.9 |
| Late-2 | Rgcc | 2.8 |
| | Rsrp1 | 2.8 |
| | Nfkbia | 2.4 |
| | Mmp19 | 2.4 |
| | Cxcr4 | 2.3 |
| | Arg2 | 2.3 |
| | Lmna | 2.2 |
| | Saa3 | 2.0 |

**Table S6:** List of genes with >1.5-fold expression difference between late-stage classical monocytes, related to Fig. 3.

| CM Subtype | Markers | FC |
|---|---|---|
| Thsb1+ Cd14+ Late CMs | Cd14 | 2.4 |
| | Cepbp | 1.5 |
| | Lmna | 2.2 |
| | Btg1 | 1.5 |
| | Plaur | 1.6 |
| | Rgcc | 1.9 |
| | Nfkbia | 1.6 |
| | Tnpo3 | 1.6 |
| | Tlr2 | 1.6 |
| | Tubb6 | 1.6 |
| | C5ar1 | 1.6 |
| | Mmp19 | 1.6 |
| | Cxcr4 | 1.7 |
| | Tppp3 | 1.7 |
| | Ninj1 | 1.5 |
| | Skil | 1.5 |
| | Thbs1 | 1.6 |
| | Dmkn | 1.6 |

| CM Subtype | Markers | FC |
|---|---|---|
| Thsb1- Cd14- Late CMs | Fos | 2.8 |
| | Dusp1 | 2.4 |
| | Jun | 2.5 |
| | Atf3 | 2.1 |
| | Ccl3 | 2.7 |
| | Tsc22d3 | 1.7 |
| | Rgs2 | 1.7 |
| | S100a8 | 3.2 |
| | S100a9 | 2.6 |
| | Txnip | 1.6 |
| | Egr1 | 1.6 |
| | Ccl4 | 2.2 |
| | Ccl2 | 1.5 |
| | Wfdc21 | 1.5 |

**MULTI-seq: Universal sample multiplexing for single-cell RNA sequencing using lipid-tagged indices.**

## SUPPLEMENTAL NOTES

Aspects of the expression library analysis workflow that were shared between all datasets were outlined in the 'Expression library analysis' section of the Computational Methods. However, unique analytical techniques were applied to each presented dataset. In-depth summaries of these analyses are discussed below.

### Proof-of-Concept scRNA-seq and snRNA-seq

Testing the effects of MULTI-seq barcoding on scRNA-seq and snRNA-seq data: MULTI-seq could negatively influence scRNA-seq and snRNA-seq data in two main ways: by (1) competing with endogenous mRNAs for capture bead hybridization regions, or (2) inducing a transcriptional response to LMO or CMO labeling. To test these possibilities, we first parsed our proof-of-concept scRNA-seq and snRNA-seq datasets to include only HEK293 (HEK) cells and MEF nuclei, respectively. Focusing on individual cell types ensures that any observed performance differences are primarily due to technical and not biological reasons.

All HEK cells and MEF nuclei subsets were indistinguishable with regards to the total number of detected RNA UMIs and genes (Fig. S2C, Fig. S2H, left). Moreover, barcode and RNA UMIs were not negatively correlated (Fig. S2C, Fig. S2H, right). These observations suggest that MULTI-seq barcodes do not detrimentally compete with endogenous transcripts during mRNA capture. Additionally, LMO-, CMO-, and unlabeled HEK cells and MEF nuclei exhibited similar proportions of reads aligning to mitochondrial genes (Fig. S2C, Fig. S2H, left); therefore, LMO and CMO labeling are unlikely to induce an apoptotic cellular response. To explore whether MULTI-seq labeling perturbs endogenous gene expression in other ways, we compared the proportion of each cell/nuclei's 100 nearest neighbors in principal component (PC) space that were derived from LMO-, CMO-, or unlabeled subsets. Neighborhoods were defined by computing the Euclidean distance matrix for statistically-significant PCs with the 'rdist' R function.

For HEK cells, neighborhood analysis revealed that CMO-labeled cells preferentially co-localized in gene expression space, while LMO-labeled and unlabeled neighborhoods were nearly indistinguishable (Fig. S2D). We then performed differential gene expression analysis between HEKs from each sample group, which demonstrated that 3 and 8 genes were 1.5-fold enriched in LMO- or CMO-labeled HEKs relative to unlabeled controls, respectively (Table S1). Intriguingly, even after only < 1 hour on ice, CMO-labeled HEKs exhibited differential expression of AP2B1, which has established roles in cholesterol and sphingolipid transport. When considered along with flow cytometry analyses demonstrating that CMOs exhibit reduced live-cell membrane residency compared to LMOs (Fig. S1B), these results collectively illustrate that LMOs are the preferred reagent for scRNA-seq sample multiplexing.

In contrast to HEK cells, MEF nuclei from each labeling condition had uniform neighborhood proportions (Fig. S2I). Additionally, we did not detect any genes that were differentially expressed > 0.7-fold between LMO-, CMO-, and unlabeled nuclei. These results

demonstrate that the transcriptional response to CMO labeling observed in HEK cells was absent in nuclei. Moreover, we observed a ~10-fold increase in barcode nUMIs for CMO-labeled MEF nuclei relative to LMO-labeled nuclei (Fig. S2H, right). This observation was in-line with our previous flow cytometry titration experiments (Fig. S1C). We believe that this difference in sample barcode capture efficiency was due to the presence of BSA in nuclei resuspension buffer, which is necessary to prevent aggregation nuclei purification. BSA has a lipid-binding pocket which likely sequesters LMOs, leading to reduced sample barcode association with the nuclear membrane. When considered along with the commercial-availability of CMOs, these results collectively illustrate that CMOs are the preferred reagent for snRNA-seq sample multiplexing.

Sample classification accuracy: To analyze the accuracy of MULTI-seq classifications during snRNA-seq, we compared MULTI-seq sample classifications to cell type annotations determined by (1) mm10 pre-mRNA reference transcriptome alignment for MEFs, (2) CD3D expression for Jurkats, and (3) Xist expression for HEKs, which were isolated from a female donor. Inter-species doublets were defined as cells with > 256 RNA UMIs from both the human and mouse pre-mRNA reference transcriptomes. Sample classification accuracy was then approximated using the proportion of matching MULTI-seq classifications and cell type annotations (Fig. 1F). The mismatch rate for all three cell types was ~0.5% while 85% of known mouse-human doublets were identified.

Jurkat T-cell activation time-course: Gene expression centroids for each Jurkat time-point were computed as the mean t-SNE embedding coordinates amongst cells classified into each time-point group.

**Semi-Supervised Negative Cell Reclassification**

In its current form, MULTI-seq barcoding is an imperfect process that produces a small fraction of cells that cannot be classified into sample groups. These negative cells are of two varieties: True and false negatives. True negatives manifest in barcode space as high-density regions lacking enrichment for any particular barcode (e.g., central HMEC region in Fig. S4; top-right PDX region in Fig. S6B). True negatives result from cells with poor barcode labeling. In contrast, false negatives result from algorithmic misclassification.

False negatives manifest in barcode space as negatives intermixed amongst high-density regions exhibiting enrichment for a single barcode. Since a single inter-maxima quantile threshold is applied to all barcodes during sample classification, we believe false negatives arise because this thresholding strategy may be sub-optimal for a subset of barcode distributions. Thus, although false negatives have poor *absolute* signal in comparison to high-confidence singlets, we reasoned that false negatives could be 'rescued' by computing the *relative* strength of each barcode signal on a cell-by-cell basis.

To distinguish which negative cells are the best candidates for reclassification before reclassifying negatives into their appropriate barcode groups, we used the following strategy:

1. Repeat the original sample classification workflow, recording the total number of thresholds that each negative cell surpasses at each quantile.

2. Compute each cell's classification stability (CS) – defined as the number of quantiles across which a cell surpasses a single threshold.
3. Subset equal numbers of 'ground-truth' cells from the original classification results.
4. Perform semi-supervised k-means clustering on merged data including 'ground-truth' and negative cells. Clustering is semi-supervised because one member of each 'ground-truth' sample group is used to initialize cluster centers.
5. Compute the rate at which 'ground-truth' and negative cell classifications match the k-means results.
6. Iteratively repeat steps 4 and 5 using a different 'ground-truth' cell to initialize cluster centers during each iteration. Repeat until all 'ground-truth' cells have been used.
7. Compare k-means matching rates between 'ground-truth' and negative cells binned according to CS values. Negative cells with CS values resulting in matching rates that approximate 'ground-truth' matching rates are reclassified.

Negative cell reclassification rescues 10%-20% of negative cells across the different datasets presented in this study. While not insignificant, we believe that further optimization will improve performance. For instance, in our PDX data, two distinct clusters of negative cells remain following reclassification (Fig. S6B) – a true-negative population low for all barcodes (red outline, bottom right) and a putative false-negative population (blue outline, center right). It is unclear if these cells are doublet or singlets. However, we interpret this observation as evidence that further optimization will improve performance.

**96-plex HMEC**

Exploring transcriptional responses to cell type composition: Besides transcriptional variability between MEPs and LEPs, the cell type composition of HMEC cultures represented the most pronounced source of variability in our 96-plex HMEC experiment (Fig. 2B). To explore the transcriptional responses to cell type composition, we began by pre-processing data subsets containing only MEPs or LEPs. Separating cell types revealed distinct resting and proliferative MEP and LEP subsets discernible by enriched MKI67 expression (Fig. S5B). To assess whether co-culture influenced proliferation, we specified subsets of cells where each cell type composition (e.g., mono- or co-cultures) were equally abundant. We then determined whether mono- and co-cultured cells were evenly represented in the resting and proliferative states. Down-sampling in this fashion controls for differences in the total numbers of cells from each group. This analysis revealed that LEPs were specifically induced to proliferate in the presence of MEPs (Fig. 2D), whereas the same effect was not observed in MEPs cultured in the presence of LEPs (Fig. S5C).

Next, we pre-processed data subsets containing only resting MEPs or LEPs. Notably, we did not proceed with all resting cells but, rather, down-sampled every culture composition (e.g., mono- or co-cultures) to have equal numbers of cells from each signaling molecule perturbation. We then computed the average TGFBI expression amongst MEPs and LEPs grouped by signaling molecule exposure and observed that co-cultured LEPs and MEPs were associated with elevated TGFBI expression independent of perturbation (Fig. 2D, Fig. S5D). In our proof-of-concept scRNA-seq experiment, we observed that TGFBI expression is increased specifically in HMECs responding to TGF-β (Fig. 1E). Thus, these results suggest that co-culturing induces paracrine-mediated TGF-β signaling in both LEPs and MEPs.

Exploring transcriptional responses to signaling molecule perturbation: Using data subsets containing equal numbers of resting MEPs or LEPs from each culture composition and signaling molecule condition (described above), we next sought to characterize transcriptional responses to signaling molecules. To this end, we grouped cells according to signaling molecule exposure and performed hierarchical clustering on the average gene expression profile for each group using the 'BuildClusterTree' function in 'Seurat'. Hierarchical clustering revealed two distinct clades corresponding to cells stimulated with (1) the EGFR ligands AREG and EGF and (2) RANKL, IGF1, or WNT4. Notably, AREG/EGF stimulation dominated the effect of RANKL/IGF1/WNT4, as cells grown in media supplemented with both AREG/EGF and RANKL/IGF1/WNT4 remained members the EGFR ligand clade.

Differential gene expression analysis between these two clades revealed that AREG/EGF stimulated cells expressed elevated levels of a number of EGFR signaling target genes (Fig. 2E, S5E), as expected. Differentially-expressed genes amongst RANKL/IGF1/WNT4 stimulated cells could not be as readily connected to their corresponding signaling pathways. This observation suggests that the rich media used to culture HMECs buffered the cells against RANKL/IGF1/WNT4 induction. This notion is further supported by the fact that cells stimulated with EGFR ligands – which were purposefully depleted from the M87A media used in this experiment – represented the most pronounced transcriptional signature amongst signaling molecule conditions.

Amongst cells induced with EGFR ligands, hierarchical clustering also revealed sub-clades corresponding to cells exposed to AREG or EGF. However, differential gene expression analyses between these groups using a fold-change threshold of 1.25 were largely unsuccessful (data not shown). Higher numbers of cells per sample or more sophisticated analytical methods may be necessary to better understand transcriptional responses to distinct EGFR ligands.

Doublet analysis, comparison to computational doublet prediction methods: Recently described computational double detection methods like DoubletFinder (McGinnis et al., 2018) require parameter selection prior to being applied to scRNA-seq data. To fit DoubletFinder parameters to our 96-plex HMEC scRNA-seq data, we began by performing a parameter sweep using the 'paramSweep' function in the 'DoubletFinder' R package. Ideal parameters were then defined using the 'summarizeSweep' function, which uses receiver operating curve analysis to compute the predictive capacity of each parameter set relative to ground-truth doublet labels. We used MULTI-seq doublet classifications as ground-truth in this application.

With ideal parameters defined (e.g., pN = 0.25, pK = 0.03), we then thresholded DoubletFinder results by adjusting the total number of MULTI-seq-defined doublets to account for homotypic doublet formation. Homotypic doublets are doublets that are formed from transcriptionally-similar cells and are known to be undetectable using computational doublet detection methods that rely solely on gene expression features (McGinnis et al., 2018, Wolock et al., 2018). To account for homotypic doublets, we multiplied the total doublet number (3413) by the sum of squared cell type frequencies (0.51), resulting in 1738 total doublet predictions.

DoubletFinder and MULTI-seq doublet classifications largely co-localize in gene expression space (Fig. 2C), exhibiting enrichment amongst cells expressing high levels of both LEP and MEP markers (Fig. S5A). Beyond cells expressing both KRT19 and KRT14, doublet-

enriched regions with either LEP- or MEP-like expression patterns are also detected by MULTI-seq (Fig. 2B, arrow). These doublets are likely to represent doublets formed from LEP and MEP sub-states (e.g., EGFR-induced, proliferative, etc.), that would be overlooked when classifying doublets using marker gene analysis, alone.

Although DoubletFinder and MULTI-seq doublet classifications are generally in agreement, there are two noteworthy and expected discrepancies. First, MULTI-seq identifies many doublets that DoubletFinder classifies as singlets that are evenly interspersed amongst LEP and MEP clusters in gene expression space. These doublets likely correspond to homotypic doublets formed from transcriptionally-similar MEPs or LEPs. Because DoubletFinder predicts doublets as cells that cluster separately from real singlets in gene expression space, we expected DoubletFinder to be insensitive to homotypic doublets.

Second, DoubletFinder identifies many doublets that were classified as singlets during MULTI-seq sample classification. We expect MULTI-seq to be insensitive to the subset of doublets derived from cells barcoded with the same MULTI-seq sample index (e.g., 1/76 = ~1% false-negative rate). Sample multiplexing methods are generally insensitive to such doublets, which emphasizes how sample multiplexing and computational doublet detection methods can serve complementary roles in this context. However, we would expect these doublets to be randomly distributed amongst existing doublet clusters in gene expression space. Instead, DoubletFinder identifies many cells as doublets that are MULTI-seq-defined singlets *and* form distinct clusters in gene expression space (Fig. 2C, arrow). These discordant cells are proliferative LEPs, and thus represent DoubletFinder false-positives. The performance of computational detection methods suffer when applied to scRNA-seq data with low numbers of cell states (McGinnis et al., 2018, Wolock et al., 2018) and, hence, further emphasizes the utility of sample multiplexing for doublet detection.

Signal-to-noise ratio (SNR) computation: SNR for singlets, doublets, and negative cells was calculated as the quotient of the two most abundant raw barcode UMI abundances for each cell (Fig. S3D). Since cells are discarded as doublets when surpassing two or more barcode-specific thresholds during our sample classification workflow, we reasoned that the relative abundances of the top two barcodes was a sufficient SNR definition. In singlets, on-target barcodes are an average of 199-fold higher than the most abundant off-target barcode. Doublets have much lower SNR but higher total barcode nUMIs. This observation matches expectations, as doublet formation results in the pooling of MULTI-seq barcodes from two cells. Negative cells exhibit very low SNR and total nUMIs, indicating that negative cells were not sufficiently labeled with LMOs to enable sample classification. The same SNR trends were also observed in our 96-plex HMEC technical replicate data (Fig. S3G).

Sample drop-out and 96-plex technical replicate scRNA-seq experiment: During our initial 96-plex HMEC experiment, 20 samples were unaccounted for in the final dataset. Missing barcodes were defined as those lacking any region of significant enrichment in barcode space. The presence of missing sample barcodes is problematic during the MULTI-seq sample classification workflow, as extremely low barcode thresholds result in unrealistically-high numbers of doublets. Thus, we suggest visual inspection of barcode abundances in barcode space (e.g., as generated using t-SNE) prior to sample classification for all MULTI-seq datasets. Barcodes that are not enriched in distinct domains of barcode space should be discarded.

To determine why barcodes were missing from our 96-plex HMEC data, we checked the position of every missing barcode on the 96-well plate utilized during sample barcoding, washing, and pooling. Interestingly, 7/20 missing barcodes were positioned in a single column on the 96-well plate (Fig. S3B), suggesting that manual pipetting error (e.g., poor resuspension during pooling) caused at least a portion of the missing barcodes. To verify that MULTI-seq can indeed be scaled to large sample numbers, we performed a 96-plex technical replicate experiment. In this experiment, we pooled 96 HMEC cultures and sequenced the resulting expression library at very shallow depth. Shallow sequencing results are sufficient to identify cell-associated droplets, which we then used to perform sample classification on barcode data sequenced at standard depth (Table S2). This workflow resulted in the accurate classification of cells into all 96 sample groups (Fig. S3E-G), illustrating that MULTI-seq can be scaled to high sample numbers.

**PDX**

MULTI-seq sample classifications distinguish low-RNA from low-quality cells: Following expression library pre-processing (e.g., using CellRanger), raw RNA UMI count matrices must be parsed to define cell barcodes associated with intact cells versus ambient mRNA and cell debris. This challenge is commonly addressed by identifying the inflection point of log-log RNA UMI by RNA UMI rank distributions, which follows the assumption that droplets containing intact cells should feature elevated nUMIs. This strategy is inherently biased against cells with intrinsically low RNA content, and may be confounded by distributions with multiple inflection points (e.g., datasets with many cell types, Lun et al., 2018).

The RNA UMI distribution for mouse immune cells sequenced during our PDX experiment exemplifies this issue. Specifically, we observed a mode corresponding to cell barcodes with ~500 total RNA UMIs that was discarded by the standard CellRanger UMI threshold (Fig. 3C, top left). To assess whether this region represented intact low-RNA cells, we performed the MULTI-seq sample classification workflow on all cell barcodes with at least 100 RNA UMIs. We selected this threshold because droplets with < 100 RNA UMIs can be confidently assumed to be empty (Lun et al., 2018). Intriguingly, sample classifications produced 2,580 singlets and 583 negatives amongst cells with RNA UMIs between 100 and the CellRanger threshold (1350 RNA UMIs).

To test whether sample classification results could be used to distinguish low-RNA cells from ambient mRNA and cellular debris, we first pre-processed putative low-RNA singlets using 'Seurat' and used unsupervised clustering and differential gene expression analyses to reveal discrete clusters in gene expression space characterized by established marker genes for neutrophils, monocytes, alveolar macrophages and endothelial cells (Fig. 3C, top right, Table S4). In contrast, equivalent analyses of unclassified cell barcodes with 100-1350 RNA UMIs revealed clusters corresponding to broken cells and a small number of neutrophils. We annotated broken cells into two subsets – one with enriched mitochondrial gene expression and another with elevated levels of lncRNAs (e.g., Xist) and ribosomal RNAs (Table S4). We speculate that the latter represents nuclei released from cells due to shear stress.

Collectively, these results suggest that MULTI-seq – and sample multiplexing methods, writ-large (see Stoeckius et al., 2018 for analogous analyses) – improves scRNA-seq quality control workflows by distinguishing low-RNA cells from ambient mRNA and debris. We anticipate that this feature will further increase scRNA-seq cell throughput while diminishing the effects of systematic bias against cells with intrinsically-low RNA content.

Cell state annotation definition: Mouse lung immune cell state annotations (Fig. 3D) were defined by performing unsupervised clustering and assessing the accuracy of clustering results using marker genes identified previously in the same tissue (e.g., Reyfman et al., 2018; Tabula Muris Consortium, 2018). We then performed marker gene analysis (fold-change threshold = 2) on these cell groups and selected the top three most statistically-significant genes. Prior to heatmap visualization, we computed the average expression of each marker gene within each cell group and scaled these averages from 0 to 1 (Fig. S6F).

Immune cell proportion shifts: To assess whether lung immune cell type proportions shifted during metastatic progression in our PDX mice, we first defined a subset of cells where each tumor stage (e.g., WT, early, mid, and late) was equally represented. Down-sampling in this fashion controls for technical differences in the number of sequenced cells. We then computed the proportion of each cell type present in lung immune cells from each tumor stage (Fig 3E). Statistically-significant proportional shifts relative to WT proportions were then defined using two-proportion z-tests (as implemented in the 'prop.test' R function) with Bonferroni multiple comparison correction (as implemented in the 'p.adjust' R function).

Classical monocyte heterogeneity: Tumor stage-specific heterogeneity amongst classical monocytes (CMs) was visually discernible in mouse immune cell gene expression space (Fig. 3D, bottom, Fig. 3F). CMs are also known to be recruited to the metastatic lung in PDX breast cancer mouse models, where they exhibit metastasis-associated phenotypic heterogeneity (Catena et al., 2013, Condamine et al., 2015, Kitamura et al., 2018, Ouzounova et al., 2017). For these reasons, we sought to characterize CM heterogeneity during metastatic progression in our PDX data.

We began by pre-processing a dataset including only CMs using 'Seurat'. Unsupervised clustering of these data revealed sub-structure demarcating each tumor stage (Fig. S6G). Early-stage CMs were distinct from WT CMs despite the lack of detectable metastases, which suggests that this data could provide insight into CM transcriptional behavior during metastatic colonization (Table S5). Early-stage CMs were also transcriptionally-similar to a subset of mid-stage CMs. However, mid- and late-stage CMs manifested as two distinct sub-states featuring heterogeneous expression of many genes previously linked to metastatic/aggressive behavior (Table S5).

To explore CM heterogeneity in the presence of metastases, we further parsed CMs to include only those from late-stage PDX mice. Unsupervised clustering and differential gene expression analyses revealed that late-stage CM subsets could be distinguished according to Thbs1 and Cd14 expression (Fig. 3F, Table S6). Considering the established anti-metastatic role of Thbs1 in the lungs of metastatic PDX breast cancer models (Catena et al., 2013), we speculate that Thbs1+/Cd14+ CMs functionally inhibit metastatic progression. In contrast, Thbs1-/Cd14- CMs expressed elevated levels of known pro-metastatic genes (Ouzounova et

al., 2017), which suggests that these CMs contribute to metastatic progression. Further exploration of the functional consequences and mechanisms regulating CM heterogeneity will be the subject of a future publication.

Sample-to-sample SNR comparison: While preparing our PDX samples for scRNA-seq, we tracked the cell viability (by FACS) and number of cells (by manual counting with a hemocytometer) for each MULTI-seq labeling reaction in order to assess how these two features influence LMO labeling efficiency. We reasoned that poor viability and high cell numbers would decrease LMO labeling efficiency, resulting in lower SNR relative to cells labeled in high-viability, low-cell-number conditions.

We computed the SNR for each mouse singlet as the quotient of the top two abundant raw barcode UMIs (as described previously). Comparing SNR distributions between samples does not reveal any pronounced differences correlating with either viability or cell number (Fig. S2C). Since the viability was highly variable between all samples (Table S3), this suggests that viability does not negatively influence SNR. Moreover, SNR was not sensitive to the number of cells in the range tested in this experiment. However, we anticipate that larger inter-sample differences in cell numbers would indeed result in variable SNR.

Computing inter-sample variability using Earth Mover's Distance: Earth Mover's Distance (EMD) measures the distance in gene expression space that is required to map two distinct high-dimensional manifolds onto one another. To this end, EMD is an emerging tool to quantify differences amongst sets of cells in scRNA-seq data. We used EMD, as implemented in the 'calculate_emd' function from the 'EMDomics' R package (Nabavi et al., 2016), to quantify the variability between lung immune cells from biological replicate mice and mice from distinct tumor stages.

Specifically, we first down-sampled our existing data to include equal numbers of CMs from each tumor stage and mouse. Down-sampling in this fashion is necessary to control for differences in EMD results solely due to the total number of cells. We then extracted the PC space embeddings for this cell subset, and performed EMD on cells grouped by (1) tumor stage and (2) mouse ID. Notably, we only extracted embeddings for statistically-significant PCs (e.g., 10 for the CM-only dataset). We then scaled all of the EMD values from 0 to 1 and found the mean EMD between tumor stages and biological replicates (e.g., mice 1/4 and 2/5). CMs from biological replicates had a lower mean scaled EMD than CMs from each tumor stage (0.16 vs. 0.69), demonstrating that the observed CM heterogeneity between different tumor stages is not solely attributable to variability between individual mice.

## SUPPLEMENTAL METHODS

Analytical flow cytometry: The BD FACSCalibur instrument was used to perform analytical flow cytometry experiments measuring live-cell and nuclear membrane labeling efficiency (Fig. S1A,C), LMO and CMO membrane residency kinetics on ice (Fig. S1B,D) and at room temperature (Fig. S1E), and efficacy of BSA quenching (Fig. S1F). HEK293 cells and nuclei were utilized for all experiments. Samples were prepared using the same workflows as proof-of-concept scRNA-seq and snRNA-seq experiments (discussed above) with one exception. In place of barcode oligonucleotides, anchor LMOs or CMOs were pre-hybridized to equimolar

concentrations of FAM- or AF647-conjugated oligonucleotides matching the barcode oligonucleotide 5' PCR handle excluding the barcode and poly-A regions.

For titration experiments, $5x10^5$ cells or nuclei were suspended in 180 µL cold PBS followed by addition of 20 $\mu$L 10X anchor LMO or CMO pre-mixed with equimolar complimentary oligonucleotide conjugated to AF647 (final concentrations of 10 nM, 50 nM, 100 nM, 500 nM, or 1000 nM). Cells were incubated on ice for 5 minutes followed by addition of 20 µL of 10X stock corresponding co-anchor. The experiment was repeated three times, mean fluorescence intensity was calculated for each condition, and linear regression was performed. For exchange experiments, HEK293 cells were labeled with 200 nM LMOs or CMOs bearing FAM- or AF647-conjugated oligonucleotides. FAM- and AF647-labeled cells were then mixed and kept on ice for 2 hours in PBS with 1% BSA (2% for nuclei), during which cell aliquots were analyzed every 30 minutes. For room temperature experiments, cells were incubated for 30 minutes at room temperature and analyzed every 10 minutes. Label stability was computed as proportional differences between FAM or AF647 intensity relative to time zero. Off-target labeling was computed as FAM abundance on AF647-labeled cells (or vice versa). Fluorophore only controls were included in nuclei flow cytometry experiments because fluorophore-conjugated oligonucleotides demonstrate non-specific labeling. All analyses were performed in FlowJo and R.

For BSA quenching experiments, HEK293 cells were labeled with 200 nM LMOs or CMOs in 100 µL total volume PBS as described above. Prior to washing, each sample was diluted with ice cold PBS or PBS containing 1% BSA followed by centrifugation (160 rcf, 4 °C, 4 min). The 150 µL supernatant was removed from each primary labeling mixture and used to resuspend unlabeled HEK293 cells (secondary labeling). All primary and secondary labeled cells were washed 3X with ice cold PBS containing 1% BSA and analyzed by flow cytometry. Each secondary labeled sample was plotted as a proportion of the primary labeled sample (Fig. S1F). All analyses were performed in FlowJo and R.

Synthesis of lipid-modified oligonucleotides (LMOs): Oligonucleotides were synthesized on an Applied Biosystems Expedite 8909 DNA synthesizer, as previously described (Weber et al[24], Supplemental Information).

Specifically, Hexadecanoic (palmitic) acid, tetracosanoid (lignoceric) acid, N,N-diisopropylethylamine (DIPEA), N,N-diisopropylcarbodiimide (DIC), N,N-dimethylformamide (DMF), methylamine, ammonium hydroxide, and piperidine were obtained from Sigma-Aldrich. HPLC grade acetonitrile ($CH_3CN$), triethylamine ($NEt_3$), acetic acid, and anhydrous dichloromethane ($CH_2Cl_2$) were obtained from Fisher Scientific. 6-(4-Monomethoxytritylamino)hexyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite (5'-Amino-Modifier C6 Phopshoramidite), standard phosphoramidites, and DNA synthesis reagents were obtained from Glen Research. Controlled pore glass (CPG) supports (2-Dimethoxytrityloxymethyl-6-fluorenylmethoxycarbonylamino-hexane- 1-succinoyl)-long chain alkylamino-CPG (3'-Amino-Modifier C7 CPG 1000), 5'-Dimethoxytrityl-N-dimethylformamidine-2'-deoxyGuanosine, 3'-succinoyl-long chain alkylamino-CPG (dmf-dG-CPG 1000), and 5'-Dimethoxytrityl-N-Acetyl-2'-deoxyCytidine, 3'-succinoyl-long chain alkylamino-CPG (Ac-dC-

CPG 1000) synthesis columns were obtained from Glen Research. All materials were used as received from manufacturer.

For the anchor LMO, after synthesis of the DNA sequence, the 5' end was modified with an amine using 5'-Amino-Modifier C6 Phosphoramidite (100 mM) and a custom 15-minute coupling protocol. After synthesis of 5′ amino-modified DNA, the MMT protecting group was removed manually on the synthesizer by priming alternately with deblock and dry $CH_3CN$ at least three times until yellow color disappears. CPG beads were dried by priming several times with dry Helium gas. For the 3′ FMOC-protected amino-modified CPG, prior to oligonucleotide synthesis, the FMOC group was removed by suspending the CPG in a solution of 20% piperidine in dimethylformamide for 10 minutes at room temperature. The beads were then washed three times each with DMF and $CH_2Cl_2$. This procedure was repeated twice more to ensure complete deprotection of the FMOC protecting group prior to coupling to the fatty acid. Residual solvent was removed with reduced pressure on a Savant SPD121P SpeedVac System (ThermoFisher).

Fatty acid conjugation was performed on solid support by coupling the carboxylic acid moiety of the fatty acid to the 3' or 5' free amine—lignoceric acid and palmitic acid for the anchor and co-anchor, respectively. The solid support was transferred to a microcentrifuge tube and resuspended in a solution of anhydrous dichloromethane containing 200 mM fatty acid, 400 mM DIPEA, and 200 mM DIC. The microcentrifuge tubes were sealed with parafilm, crowned with a cap lock, and shaken overnight at room temperature. The beads were then washed 3X with $CH_2Cl_2$, 3X with DMF, and 2X $CH_2Cl_2$. Oligonucleotides were then deprotected and cleaved from solid support by suspending the resin in a 1:1 mixture of ammonium hydroxide and 40% methylamine (AMA) for 15 minutes at 65 °C with a cap lock followed by evaporation of AMA with a Savant SPD121P SpeedVac System. Cleaved oligonucleotides were dissolved in 0.7 mL of 0.1 M triethylammonium acetate (TEAA) and filtered through 0.2 µM Ultrafree-MC Centrifugal Filter Units (Millipore) to remove any residual CPG support prior to HPLC purification.

Fatty acid-modified oligonucleotides were purified from unmodified oligonucleotides by reversed-phase high-performance liquid chromatography (HPLC) using an Agilent 1200 Series HPLC System outfitted with a C8 column (Hypersil Gold, Thermo Scientific) and equipped with a diode array detector (DAD) monitoring at 230 and 260 nm. For HPLC purification, Buffer A was 0.1 M TEAA at pH 7 and buffer B was $CH_3CN$. running a gradient between 8 and 95% $CH_3CN$ over 30 minutes. Pure fractions were collected manually and lyophilized. The resulting powder was then resuspended in distilled water and lyophilized again two more times to remove residual TEAA salts prior to use. Purified fatty acid-modified oligonucleotides were resuspended in distilled water and concentrations were determined by measuring their absorbance at 260 nm on a Thermo-Fischer NanoDrop 2000 series.