DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

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

Citation
Gierahn, Todd M., Marc H. Wadsworth, Travis K. Hughes, Bryan D. Bryson, Andrew Butler, Rahul Satija, Sarah Fortune, J. Christopher Love, and Alex K. Shalek. 2017. "Seq-Well: A Portable, Low-Cost Platform for High-Throughput Single-Cell RNA-Seq of Low-Input Samples." Nature methods 14 (4): 395-398. doi:10.1038/nmeth.4179. http://dx.doi.org/10.1038/nmeth.4179.

Published Version
doi:10.1038/nmeth.4179

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:34375305

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Methods. Author manuscript; available in PMC 2017 August 13.

Published in final edited form as:

Nat Methods. 2017 April ; 14(4): 395–398. doi:10.1038/nmeth.4179.

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

Abstract

Single-cell RNA-Seq can precisely resolve cellular states but application to sparse samples is challenging. Here, we present Seq-Well, a portable, low-cost platform for massively-parallel single-cell RNA-Seq. Barcoded mRNA capture beads and single cells are sealed in an array of subnanoliter wells using a semi-permeable membrane, enabling efficient cell lysis and transcript capture. We characterize Seq-Well using species-mixing experiments and PBMCs, and profile thousands of primary human macrophages exposed to tuberculosis.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*To whom correspondence should be addressed: shalek@mit.edu (AKS), clove@mit.edu (JCL).
#These authors contributed equally to this work
*These senior authors contributed equally to this work

DATA AVAILABILITY.

All RNA-Seq data are available in GEO under Accession Number GSE92495.

AUTHOR CONTRIBUTIONS.

TMG, MHW, TKH, JCL, and AKS developed the concepts and designed the study. TG, MHW, TKH, and BDB performed the experiments. All authors analyzed and interpreted the data. TMG, MHW, TKH, JCL, and AKS wrote the manuscript with feedback from all authors.

COMPETING FINANCIAL INTERESTS.

T.M. Gierahn, M.H. Wadsworth II, T.K. Hughes, J.C. Love, A.K. Shalek, and Institutions The Broad Institute and the Massachusetts Institute of Technology have filed a patent application that relates to Seq-Well, compositions of matter, the outlined experimental and computational methods, and uses thereof.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

MAIN

The emergence of single-cell genomics has empowered new strategies for identifying the cellular and subcellular drivers of biological phenomena[1–19]. Patterns in genome-wide mRNA expression measured by single-cell RNA-Seq (scRNA-Seq) can be leveraged to uncover distinct cell types, states and circuits within cell populations and tissues[1–5,9–13]. The unprecedented view of cellular phenotypes scRNA-Seq affords could help transform our understanding of healthy and diseased behaviors, and guide the rational selection of precision diagnostics and therapies, if it could be broadly and easily applied to low-input ($\leq 10^4$ cells) clinical specimens.

Typically, scRNA-Seq has involved isolating and lysing individual cells, then independently reverse transcribing and amplifying their mRNA before generating barcoded libraries that are pooled for sequencing. Although manual picking[2,5,8], FACS-sorting[1,3,4] or integrated microfluidic circuits[7,9,10] can isolate single cells, one-cell-one-sample approaches are constrained fundamentally in scale by costs, time, and labor. Recently, massively-parallel methods have emerged that assign unique barcodes to each cell's mRNAs during reverse transcription, enabling ensemble processing while retaining single-cell resolution. These methods typically yield single-cell libraries of lower complexity, but higher throughput reduces the impact of the technical and intrinsic noise associated with each cell in analyses[11,12]. The most common variant is microfluidic devices that generate reverse-emulsion droplets to serially couple single cells with uniquely-barcoded mRNA capture beads[11,12]. Droplet-based techniques, however, can have inefficiencies in encapsulation, introduce technical noise through differences in cell lysis time, and require specialized equipment, limiting where, when, and with what scale scRNA-Seq can be performed.

One alternative is to use arrays of subnanoliter wells loaded by gravity. Operational simplicity reduces the need for peripheral equipment, decreases dead volumes, and facilitates parallelization. As proof-of-principle, cells and beads have been co-confined in unsealed nanowell arrays to perform targeted single-cell transcriptional profiling[13], yet the use of an open-array format significantly limits capture efficiency and increases cross-contamination (Supplementary Fig. 1). To avoid these issues, nanowells have also been combined with microfluidic channels that facilitate oil-based single-cell isolation via fluid exchange[14]. Nevertheless, this design limits buffer exchange and necessitated integrated temperature and pressure controllers, impacting ease-of-use and portability[15]. Semi-porous-membrane-covered nanowells have been used to link pairs of specific transcripts from single cells[16]; however, transcript capture and sealing efficiency were not addressed, and unique single-cell libraries were not achieved using many beads per well.

To overcome these assorted challenges, we have developed Seq-Well, a portable, simple platform for massively-parallel scRNA-Seq (Supplementary Fig. 2). Similar to previous nanowell-based implementations, Seq-Well confines single cells and barcoded poly(dT) mRNA capture beads in a PDMS array of ~86,000 subnanoliter wells. Designing well dimensions to accommodate only one bead enables single-bead loading efficiencies of ~95% (Figure 1a, Supplementary Fig. 3a; Supplementary Video 1). A simplified cell-loading scheme, in turn, enables capture efficiencies around 80% (**Methods**; Supplementary Fig.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3b), with a rate of dual occupancy that can be tuned by adjusting the number of cells loaded and visualized prior to processing (Supplementary Fig. 3c).

Importantly, Seq-Well uniquely leverages selective chemical functionalization to facilitate reversible attachment of a semi-permeable polycarbonate membrane (10 nm pore size) in physiologic buffers. This trait enables rapid solution exchange for efficient cell lysis but traps biological macromolecules, increasing transcript capture during hybridization and reducing cross-contamination (Supplementary Fig. 4a; Supplementary Protocol; Supplementary Video 2). The array's unique three-layer surface functionalization comprises an amino-silane base[20] crosslinked to bifunctional poly(glutamate)/chitosan top via a *p*-Phenylene diisothiocyante intermediate (**Methods**; Supplementary Fig. 4); this bifunctional top, with poly(glutamate) coating the inner surfaces of the nanowells (where cells are lysed) and chitosan the array's top surface (where the membrane binds), prevents non-specific binding of RNA to the array and efficient sealing, respectively (**Methods**; Supplementary Protocol; Supplementary Fig. 4b,c). To test sealing and buffer exchange, we monitored the fluorescence of dye-labeled, cell-bound antibodies before and after adding a guanidinium-based lysis buffer. We observed rapid diffusion of the antibodies throughout the wells within five minutes of buffer addition and, unlike unsealed or previously-described, membrane-covered BSA-blocked arrays[16], little change in fluorescent signal over 30 minutes, suggesting robust retention of biological macromolecules despite use of a strong chaotrope (**Methods**; Supplementary Fig. 5).

After lysis, cellular mRNAs are captured by bead-bound poly(dT) oligonucleotides that also contain a universal primer sequence, a cell barcode, and a unique molecular identifier (UMI) (**Methods**; Supplementary Table 1). Next, the membrane is peeled off and the beads are removed for subsequent bulk reverse transcription, amplification, library preparation and paired-end sequencing, as previously described[12] (**Methods**). Critically, beyond a disposable array and membrane, Seq-Well only requires a pipette, a manual clamp, an oven, and a tube rotator to achieve stable, barcoded single-cell cDNAs (Fig. 1a), enabling it to be performed almost anywhere.

To assess transcript capture efficiency and single-cell resolution, we profiled a mixture of 5×10^3 human (HEK293) and 5×10^3 mouse (3T3) cells using Seq-Well. The average fraction of reads mapping to exonic regions was 77.5% (Supplementary Fig. 6), demonstrating high quality libraries. Shallow sequencing from a fraction of an array revealed highly organism-specific libraries, suggesting single-cell resolution and minimal cross-contamination (Fig. 1b; Supplementary Fig. 7a–c). In the absence of membrane sealing, by comparison, we obtained poor transcript and gene detection, and substantial cross-contamination (Supplementary Fig. 1). From deeper sequencing of a fraction of a second array, we detected an average of 37,878 mRNA transcripts from 6,927 genes in HEK cells and 33,586 mRNA transcripts from 6,113 genes in 3T3 cells, comparable to a droplet-based approach using the same mRNA capture beads (Drop-Seq)[12] (Fig. 1c,d & Supplementary Fig. 7&8). Upon matched-read downsampling, we also observed levels of transcript and gene detection consistent with other massively-parallel bead-based scRNA-Seq methods (**Methods**; Supplementary Fig. 7d–g). Moreover, we found strong correlations between bulk RNA-seq data and populations constructed *in silico* from individual HEK cells (R=0.751±0.073–

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

0.983±0.0001 for populations of 1–1,000 single cells, respectively), suggesting representative cell and transcript sampling **(Methods**; Supplementary Fig. 9).

Next, to examine the ability of Seq-Well to resolve populations of cells in complex primary samples, we loaded human peripheral blood mononuclear cells (PBMCs) into arrays in triplicate prior to beads**,** allowing us to perform on-array multi-color imaging cytometry (**Methods;** Fig. 2a,b, Supplementary Tables 2&3). Sequencing one-third of the beads recovered from each array yielded 3,694 high-quality single-cell libraries (**Methods**). Unsupervised graph-based clustering revealed unique subpopulations corresponding to major PBMC cell types (**Methods;** Fig. 2b, Supplementary Fig. 10–12; Supplementary Table 4). Each array yielded similar subpopulation frequencies (Fig. 2c), with detection efficiencies comparable to other massively-parallel technologies (Supplementary Fig. 13). The proportion of each subpopulation determined by sequencing also matched on-array immunophenotyping results (Fig. 2a,b). Critically, sequencing provides additional information: in addition to resolving dendritic cells from monocytes (Fig. 2b), we found significant variation among the monocytes (captured in PC3) due to differential expression of inflammatory and anti-viral gene programs (Fig. 2d)[1,3]. Overall, characterizing a sample in two ways using a single platform increases the amount of the information that can be extracted from a precious specimen, while also allowing analysis of one measurement in light of the other.

Finally, to test the portability of Seq-Well, we profiled primary human macrophages exposed to *M. tuberculosis* (H37Rv) in a BSL3 facility **(Methods)**. In total, we recovered 14,218 (of 40,000 possible) macrophages with greater than 1,000 mapped transcripts from a TB-exposed and an unexposed array. Unsupervised analysis of 4,638 cells with greater than 5,000 transcripts per cell revealed five distinct clusters (Fig. 3a,b & Supplementary Fig. 14a,b; Supplementary Table 5). Two had lower transcript capture and high mitochondrial gene expression (suggestive of low quality libraries)[17], and were removed; the remaining three (2,560 cells) were identified in both the exposed and unexposed samples (Fig. 3a, Supplementary Fig. 14c,d&15), and likely represent distinct sub-phenotypes present in the initial culture.

We next examined common and cluster-specific gene enrichments (**Methods).** Although Clusters 1 and 3 did not present strong stimulation-independent enrichments, Cluster 2 uniquely expressed several genes associated with metabolism (Supplementary Tables 6&7). Intriguingly, within each cluster, we observed pronounced shifts in gene expression in response to *M. tuberculosis* (**Methods;** Fig. 3c & Supplementary Table 8), with common enrichments for gene sets previously observed in response to intracellular infection, LPS stimulation, and activation of TLR7/8 (Supplementary Tables 9&10). Cluster 1 uniquely displayed stimulation-induced shifts in several genes associated with cell growth, Cluster 3 in transcripts associated with hypoxia, and Cluster 2, again, in genes linked to metabolism. Overall, these data suggest that basal cellular heterogeneity may influence ensemble tuberculosis responses. Equally importantly, they demonstrate the ability of Seq-Well to acquire large numbers of single-cell transcriptomes in challenging experimental environments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In conclusion, Seq-Well is a robust platform for scalable, single-cell transcriptomics applicable to almost any cellular suspension for which a reference genome or transcriptome exists. The technique is inexpensive, user-friendly, portable, and efficient, enabling scRNA-Seq to accelerate scientific and clinical discovery, even when working with limited samples. Furthermore, the ability to measure protein secretion and cell surface expression on the same platform[18,19] foreshadows multi-omic single-cell measurements at scale. As such, our platform may prove to be a potent tool for empowering a new era of precision science and medicine.

METHODS

Methods and any associated references are available in the online version of the paper.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank K. Shekar, T. Tickle, and M. Xie for fruitful discussions.

This work was supported by the Searle Scholars Program (AKS), the Beckman Young Investigator Program (AKS), a NIH New Innovator Award DP2 OD020839 (AKS), U24 AI11862-01 (AKS), P50 HG006193 (AKS), the Bill and Melinda Gates Foundation grant 03629000189 (AKS, JCL, SF), the Ragon Institute (AKS, SF), the Burroughs Wellcome Foundation (SF), P30 AI060354 (SF), DP3 DK09768101 (JCL), P01 AI045757 (JCL), R21 AI106025 (JCL), R56 AI104274 (JCL), the W.M. Keck Foundation (JCL), and the U. S. Army Research Office through the Institute for Soldier Nanotechnologies, under contract number W911NF-13-D-0001 (JCL). This work was also supported in part by the Koch Institute Support (core) Grant P30-CA14051 from the National Cancer Institute. JCL is a Camille Dreyfus Teacher-Scholar.

References

1. Shalek AK, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature. 2013; 498:236–240. [PubMed: 23685454]
2. Lohr JG, et al. Whole exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nature biotechnology. 2014; 32:479.
3. Shalek AK, et al. Single cell RNA Seq reveals dynamic paracrine control of cellular variation. Nature. 2014; 510:363. [PubMed: 24919153]
4. Tirosh I, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016; 352:189–196. [PubMed: 27124452]
5. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell reports. 2012; 2:666–673. [PubMed: 22939981]
6. Bendall SC, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science. 2011; 332:687–696. [PubMed: 21551058]
7. Buenrostro JD, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature. 2015; 523:486–490. [PubMed: 26083756]
8. Smallwood SA, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nature methods. 2014; 11:817–820. [PubMed: 25042786]
9. Zeisel A, et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015; 347:1138–1142. [PubMed: 25700174]
10. Treutlein B, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature. 2014; 509:371–375. [PubMed: 24739965]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

11. Klein AM, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015; 161:1187–1201. [PubMed: 26000487]
12. Macosko EZ, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015; 161:1202–1214. [PubMed: 26000488]
13. Fan HC, Fu GK, Fodor SP. Combinatorial labeling of single cells for gene expression cytometry. Science. 2015; 347:1258367. [PubMed: 25657253]
14. Bose S, et al. Scalable microfluidics for single-cell RNA printing and sequencing. Genome biology. 2015; 16:1. [PubMed: 25583448]
15. Yuan J, Sims PA. An Automated Microwell Platform for Large-Scale Single Cell RNA-Seq. Scientific Reports. 2016; 6
16. DeKosky BJ, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. Nature biotechnology. 2013; 31:166–169.
17. Ilicic T, et al. Classification of low quality cells from single-cell RNA-seq data. Genome biology. 2016; 17:1. [PubMed: 26753840]
18. Yamanaka YJ, et al. Single-cell analysis of the dynamics and functional outcomes of interactions between human natural killer cells and target cells. Integrative Biology. 2012; 4:1175–1184. [PubMed: 22945136]
19. Han Q, et al. Polyfunctional responses by human T cells result from sequential release of cytokines. Proceedings of the National Academy of Sciences. 2012; 109:1607–1612.
20. Steinberg G, Stromsborg K, Thomas L, Barker D, Zhao C. Strategies for covalent attachment of DNA to beads. Biopolymers. 2004; 73:597–605. [PubMed: 15048783]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

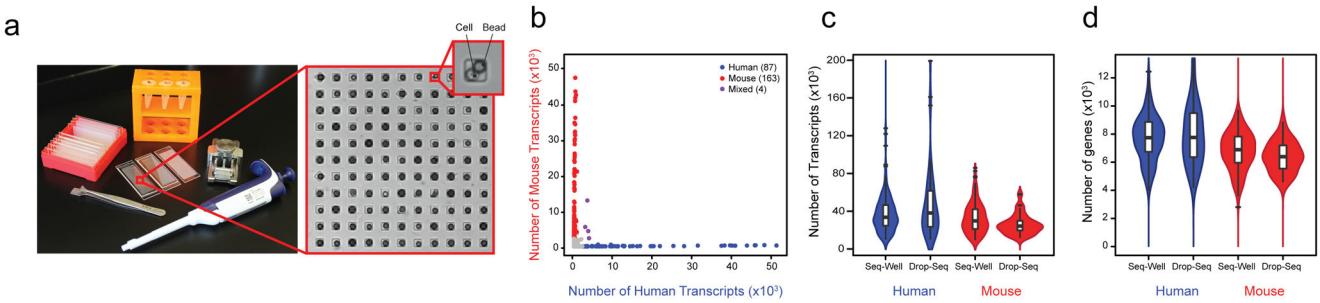

Figure 1. Seq-Well: A Portable, Low-Cost Platform for High-Throughput Single-Cell RNA-Seq of Low-Input Samples

(**a**) Photograph of equipment and array used to capture and lyse cells, respectively. (b) Transcripts captured from a mix of human (HEK293) and mouse (NIH/3T3) cells reveal distinct transcript mapping and single-cell resolution. Human (mouse) cells (> 2,000 human (mouse) transcripts and < 1,000 mouse (human) transcripts) are shown in blue (red). Among the 254 cells identified, 1.6% (shown in purple) had a mixed phenotype. (**c,d**) Violin plots of the number of transcripts (**c**) and genes (**d**) detected in human or mouse single-cell libraries generated by Seq-Well or Drop-Seq (Ref. 12; Center-line: Median; Limits: 1st and 3rd Quartile; Whiskers: +/− 1.5 IQR; Points: Values > 1.5 IQR). Using Seq-Well (Drop-Seq), an average of 37,878 (48,543) transcripts or 6,927 (7,175) genes were detected among human HEK cells (n = 159 for Seq-Well; n = 48 for Drop-Seq) and an average of 33,586 (26,700) transcripts or 6,113 (5,753) genes were detected among mouse 3T3 cells (n = 172 for Seq-Well; n = 27 for Drop-Seq) at an average read depth of 164,238 (797,915) reads per human HEK cell and an average read depth of 152,488 (345,117) read per mouse 3T3 cell.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

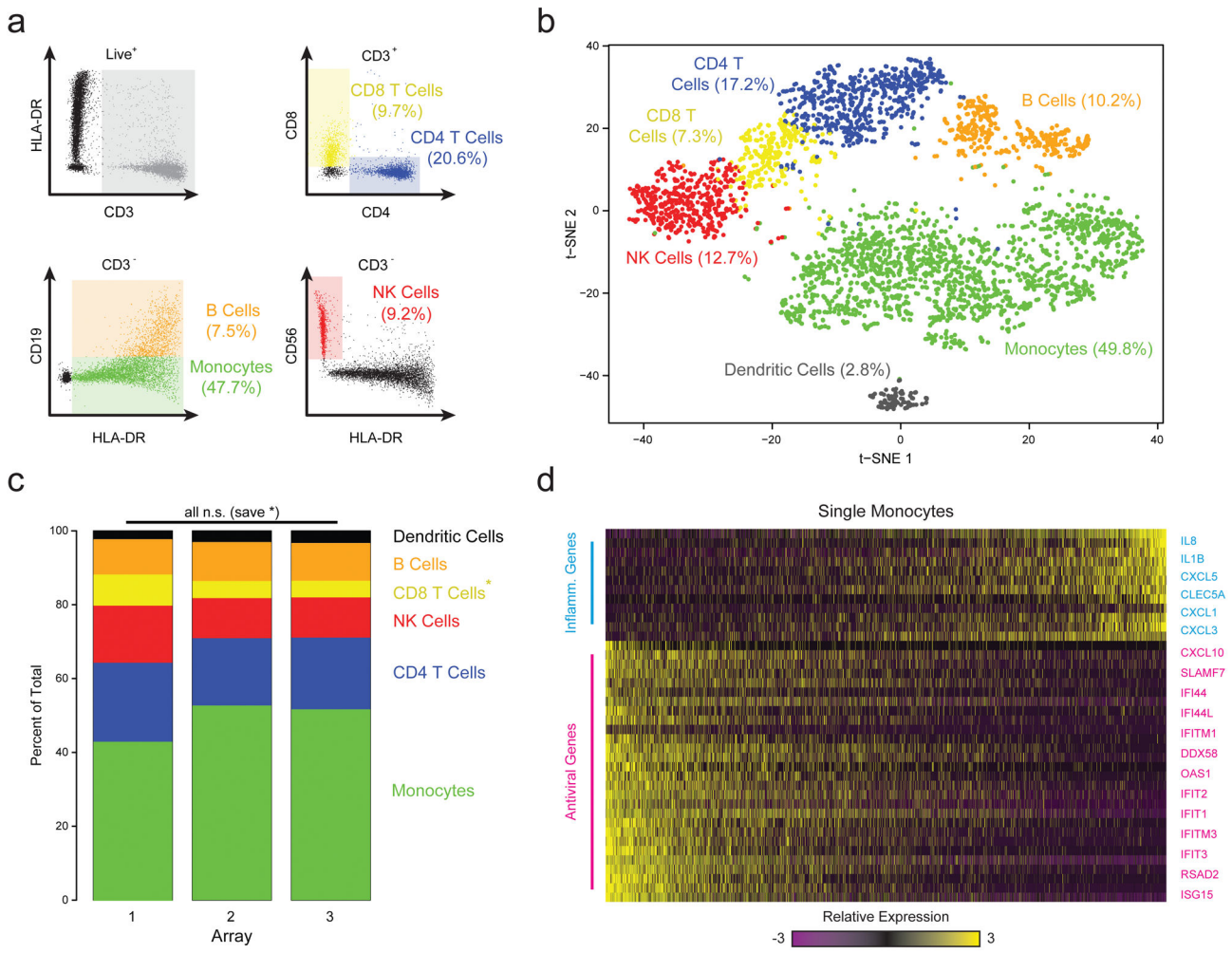

Figure 2. Combined Image Cytometry and scRNA-Seq of Human PBMCs

(**a**) The hierarchical gating scheme (with the frequencies of major cell subpopulations) used to analyze PBMCs that had been labeled with a panel of fluorescent antibodies, loaded onto three replicate arrays and imaged prior to bead loading and transcript capture (**Methods**). Myeloid cells (green) were identified as the population of hCD3(-) HLA-DR(+) CD19(-) cells; B cells (orange) as the subset of hCD3(-) HLA-DR(+) CD19 (+) cells; CD4 T cells (blue) as the subset of CD3(+) CD4(+) cells; CD8 T cells (yellow) as the CD3(+) CD8(+) subset of cells; and, NK cells (red) as the subset of CD3(-) HLA-DR (-) CD56 (+) CD16(+) cells. (**b**) t-SNE visualization of single-cell clusters identified among 3,694 human Seq-Well PBMCs single-cell transcriptomes recovered from the imaged array and the two additional ones (**Methods**; Supplementary Fig. 10–12). Clusters (subpopulations) are labeled based on annotated marker gene (Supplementary Fig. 10). (**c**) The distribution of transcriptomes captured on each of the 3 biological replicate arrays, run on separate fractions of the same set of PBMCs. All shifts are insignificant save for a slightly elevated fraction of CD8 T cells in array 1 (*, $p=1.0\times10^{-11}$; Chi-square Test, Bonferroni-corrected). (**d**) A heatmap showing the relative expression level of a set of inflammatory and antiviral genes among cells identified as monocytes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

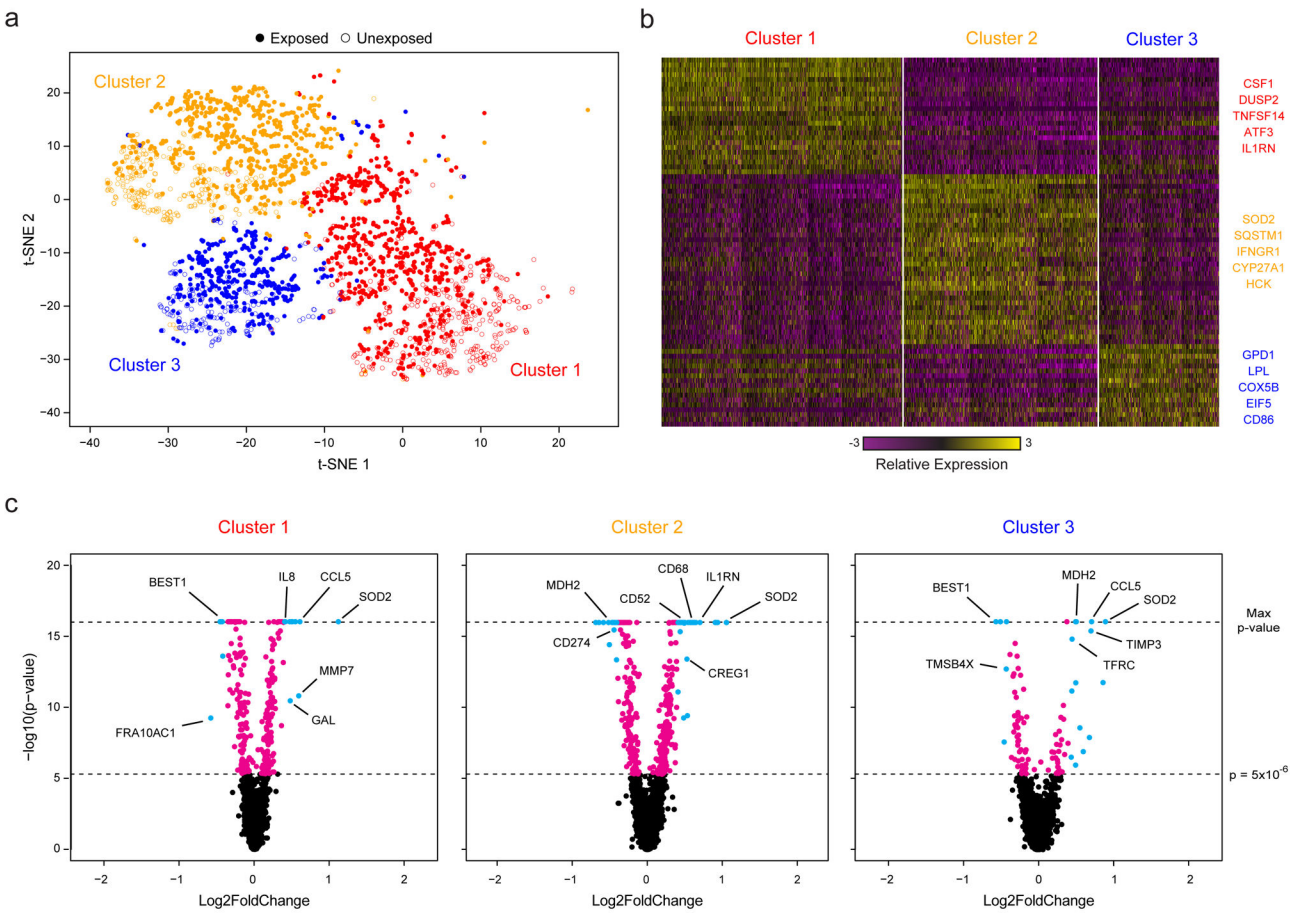

Figure 3. Sequencing of TB-Exposed Macrophages in a BSL3 Facility Using Seq-Well

(**a**) t-SNE visualization of single-cell clusters identified among 2,560 macrophages (1,686 exposed, solid circles; 874 unexposed, open circles) generated using 5 principal components across 377 variable genes (**Methods**). (**b**) Marker genes for the 3 phenotypic clusters of macrophages highlighted in (**a**). (**c**) Volcano plots of differential expression between exposed and unexposed macrophages within each cluster showing genes enriched in cells exposed to *M. tuberculosis*. In each plot, a p-value threshold of 5.0×10^{-16} based on a likelihood ratio test was used to establish statistical significance, while a log2-fold change threshold of 0.4 was used to determine differential expression. Genes with p-values less than 5.0×10^{-6} are shown in cyan and absolute log2-fold changes greater than 0.4; In magenta are genes with p-values less than 5.0×10^{-6} but absolute log2-fold changes less than 0.4; and, in black, are genes with p-values greater than 5.0×10^{-6} and absolute log2-fold changes less than 0.4.