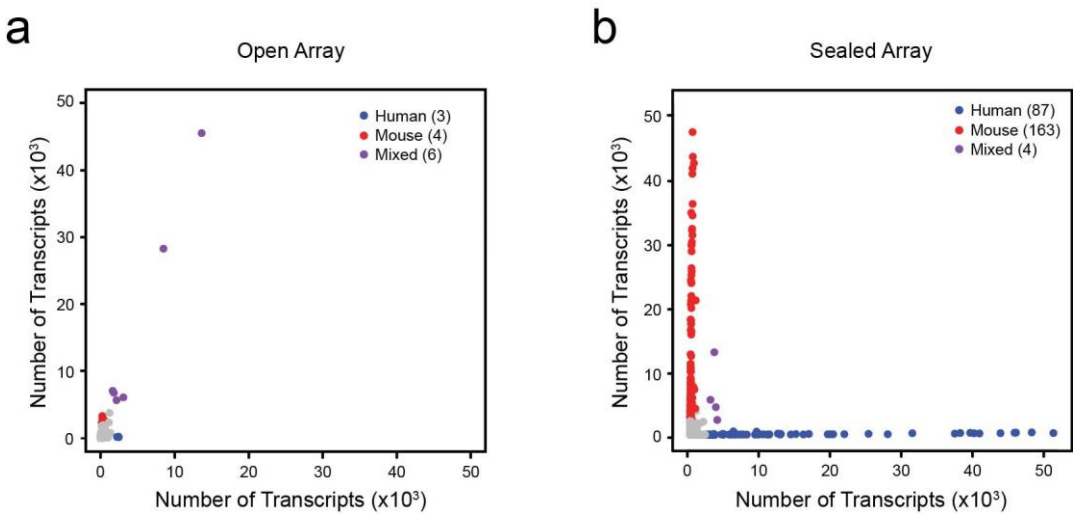
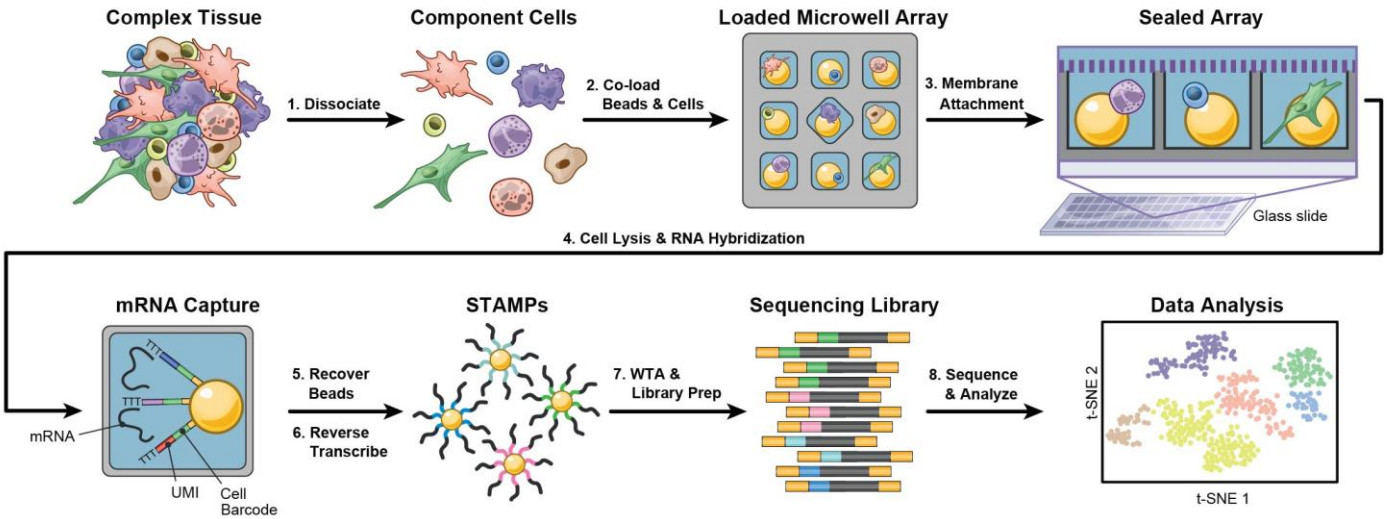


Supplementary Figure 1

Open Array Gene and Transcript Capture

a) An open array format results in decreased gene and transcript capture, and increased cross-contamination, relative to the membrane sealing implemented in Seq-Well. **(b)** Species mixing experiments with reversible membrane sealing using Seq-Well provides increased gene/transcript capture and improved single-cell resolution.

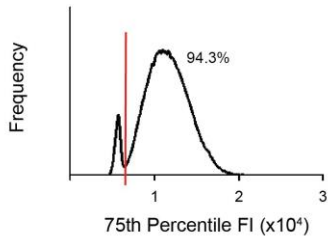


Supplementary Figure 2

Seq-Well Experimental Workflow (see Seq-Well Protocol – Ref. 21)

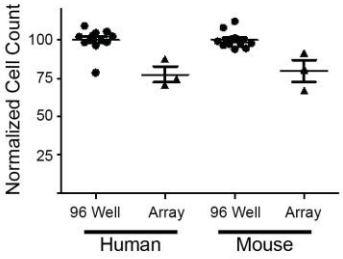
Cells are obtained from complex tissues or clinical biopsies, and digested to form a single-cell suspension. Barcoded mRNA capture beads are added to the surface of the microwell device, settling into wells by gravity, and then a single-cell suspension is applied**.** The device is sealed using a semi-permeable membrane that, upon addition of a chemical lysis buffer, confines cellular mRNAs within wells while allowing efficient buffer exchange. Liberated cellular transcripts hybridize to the bead-bound barcoded poly(dT) primers that contain a cell barcode (shared by all probes on the same bead but different between beads) and a unique molecular identifier (UMI) for each transcript molecule. After hybridization, the beads are removed from the array and bulk reverse transcription is performed to generate single-cell cDNAs attached to beads. Libraries are then made by a combination of PCR and tagmentation, and sequenced. After, single-cell transcriptomes are assembled *in silico* using cell barcodes and UMIs.

a



	Beads/well		
	0	1	>1
Array 1	5.0%	94.5%	0.5%
Array 2	5.7%	93.9%	0.4%

b



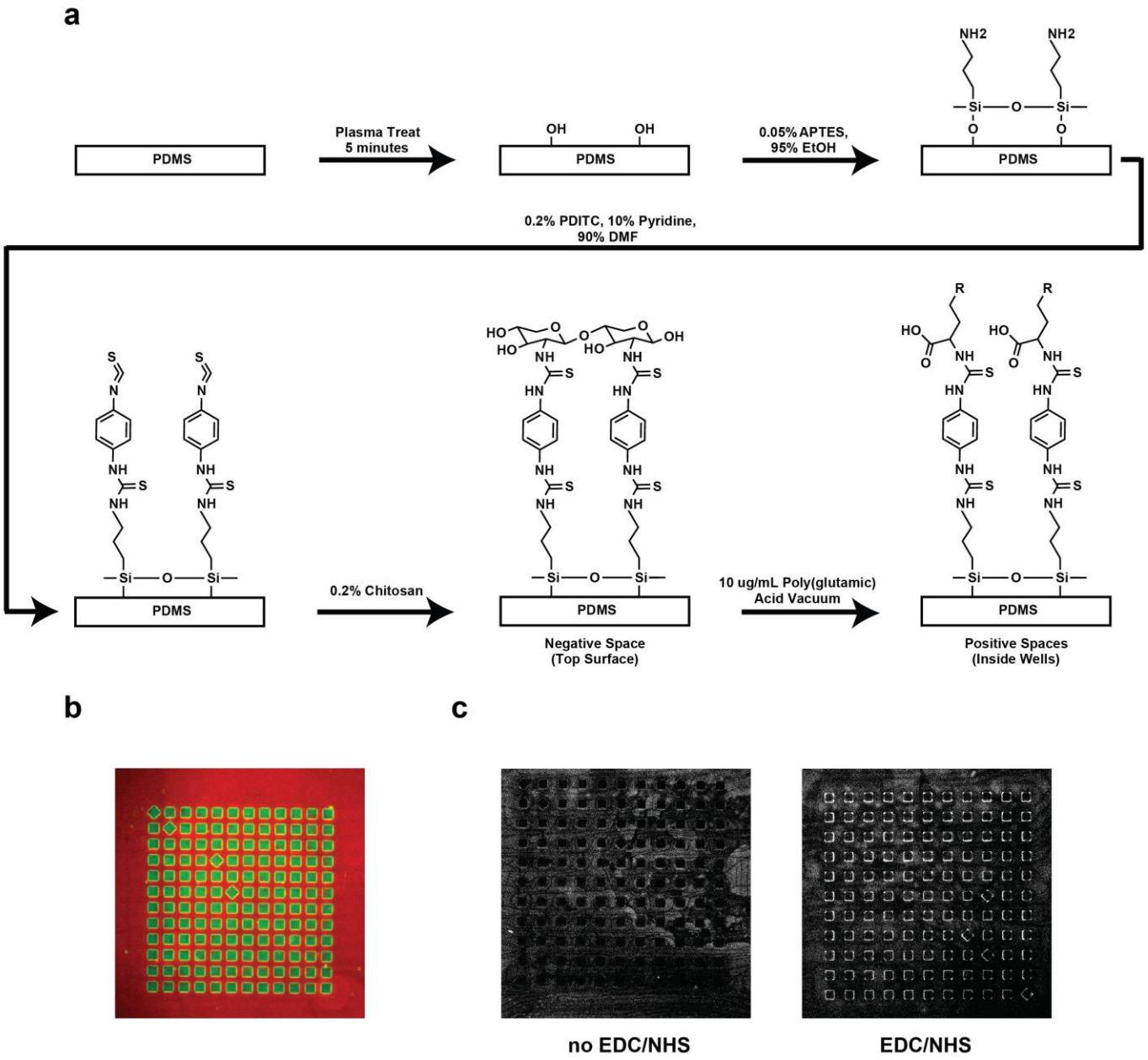
c

Loading Density	Empty	1 Human	1 Mouse	2+ Human	2+ Mouse	Human & Mouse Multiplet	Total Multiplet	Human & Mouse Multiplet Freq. (%)	Total Multiplet Freq. (%)
200	83,200 ± 293	134 ± 23	177 ± 53	0.33 ± 0.6	10.3 ± 17.9	0.7 ± 0.6	11.3 ± 19	0.1 ± 0.2	2.7 ± 4.4
2,000	81,100 ± 320	1,310 ± 150	1,080 ± 170	9.7 ± 3.2	23.3 ± 13	10.3 ± 3.1	43.3 ± 11	0.4 ± 0.1	1.8 ± 0.5
20,000	60,700 ± 560	13,100 ± 340	9,190 ± 250	976 ± 100	598 ± 23	1,100 ± 160	2,603 ± 250	4.8 ± 0.6	11.4 ± 0.9

Supplementary Figure 3

Bead And Cell Loading Efficiency

(a) Two arrays were loaded with barcoded beads through intermittent rocking. After washing, arrays were imaged in transmitted light and AF488 channel to capture bead autofluorescence. A plot of the frequency of the 75th percentile AF488 well intensity across the array (Panel 1) and the frequency of wells containing zero, one and multiple beads is displayed (Panel 2). **(b)** 200 μL of a 1:1 mix of fluorescently labeled human (HEK 293) and mouse (3T3) cell solution was loaded into 3 arrays and 12 wells of a 96 well plate. The number of cells loaded into each array and well as enumerated by fluorescent imaging is plotted, normalized to the average number of cells/well in the 96 well plate. Mean and standard error are denoted by line and error bars respectively. **(c)** 2x10^2, 2x10^3, and 2x10^4 total cells of a 1:1 mixture of fluorescently labeled HEK 293T and 3T3 cells were loaded onto three functionalized arrays each. All arrays were fluorescently imaged to enumerate the number of each cell line in each array microwell. The mean ± standard deviation of the number of empty, single and multiple occupancy wells across the three replicate arrays for each loading density is displayed along with the mean ± standard deviation of the percentage of occupied wells containing a cell from each species

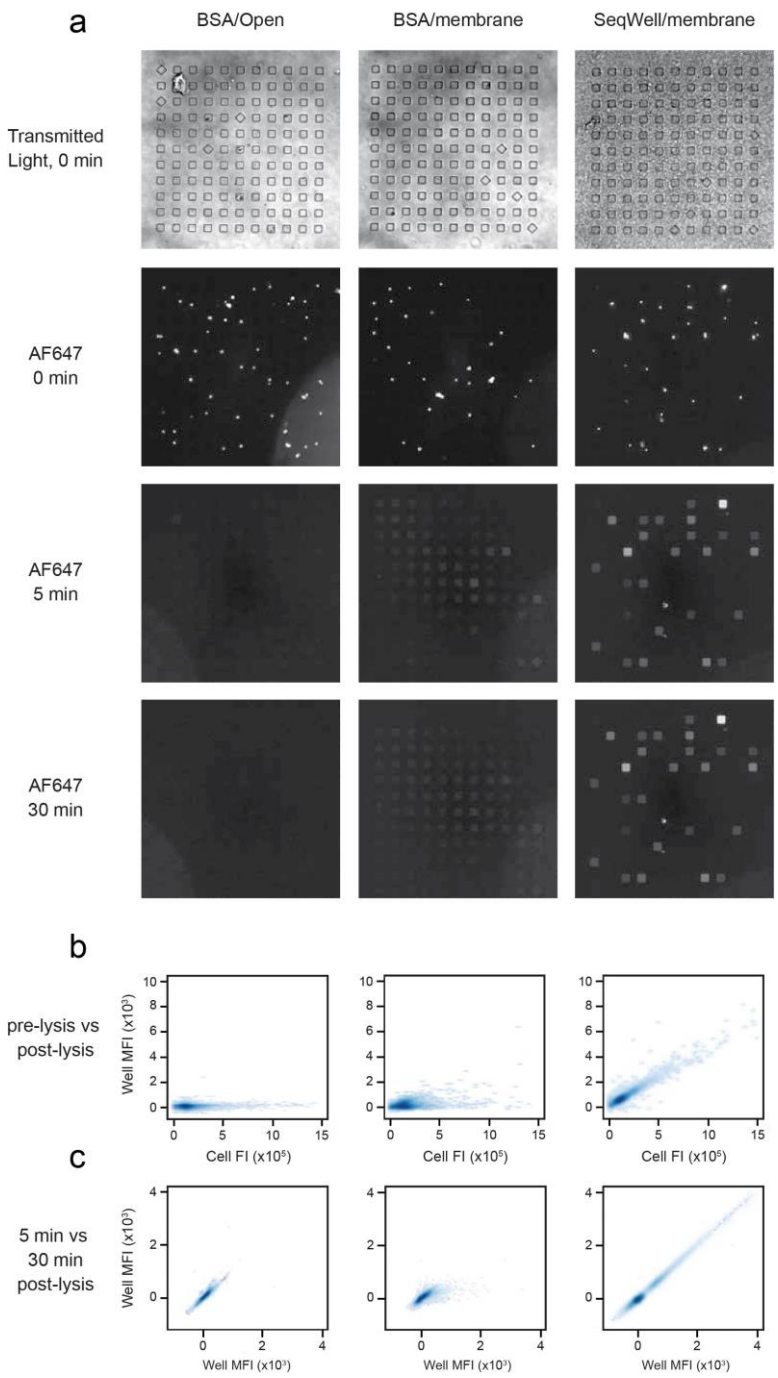


Supplementary Figure 4

PDMS Surface Chemistry Functionalization Protocol and Differential Functionalization of Microwell Arrays

(**a**) The surface of the PDMS device is initially treated with an air plasma under mild vacuum, terminating the surface in hydroxyls. This PDMS surface is aminated using (3-Aminopropyl)triethoxysilane (APTES). The amine surface is then activated with PDITC to create an isothiocyanate surface. The isothiocyanate on the top surface of the array (negative space) is covalently linked to chitosan polymers through their amine group. The hydrophobicity of the isothiocyanate surface prevents solvation of the microwells with the aqueous chitosan solution, preventing chitosan from reacting with the inner well surfaces (positive space). These surfaces are subsequently reacted with the free amine of poly(glutamic) acid polymers under vacuum to drive the solvation of the wells. (**b**) The top surface of a PDITC-activated array was coated with streptavidin-PE (red) and the inner well surfaces were coated with streptavidin-AF488 (green) using same method used to functionalize with chitosan and poly(glutamate). (**c**) Two chitosan/poly(glutamate) bifunctionalized arrays were submerged in MES buffer without (Panel 1) or with (Panel 2) 100 µg/mL EDC and 10 µg/mL NHS for 10 minutes. The arrays were washed and then submerged in PBS solution containing 1 µg/mL AF568-labeled antibody overnight. After washing, arrays were imaged

for AF568 fluorescence.

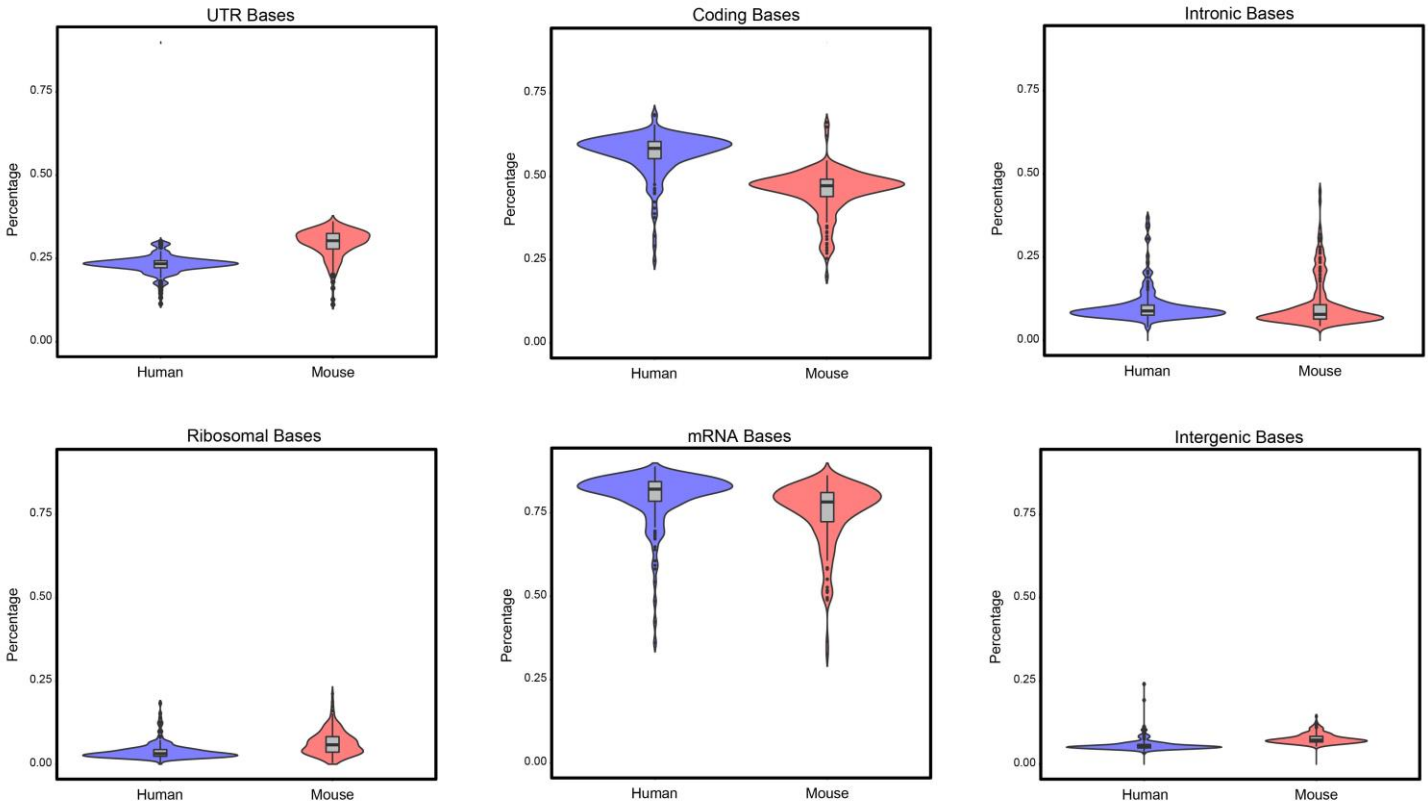


Supplementary Figure 5

Microwell Sealing With Semipermeable Membrane

PBMCs labeled with αCD45-AF647 were loaded into two BSA-blocked arrays and one array functionalized with chitosan and poly(glutamate). A semipermeable membrane was attached to one of the BSA-blocked arrays and the chitosan:polyglutamate functionalized array prior to addition of lysis buffer. (**a**) Example images of transmitted light and AF647 fluorescence of the arrays before, and 5 and 30 minutes after addition of lysis buffer are displayed for each array. (**b**) The total fluorescence intensity (FI) of all pixels associated with cells within a well is plotted against the median fluorescent intensity (MFI) of the volume of the same well 5 minutes after lysis for 12,100 wells from each array. (**c**) The MFI of the well volume 5 minutes after lysis is plotted against the MFI of

the volume of the same well 30 minutes after lysis for the same 12,100 wells from each array.

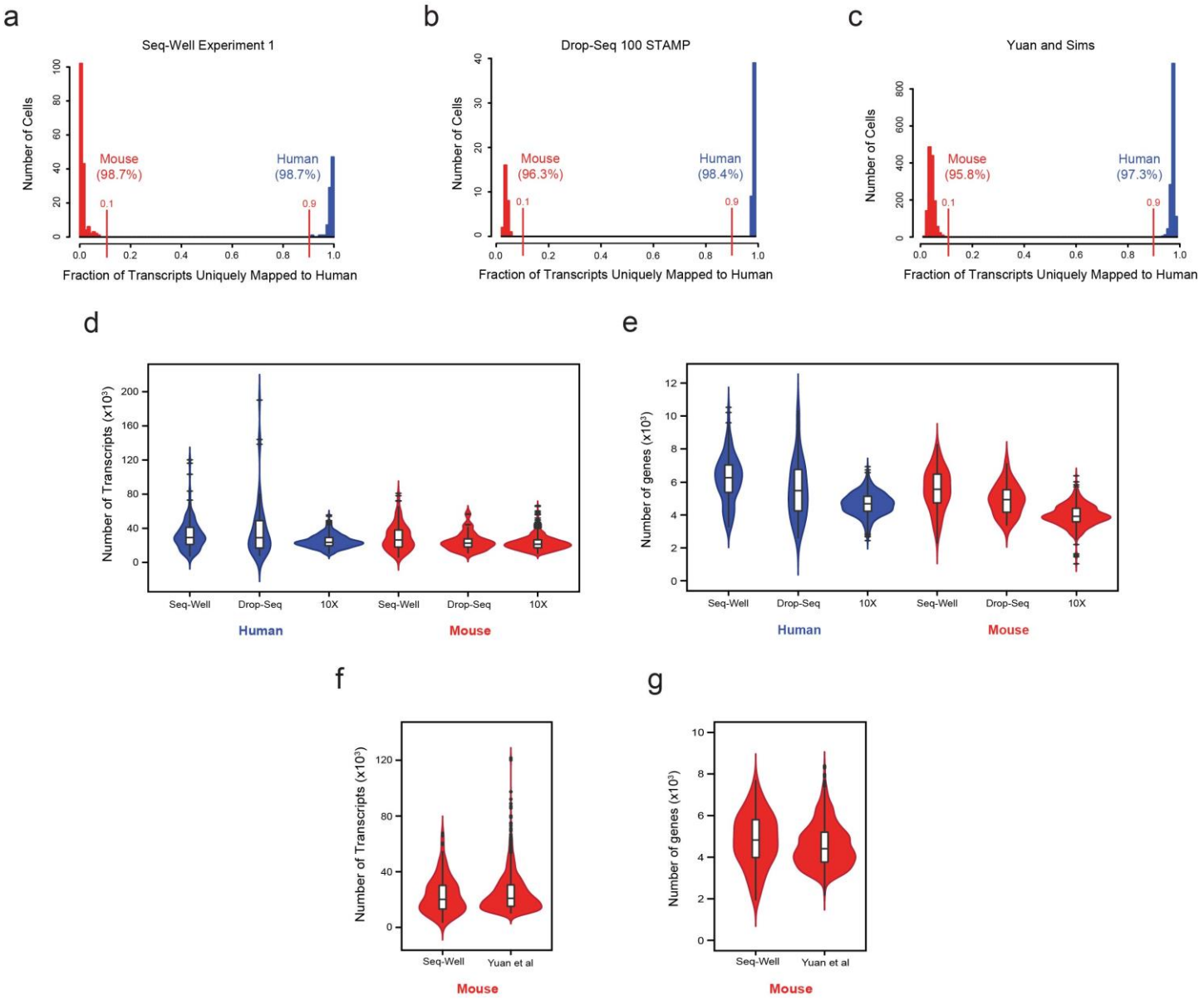


Species	Ribosomal Bases	mRNA Bases	Intergenic Bases	Intronic Bases	Coding Bases	UTR Bases
Human	3.86%	80.0%	5.79%	10.4%	56.9%	23.1%
Mouse	6.32%	75.2%	7.68%	10.9%	45.6%	29.6%

Supplementary Figure 6

Read Mapping Quality

Read mapping quality matrices were generated for each sample for human (blue) and mouse (red) cells, aligned to hg19 and mm10, respectively. High quality samples had relatively higher percentages of annotated genomic (genic) and exonic transcripts and low percentages of annotated intergenic and ribosomal transcripts (Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR).

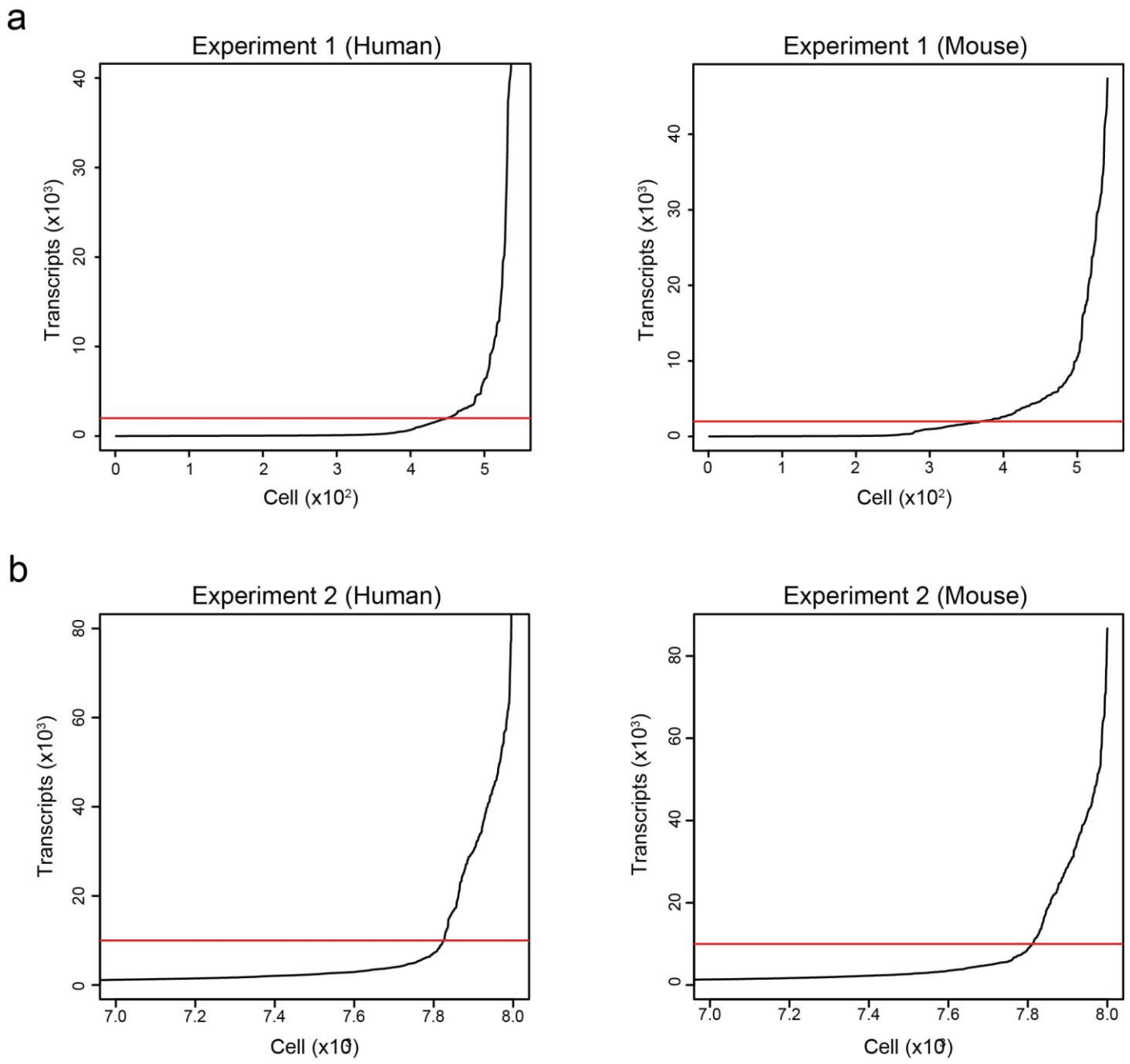


Supplementary Figure 7

Comparison of Gene and Transcript Capture and Percent Contamination Among Massively-Parallel scRNA-Seq Methods Using Mouse and Human Cell Lines

Histograms of the percent cross-species contamination in (**a**) Seq-Well, (**b**) Drop-Seq (Ref. 12), and (**c**) Yuan and Sims (Ref. 15). In each plot, cells with greater than 90% of human transcripts are displayed in blue and cells with less than 10% human transcripts are displayed in red. (**d**) Transcript capture in human (blue) and mouse (red) cell lines across three massively-parallel, bead-based single-cell sequencing platforms (Seq-Well, Drop-Seq, and 10X Genomics, with downsampling to an average read-depth of 80,000 reads per cell, consistent with 10X genomics data (Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR). We detect an average of 32,841 human transcripts and 29,806 mouse transcripts using Seq-Well compared to an average of 39,400 human transcripts and 24,384 mouse transcripts using Drop-Seq, an average of 24,751 human transcripts and 22,971 mouse transcripts using 10X Genomics (available from http://support.10xgenomics.com/single-cell/datasets/hgmm). (**e**) Gene detection across human and mouse cell lines across the same three single-cell sequencing platforms with down-sampling to the average read-depth of 80,000 reads per cell, consistent with 10X genomics (Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR). We detect an average of 6,174 human genes and 5,528 mouse genes using Seq-Well, an average of 5,561 human genes and 4,903 mouse genes using Drop-Seq and an average of 4,655 human genes and 3,950 mouse genes using 10X Genomics. (**f**) Downsampling to an average of 42,000 reads per cell consistent with data published in Yuan and Sims 2016, results in average

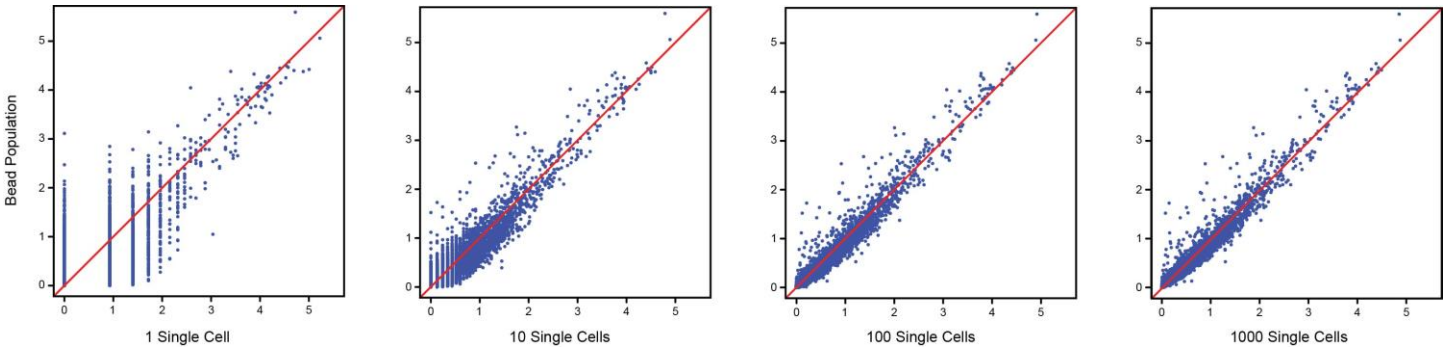
detection of 23,061 mouse transcripts using Seq-Well compared to an average of 24,761 mouse transcripts using the Yuan and Sims platform (Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR). (**g**) Downsampling to an average of 42,000 reads per cell results in average detection of 4,827 mouse genes using Seq-Well compared to an average of 4,569 mouse genes using the Yuan and Sims platform (Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR).



Supplementary Figure 8

Transcript Cutoff For Species-Mixing Validation

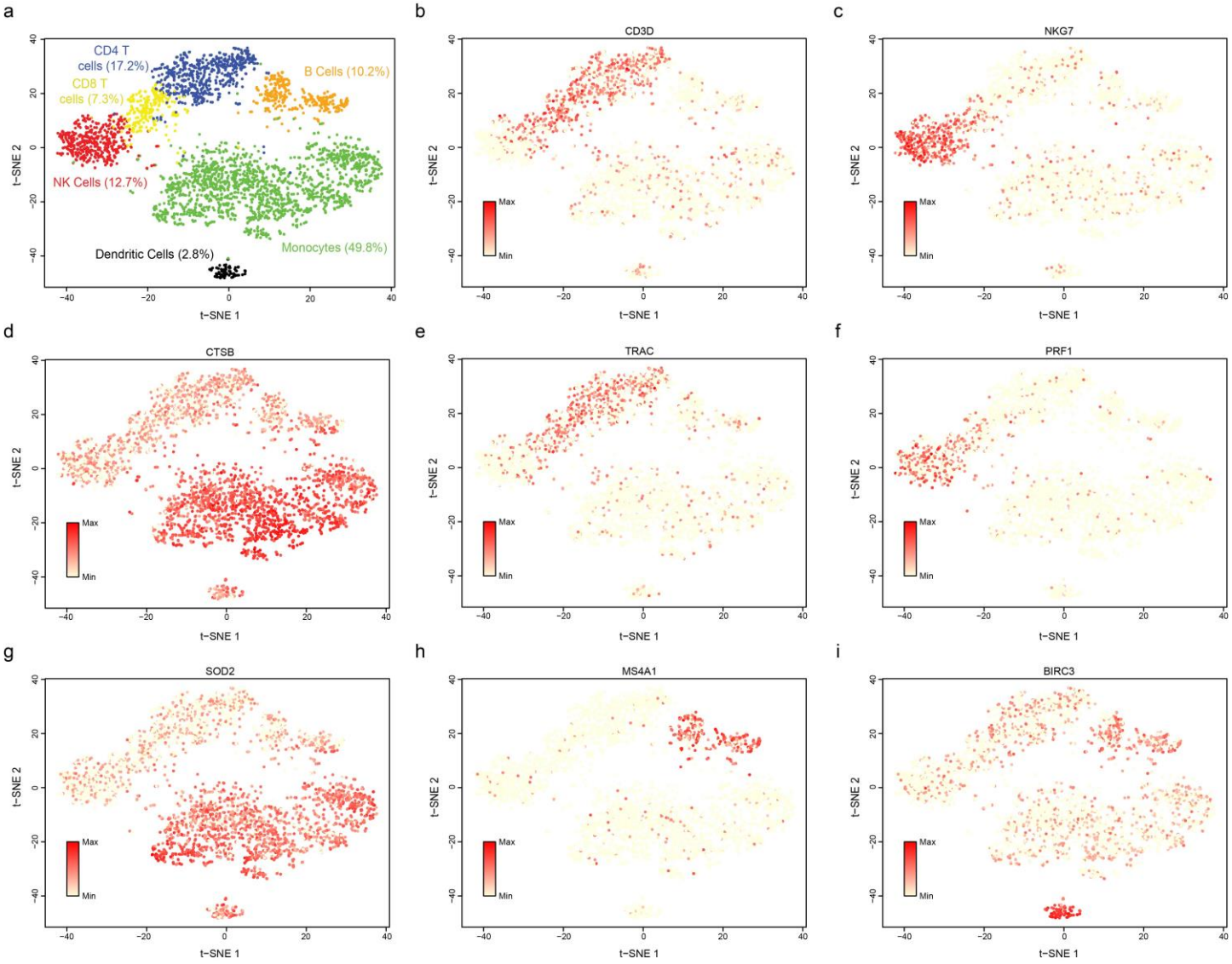
We sequenced two arrays (**a & b**) to confirm single-cell resolution and minimal cross-contamination between mouse and human cells. We called cells by plotting the cumulative distribution of transcripts and making a cutoff at the elbow in the curve. In the first experiment (**a**), which was used to validate our single-cell resolution, we shallowly sequenced the array and made the cutoff at 2,000 transcripts. In the second experiment (**b**), where we sequenced the array deeply to allow a competitive comparison to Drop-Seq, we made our cutoff at 10,000 transcripts.



Supplementary Figure 9

Comparison Of In-Silico HEK293 Populations With Bulk Populations

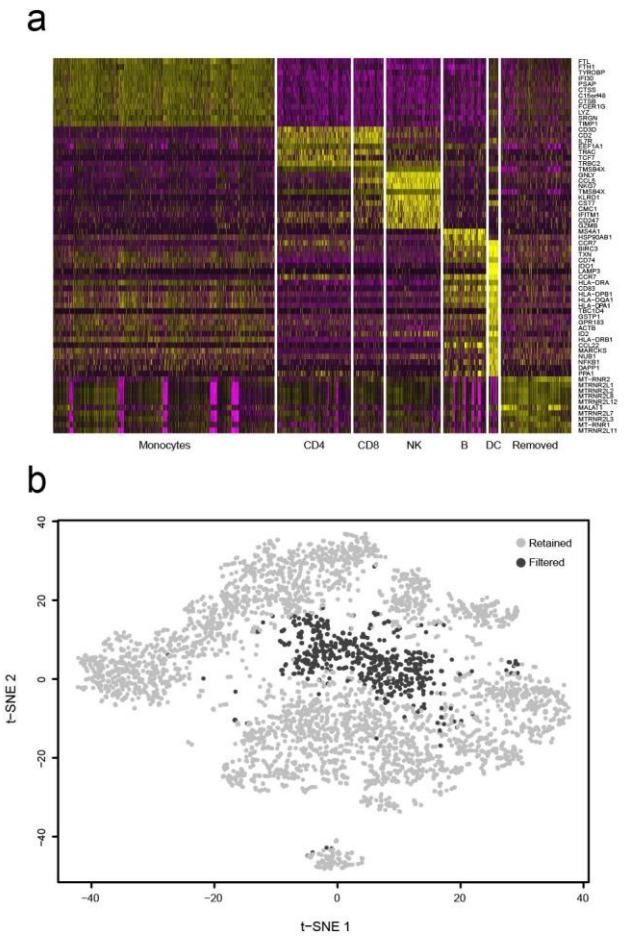
Scatterplots showing the correlation between gene expression estimates from bulk populations (40,000 HEK cells and 40,000 mRNA capture) and populations generated in-silico from 1, 10, 100, and 1,000 randomly-sampled single HEK293 cells (1 Cell: R = 0.751 ± 0.0726; 10 Cells: R = 0.952 ± 0.008; 100 Cells: R = 0.980 ± 0.0006; 1000 Cells: R = 0.983 ± 0.0001)**.**



Supplementary Figure 10

Mapping Lineage Defining Transcripts To PBMC Clusters

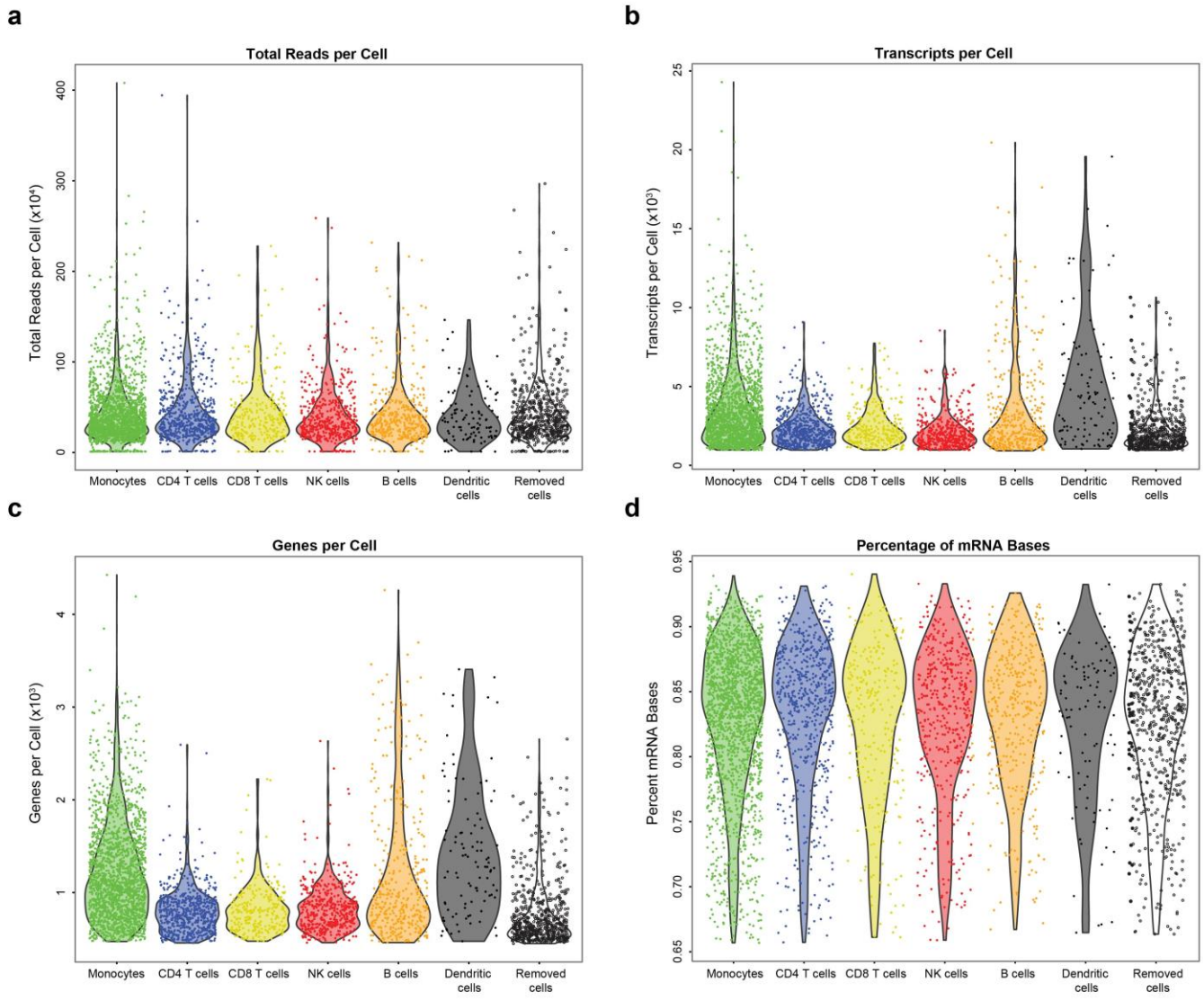
(**a**) Clusters identified through graph-based clustering (**Methods**) correspond to major immune cell populations. (**b,e**) CD4 T cells are characterized by expression of CD3D and T-cell receptor expression without pronounced expression of cytoxic genes NKG7 and PRF1. (**c,f**) CD8 T cells are defined by expression of NKG7 and PRF1. (**d,g**) Monocytes are defined by expression of cathepsin B (CTSB) and SOD2. (**e**) Natural killer cells are characterized by expression of cytotoxic genes in the absence of T cell receptor expression. (**h**) B cells are marked by elevated expression of MS4A1 (CD20) transcripts. (**i**) Dendritic cells are enriched for expression of BIRC3.



Supplementary Figure 11

Heatmap Of PBMCs

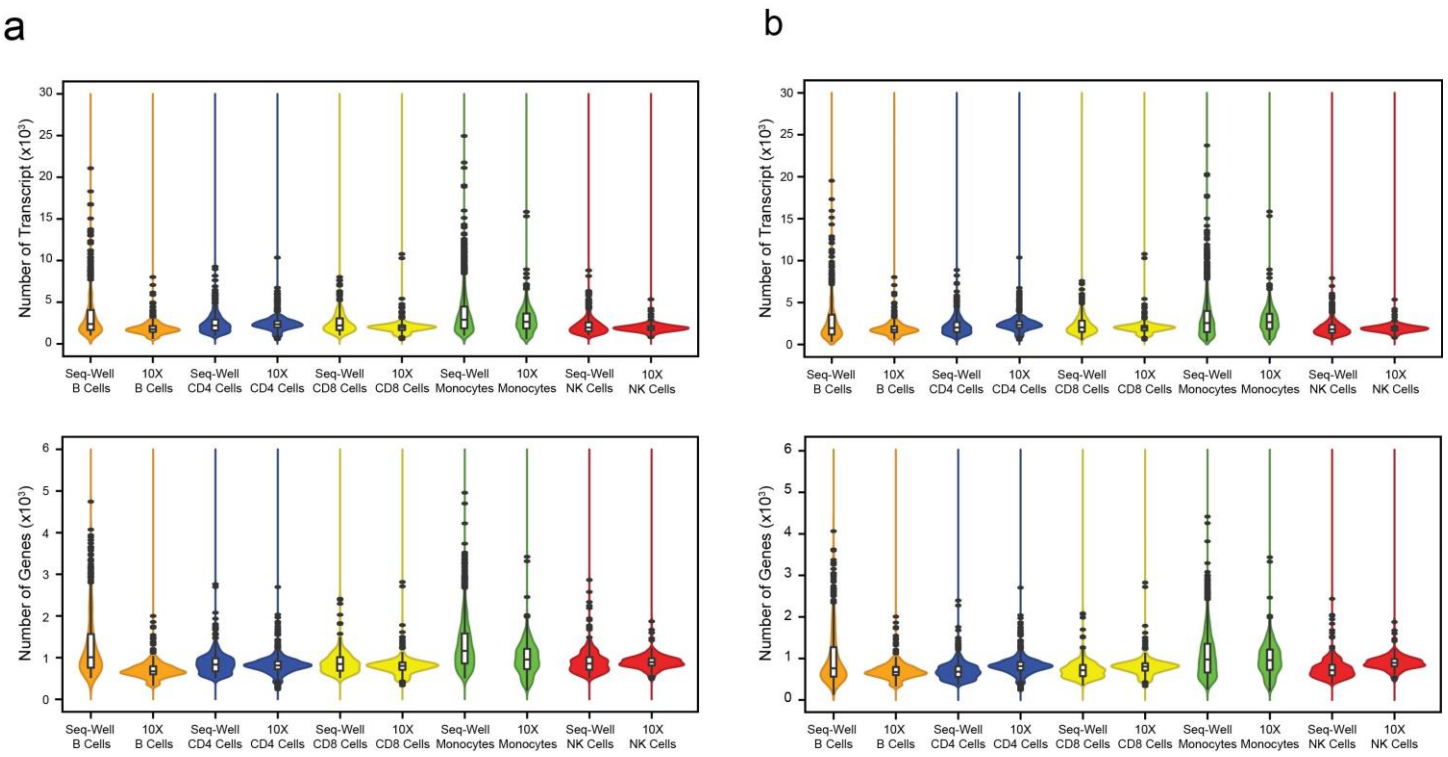
(**a**) Genes enriched in each cluster were identified using an “ROC” test in Seurat, comparing cells assigned to each cluster to all other cells. A heatmap was constructed using enriched genes found to define each cluster. One cluster of 602 cells that demonstrated exclusive enrichment of mitochondrial genes was removed as these likely represent low-quality or dying cells. (**b**) We generated a t-SNE projection of 4,296 cells with greater than 10,000 reads, 1,000 transcripts, 500 genes, and 65% transcript mapping. We removed a total of 602 cells from the final analysis found to be strongly enriched for expression of mitochondrial genes. The remaining 3,694 cells form distinct clusters enriched for lineage-defining that distinguish cells types from one another.



Supplementary Figure 12

Read Mapping Quality In PBMCs

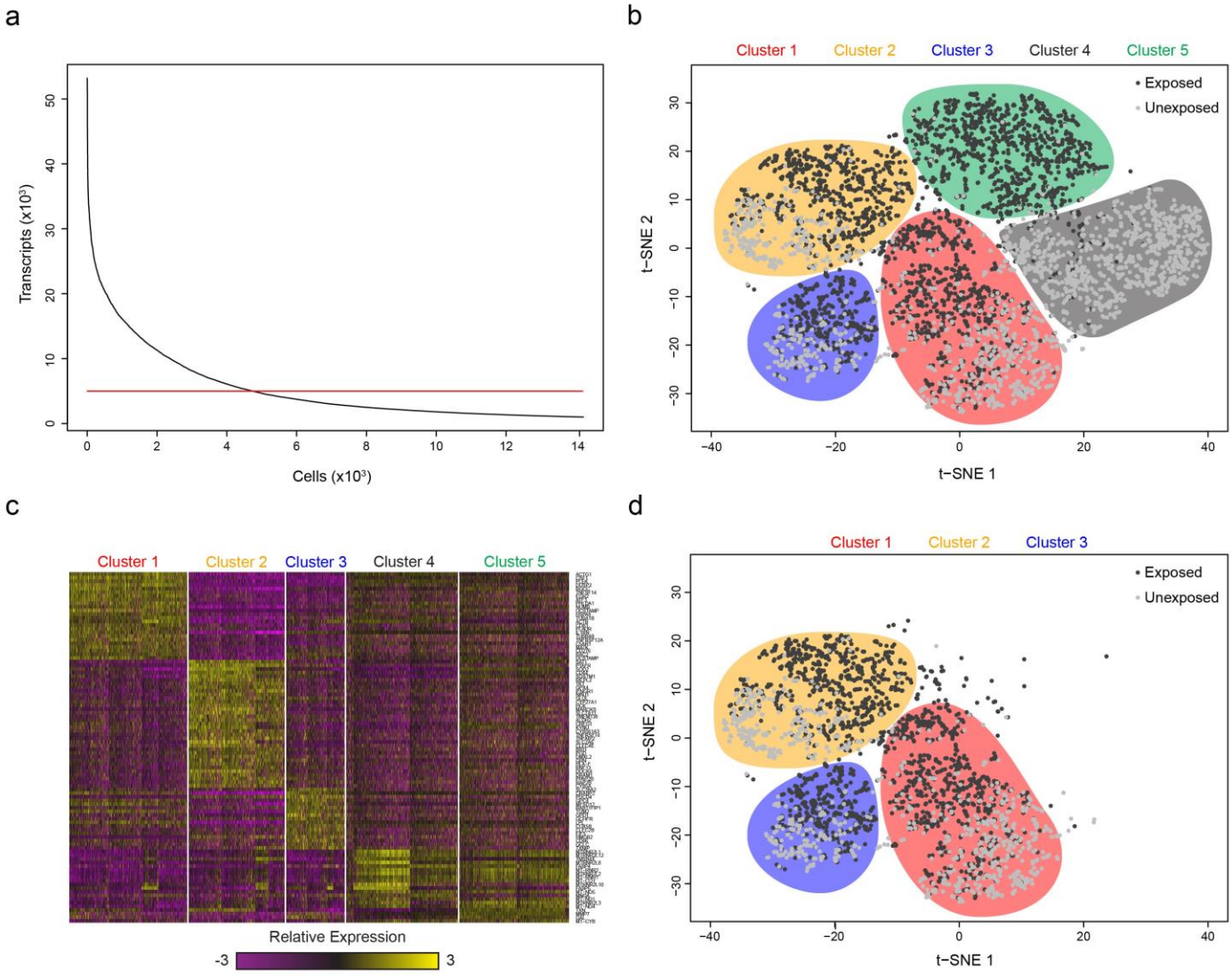
(**a-c**) Violin plots depicting (**a**) reads, (**b**) transcripts, and (**c**) genes per cell, separated by cell type. (**d**) Percent mRNA bases per cell, separated by cell type.



Supplementary Figure 13

Comparison Of Human PBMC Gene And Transcript Capture With Other Massively-Parallel scRNA-Seq Methods

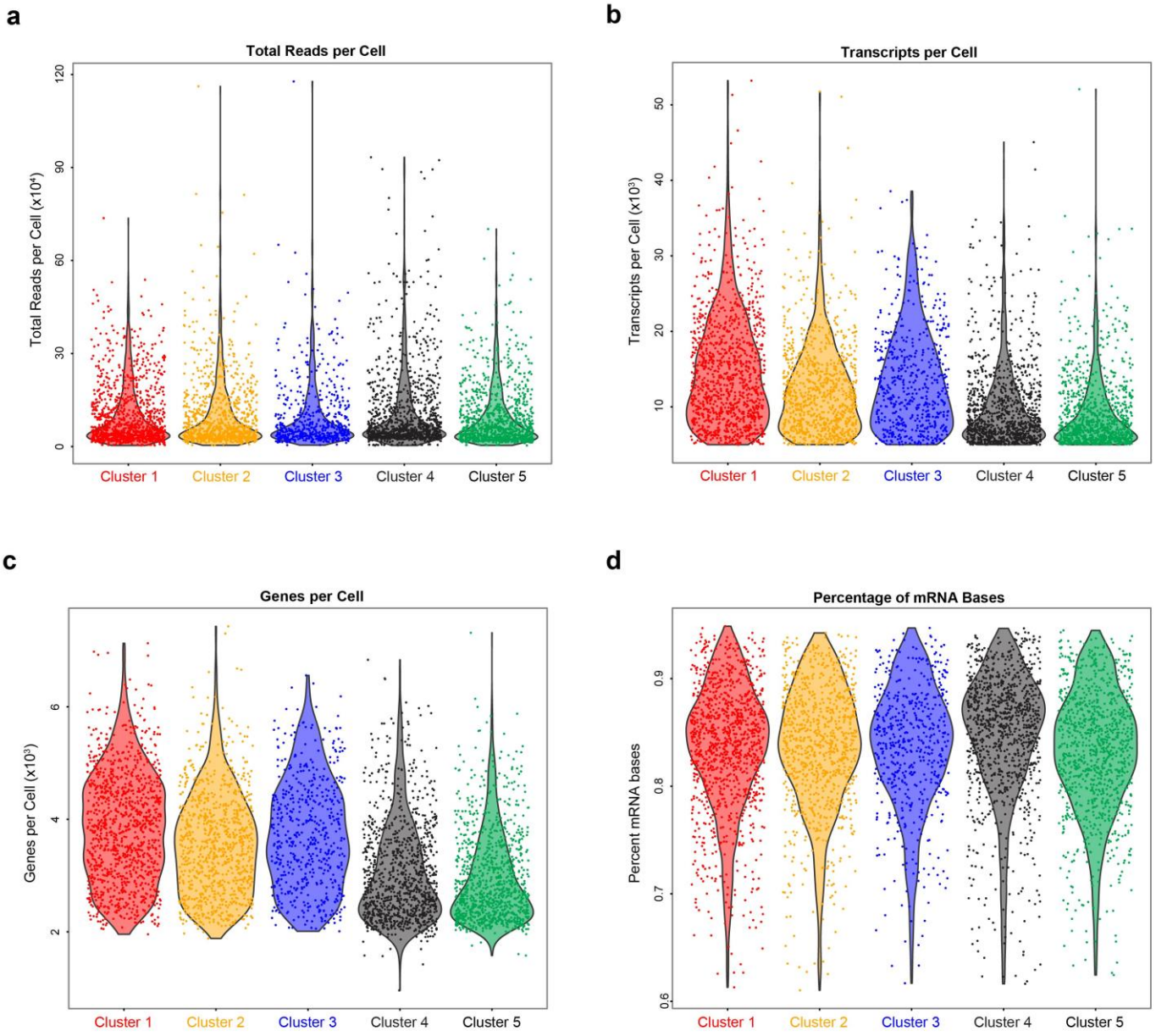
(a) Comparison of transcript capture (top) and gene detection (bottom) between Seq-Well and 10X Genomics within PBMC cell types prior to downsampling (colored as in Figure 2; Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR). Among B cells (orange), an average of 1,315 genes and 3,632 transcripts were detected using Seq-Well and an average of 710 genes and 1,910 transcripts were detected in 10X Genomics data. Among CD4 T cells (blue), an average of 861 genes and 2,444 transcripts were detected using Seq-Well and an average of 815 genes and 2,370 transcripts were detected in 10X Genomics data. Among CD8 T cells (yellow), an average of 885 genes and 2,574 transcripts were detected using Seq-Well and an average of 809 genes and 2,029 transcripts were detected in 10X Genomics data. Among Monocytes (green), an average of 1,288 genes and 3,568 transcripts were detected using Seq-Well and an average of 974 genes and 2,835 transcripts were detected in 10X Genomics data. Among NK cells (red), an average of 902 genes and 2,338 transcripts were detected using Seq-Well and an average of 907 genes and 1,943 transcripts were detected in 10X Genomics data. **(b)** Transcript capture (top) and gene detection (bottom) upon downsampling of Seq-Well data to an average read depth 69,000 reads per cell (Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/- 1.5 IQR; Points: Values > 1.5 IQR). Upon downsampling, in Seq-Well, an average of 1,048 genes and 3,103 transcripts were detected among B cells, 735 genes and 2,221 transcripts among CD4 T cells, 763 genes and 2,353 transcripts among CD8 T cells, 1,052 genes and 3,105 transcripts among monocytes, and 789 genes and 2,041 transcripts among NK cells.



Supplementary Figure 14

T-SNE Visualization Of Exposed And Unexposed Macrophages Using A 5,000 Transcript Cutoff

(**a**) Using a threshold of 5,000 detected transcripts, we identified 4,638 macrophages. (**b**) Among these 4,638 cells, we identified 5 distinct clusters of macrophages by performing graph-based clustering over 5 principal components (377 variable genes). (**c**) Clusters 1-3 are defined by unique gene expression signatures, while Clusters 4 and 5 are defined by expression of mitochondrial genes, suggesting low-quality cells. (**d**) Following removal of cells within Clusters 4 and 5, there remain a total of 2,560 cells in Clusters 1-3.



Supplementary Figure 15

Quality By Cluster Among TB Macrophages

(**a-c**) Violin plots depicting (**a**) reads, (**b**) transcripts, and (**c**) genes per cell, separated by cluster. (**d**) Percent mRNA bases per cell, separated by cluster.

SUPPLEMENTARY TABLES AND LEGENDS FOR:

Seq-Well: A Portable, Low-Cost Platform for High-Throughput Single-Cell RNA-Seq of Low-Input Samples

Todd M. Gierahn[1,#], Marc H. Wadsworth II[2,3,4,#], Travis K. Hughes[2,3,4,#], Bryan D. Bryson[4,5], Andrew Butler[6,7], Rahul Satija[6,7], Sarah Fortune[4,5], J. Christopher Love[1,3,4,*], and Alex K. Shalek[2,3,4,*]

[1] Koch Institute for Integrative Cancer Research, MIT, Cambridge, Massachusetts, USA

[2] Institute for Medical Engineering & Science (IMES) and Department of Chemistry, MIT, Cambridge, Massachusetts, USA

[3] Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

[4] Ragon Institute of MGH, MIT and Harvard, Cambridge, Massachusetts, USA

[5] Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, USA

[6] Center for Genomics and Systems Biology, Department of Biology, New York University, New York City, New York, USA

[7] New York Genome Center, New York City, New York, USA

[#] These authors contributed equally to this work

* These senior authors contributed equally to this work

* To whom correspondence should be addressed: shalek@mit.edu (AKS), clove@mit.edu (JCL)

Nature Methods: doi:10.1038/nmeth.4179

Oligo Name	Sequence 5’ to 3’
1. Barcoded Bead Sequence	5’–Bead–Linker-TTTTTTTAAGCAGTGGTATCAAC GCAGAGTACJJJJJJJJJJJJNNNNNNNN TTTTTTTTTTTTTTTTTTTTTTTTTTTT-3’
2. Template Switching Oligo	AAGCAGTGGTATCAACGCAGAGTGAATrGrGrG
3. SMART PCR Primer	AAGCAGTGGTATCAACGCAGAGT
4. New P5-SMART PCR Hybrid Oligo	AATGATACGGCGACCACCGAGATCTACACGCCT GTCCGCGGAAGCAGTGGTATCAACGCAGAGT* A*C
5. Custom Read 1 Primer	GCCTGTCCGCGGAAGCAGTGGTATCAACGCAG AGTAC

Supplementary Table 1 | Oligo Sequences. Sequences of oligos used in Seq-Well. (**1**) The barcoded bead sequence is constructed on the surface of the bead, and cell barcodes are generated through split and pool synthesis. (**2**) The template switching oligo (TSO) is used to tag the 5’ end of captured mRNA using a reverse transcriptase enzyme with terminal transferase activity. (**3**) Sequence for PCR primer used to perform whole-transcriptome amplification (WTA) PCR reaction following reverse transcription and ExoI digestion. (**4**) Sequence that selectively primes the bead-specific SMART sequence during the post-tagmentation step-out PCR, which appends a P5 sequencing adapter. (**5**) Primer used during sequencing that selectively primes the bead-specific primer site to initiate sequencing of the barcode and UMI in Illumina Read 1.

Name	Excitation, nm	Emission, nm	Exposure, ms	Intensity, %	Gain, abs.
Transmitted Light	-	-	50	5	10
CerCP710	485	725/40	100	100	50
PECy5.5	560	725/40	100	100	20
APCCy7	650	775LP	100	100	100
PECy7	560	775LP	100	100	20
PECy5	560	680/42	150	100	10
PerCP650	485	680/42	100	100	30
AF647	650	680/42	150	100	10

Nature Methods: doi:10.1038/nmeth.4179

AF568	560	607/36	100	100	10
AF488	485	525/39	100	100	10

Supplementary Table 2 | Microscope Settings. The excitation light wavelengths, emission filters, exposure times, light source intensity and camera gain settings used to capture the fluorescence of the indicated fluorophore are displayed.

Supplementary Table 3 | Gene Expression Matrix for PBMCs. UMI count matrix for the 4,296 PBMCs, labeled by array, that had at least 10,000 reads, 1,000 transcripts, and 500 genes, with at least 65% bases mapping to the transcriptome.

Supplementary Table 4 | PBMC Cluster Enrichments. Lists of genes enriched within each PBMC cluster (B cells, CD4 T Cells, CD8 T cells, Dendritic cells, Monocytes, NK Cells) based on a likelihood-ratio test in which members of each cluster are compared to members of all other clusters.

Supplementary Table 5 | Gene Expression Matrix for Mtb-Exposed Monocyte-Derived Macrophages and Unexposed Control Cells. UMI count matrix for the 4,638 monocyte-derived macrophages, labeled by exposure, that had at least 5,000 mapped transcripts.

Supplementary Table 6 | TB Cluster Enrichments. We examined sets of enriched sets of genes among all cells (irrespective of TB exposure) within Cluster 1, 2 and 3 using the find.markers function in Seurat, which implements a 'roc' test to identify relative expression differences. (**a**) From this analysis, we identified sets of genes exclusively enriched in each cluster but not the others. For each of the identified gene sets, we performed gene set enrichment analysis in DAVID and GSEA. For analysis in DAVID, we compared each gene list for enrichment among GO terms and curated pathways against a background list of 9045 genes contained in the DAVID database that were detected in at least 5% of cells. For analysis using GSEA, we compared each gene list to the complete database of gene sets contained within GSEA. (**b, c**) Within Cluster 1, we observe unique enrichment of 134 genes related to TNF-alpha signaling,

Nature Methods: doi:10.1038/nmeth.4179

inflammation, immune response and LPS response among 1099 cells. (**d, e**) In Cluster 2, we observe exclusive enrichment of 251 genes among 904 cells that distinguish monocytes from dendritic cells in culture and characterize TNF-alpha signaling. (**f, g**) Finally, in Cluster 3, we observe unique enrichment of 118 genes among 557 cells related to specifically to hypoxia, LPS stimulation, TNF-alpha signaling and apoptosis.

Supplementary Table 7 | Cluster Enrichments between Exposure Groups. Initially, we separately identified enriched genes among exposed and unexposed cells within Clusters 1, 2 and 3 using the find.markers function in Seurat. In total, we identified 18 genes with conserved enrichment among TB exposed cells within Clusters 1, 2 and 3. For each cluster, we identified the set of genes enriched among exposed cells and unexposed cells within each cluster. (**a**) We identified 28 enriched genes that were conserved among exposed cells across clusters and 31 conserved genes among unexposed, of which 18 were conserved between exposed and unexposed. We identified 38 genes uniquely enriched among exposed cells in Cluster 1 and 54 genes uniquely enriched among unexposed cells, of which 5 were conserved between exposed and unexposed cells within Cluster 1. We identified 134 genes unique to Cluster 2 among exposed cells and 200 genes unique to Cluster 2 among unexposed cells, of which 35 were conserved between exposed and unexposed cells within Cluster 2. In Cluster 3, we identified 43 genes unique to cluster 3 among exposed cells and 44 genes unique to Cluster 3 among unexposed cells, of which 9 were conserved between exposed and unexposed cells within Cluster 3. We performed gene set enrichment analyses for single gene list using DAVID and GSEA. In DAVID, we specified a background list of 9045 genes and examined enrichments within GO terms and curated pathways. For the analysis in GSEA, we made comparisons of gene lists to the complete database of gene sets within GSEA. (**b, c**) Among the 18 genes conserved across clusters in both exposed and unexposed cells, we observed strong enrichment for LPS response, TNF-alpha signaling, phagosome formation and macrophage activation. (**d, e**) Among the 5 genes unique to Cluster 1 conserved between exposed and unexposed cells, we observed enrichment of PI3K-Akt signaling and immune activation. (**f, g**) Among the 35 genes unique to Cluster 2 conserved between exposed and unexposed cells, we observed enrichment of genes related to monocyte culture and the coagulation

Nature Methods: doi:10.1038/nmeth.4179

cascade. (**h, i**) Among the 9 genes unique to Cluster 3 conserved between exposed and unexposed, we observed enrichment of genes up-regulated by HGF and apoptosis.

Supplementary Table 8 | Differentially Expressed Genes between TB Exposed and Unexposed Cells within Each Cluster. We performed a likelihood ratio test to identify genes differentially expressed between TB exposed and unexposed cells within each cluster. (**a**) Differential expression results between 673 TB-exposed and 426 unexposed cells in Cluster 1. (**b**) Differential expression results between 627 TB-exposed and 277 unexposed cells in Cluster 2. (**c**) Differential expression results between 386 TB-exposed and 171 unexposed cells in Cluster 3.

Supplementary Table 9 | TB Infection by Cluster Enrichments. Initially, we performed a LRT within each cluster to identify genes differentially expressed between TB exposed and unexposed cells. For each cluster, we created lists of genes differentially expressed with p-values less than 5.0×10^{-6} within each cluster (**Figure 3c**). (**a**) We then compared these lists to identify genes that are differentially expressed genes exclusively within each cluster. We also identified a list of 37 genes detected as differentially expressed across all clusters. We then performed gene set enrichment analysis in DAVID and GSEA to examine functional enrichment of the identified gene sets (i.e. Genes conserved across and unique to each cluster). We performed analysis in DAVID, comparing genes 37 conserved genes, 22 genes unique to Cluster 1, 142 genes unique to Cluster 2, and 40 genes unique to Cluster 3 to a background list of 9,381 genes expressed in at least 5% of filtered cells (**Methods**). Using GSEA, we made comparisons of the above gene list to the complete list of curated gene sets within the GSEA database (MSigDB v5.1: http://software.broadinstitute.org/gsea/msigdb/index.jsp). (**b,c**) Within Cluster 1, we observed unique enrichment of genes related to growth, proliferation and cell cycle. (**d, e**) Within Cluster 2, we observed enrichment of genes that identify monocyte and dendritic cell culture in addition to proliferation. (**f, g**) Within Cluster 3, we observed unique enrichment of genes related to hypoxia, oxidative stress and oxygen homeostasis.

Nature Methods: doi:10.1038/nmeth.4179

Supplementary Table 10 | GSEA Comparisons of Exposed and Unexposed Cells within Each Cluster. We performed comparisons between *M. tuberculosis* exposed and unexposed cells within each cluster using GSEA. For each cluster we created .gct files containing normalized expression data for every cells within each cluster and assigned phenotypes (i.e. TB exposed vs. unexposed) to each cell using a .cls file. We then performed gene set enrichment analysis for each cluster across the complete gene set database in GSEA with 1000 permutations of assigned phenotype. (**a,b**) In cluster 1, we observed enrichment of dendritic cell maturation, monocytes in culture, response to L. donovani, and TNF-alpha signaling among 673 TB exposed cells and relative enrichment of ribosomal genes and protein synthesis among 426 unexposed cells. (**c,d**) In Cluster 2, we observed enrichment of LPS response, dendritic cell maturation, IL1 stimulation and response to TGF-beta among 627 TB exposed cells and relative enrichment of housekeeping functions, ribosomal genes and translation among 277 TB unexposed cells. (**e,f**) In Cluster 3, we observed enrichment of among delayed response to LPS (48 response), TLR7/8 stimulation, inflammatory response, intracellular infection and TNF signaling among 386 TB exposed cells and relative enrichment of housekeeping functions (ribosome, translation, actin) among 171 unexposed cells. (**g,h**) In Cluster 4, we observed enrichment of mitochondrial gene signatures, oxidative phosphorylation, hypoxia response and interferon response among 74 TB exposed cells and enrichment of ribosomal genes and translation among 975 unexposed cells. (**i,j**) In Cluster 5, we observed enrichment of LPS stimulation, TNF signaling, sepsis and dendritic cells maturation among 988 TB exposed cells and enrichment of ribosomal proteins and translation among 41 unexposed cells.

Nature Methods: doi:10.1038/nmeth.4179