ED SUM: A multicenter study compares 13 commonly used single-cell RNA-seq protocols.

Benchmarking single-cell RNA sequencing protocols for cell atlas projects

Elisabetta Mereu[1,+], Atefeh Lafzi[1,+], Catia Moutinho[1], Christoph Ziegenhain[2], Davis J. McCarthy[3,4,5], Adrián Álvarez-Varela[6], Eduard Batlle[6,7,8], Sagar[9], Dominic Grün[9], Julia K. Lau[10], Stéphane C. Boutet[10], Chad Sanada[11], Aik Ooi[11], Robert C. Jones[12], Kelly Kaihara[13], Chris Brampton[13], Yasha Talaga[13], Yohei Sasagawa[14], Kaori Tanaka[14], Tetsutaro Hayashi[14], Caroline Braeuning[15], Cornelius Fischer[15], Sascha Sauer[15], Timo Trefzer[16], Christian Conrad[16], Xian Adiconis[17,18], Lan T. Nguyen[17], Aviv Regev[17, 19, 20], Joshua Z. Levin[17,18], Swati Parekh[21], Aleksandar Janjic[22], Lucas E. Wange[22], Johannes W. Bagnoli[22], Wolfgang Enard[22], Marta Gut[1], Rickard Sandberg[2], Itoshi Nikaido[14,23], Ivo Gut[1,24], Oliver Stegle[3,4,25], Holger Heyn[1,24,*]

[1] CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

[2] Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden

[3] European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK

[4] European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg, Germany

[5] St Vincent's Institute of Medical Research, Fitzroy, Victoria 3065, Australia.

[6] Institute for Research in Biomedicine (IRB Barcelona). The Barcelona Institute of Science and Technology, Barcelona, Spain.

[7] ICREA, Barcelona, Spain.

[8] Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Barcelona, Spain.

[9] Max-Planck-Institute of Immunobiology and Epigenetics, Freiburg, Germany

[10] 10x Genomics, Pleasanton, CA, USA

[11] Fluidigm Corporation, South San Francisco, USA

[12] Department of Bioengineering, Stanford University, Stanford, CA, USA

[13] Bio-Rad, Hercules, CA, USA

[14] Laboratory for Bioinformatics Research RIKEN Center for Biosystems, Dynamics Research, Saitama, Japan
[15] Max Delbrück Center for Molecular Medicine / Berlin Institute of Health, Berlin, Germany
[16] Digital Health Center, Berlin Institute of Health (BIH), Charité-Universitätsmedizin Berlin, Berlin, Germany
[17] Klarman Cell Observatory, Broad Institute of MIT & Harvard, Cambridge, MA, USA
[18] Stanley Center for Psychiatric Research, Broad Institute of MIT & Harvard, Cambridge, MA, USA
[19] Koch Institute of Integrative Cancer Research, MIT, Cambridge, MA, USA
[20] Howard Hughes Medical Institute, Dept. of Biology, MIT, Cambridge, MA, USA
[21] Max-Planck-Institute for Biology of Ageing, Cologne, Germany
[22] Anthropology & Human Genomics, Department of Biology II, Ludwig-Maximilians-University, Martinsried, Germany
[23] Bioinformatics Course, Master's/Doctoral Program in Life Science, Innovation (T-LSI), School of Integrative and Global Majors (SIGMA), University of Tsukuba, Wako, Saitama, Japan
[24] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[25] Division of Computational Genomics and Systems Genetics, German Cancer Research Center (DKFZ), Heidelberg, Germany

[+] These authors contributed equally

* Correspondence should be addressed to HH (holger.heyn@cnag.crg.eu)

Single-cell RNA sequencing (scRNA-seq) is the leading technique for characterizing the transcriptome of individual cells in a sample. The latest protocols are scalable to thousands of cells and are being used to compile cell atlases of tissues, organs and organisms. However, the protocols differ substantially with respect to their RNA capture efficiency, bias, scale and costs, and their relative advantages for different applications are unclear. Here, we generated benchmark datasets to systematically evaluate protocols in terms of their power to comprehensively describe cell types and states. We performed a multi-center study comparing 13 commonly used single-cell and single-nucleus RNA-seq protocols applied to a heterogeneous reference sample resource. Comparative analysis revealed marked differences in protocol performance. The protocols differed in library complexity and their ability to detect cell type marker, impacting on their predictive value and suitability for integration into reference cell atlases. These results provide guidance both for individual researchers and for consortia projects such as the Human Cell Atlas.

Single-cell genomics provides an unprecedented view of the cellular makeup of complex and dynamic systems. Single-cell transcriptomics approaches in particular have led the technological advances that allow unbiased charting of cell phenotypes[1]. The latest improvements in scRNA-seq allow these technologies to scale to thousands of cells per experiment, providing comprehensive profiling of tissue composition[2,3]. This has led to the identification of novel cell types[4–6] and the fine-grained description of cell plasticity in dynamic systems, such as development[7,8]. Recent large-scale efforts, such as the Human Cell Atlas (HCA) project[9], are attempting to produce cellular maps of entire cell lineages, organs, and organisms[10,11] by conducting phenotyping at the single-cell level. The HCA project aims to advance our understanding of tissue function and to serve as a reference for defining variation in human health and disease. In addition to methods that capture the spatial organization of tissues[12,13], the main approach it is using is scRNA-seq analysis of dissociated cells. Therefore, tissues are disaggregated and individual cells captured by cell sorting or using microfluidic systems[1]. In sequential processing steps, cells are lysed, the RNA is reverse transcribed to cDNA, amplified, and processed to sequencing-ready libraries.

Continuous technological development has improved the scale, accuracy and sensitivity of scRNA-seq methods, and now allows us to create tailored experimental designs by selecting from a plethora of different scRNA-seq protocols. However, there are marked differences between these methods, and it is not clear which protocols are best for different applications. For large-scale consortium projects, experience has shown that neglecting benchmarking, standardization and quality control at the beginning can lead to major problems later on in the analysis of the results[14]. Thus, success depends critically on implementing a high, common standard. A comprehensive comparison of available scRNA-seq protocols will benefit both large- and small-scale applications of scRNA-seq.

The available scRNA-seq protocols vary in the efficiency of RNA molecule capture, which results in differences in sequencing library complexity and in the sensitivity of the method to identify transcripts and genes[15–17]. There has been no systematic testing of how their performance varies between cell types, and how this affects the resolution of cell phenotyping in complex samples. In this paper, we extend previous efforts to compare the molecule capture efficiency of scRNA-seq protocols[15,16] by systematically evaluating the capability of these techniques to describe tissue complexity and their suitability for creating a cell atlas. We performed a multi-center benchmarking study to compare scRNA-seq protocols using a unified reference sample resource. Our reference sample contained: i) a high degree of cell type heterogeneity with various frequencies, ii) closely related subpopulations with subtle differences in gene expression, iii) a defined cell composition with trackable markers, and iv) cells from different species. By analyzing human peripheral blood and mouse colon tissue, we have covered a broad range of cell types and states from cells in suspension and solid tissues, in order to represent common scenarios in cell atlas projects. We have also added spike-in cell lines to allow us to assess batch effects, and have combined different species to pool samples into a single reference. We performed a comprehensive comparative analysis of 13 different scRNA-seq protocols, representing the most commonly used methods. We applied a wide range of different quality control metrics to evaluate datasets from different perspectives, and to test their suitability for producing a reproducible, integrative and predictive reference cell atlas.

We observed striking differences between protocols in converting RNA molecules into sequencing libraries. Varying library complexities affected the protocol's power to quantify gene expression levels and to identify cell type markers, a trend consistently observed across cell and tissue types. This critically impacted on the resolution of tissue profiles and the predictive value of the datasets. Protocols further differed in their capacity to be integrated into reference tissue atlases and, thus, their suitability for consortia-driven projects with flexible production designs.

Results

Reference sample and experimental design.

We benchmarked current scRNA-seq protocols to inform the methodology selection process of cell atlas projects. Ideally, methods should a) be accurate and free of technical biases, b) be applicable across distinct cell properties, c) fully disclose tissue heterogeneity, including subtle differences in cell states, d) produce reproducible expression profiles, e) comprehensively detect population markers, f) be integratable with other methods, and g) have predictive value with cells mapping confidently to a reference atlas.

For a systematic comparison of protocols, we designed a reference sample containing human peripheral blood mononuclear cells (PBMC) and mouse colon, which are tissue types with highly heterogeneous cell populations, as determined by previous single-cell sequencing studies[18,19]. In addition to the well-defined cell types, the tissues contain cells in transition states (e.g. colon transit amplifying or enterocyte progenitor cells) that show transcriptional differences during their differentiation trajectory[20]. The reference sample also included a wide range of cell sizes (e.g. B-cells: ~7 µm; HEK293: ~15 µm) and RNA content, which are key parameters that affect performance in cell capture and library preparation. Interrogating tissues from different species allowed us to pool a large variety of cell types in a single reference sample in order to maximize complexity while minimizing variability introduced during sample preparation. In addition to the intra-tissue complexity, the fluorescence-labeled spiked-in cell lines allowed us to monitor cell type composition during sample processing, and to identify batch effects and biases introduced during cell capture and library preparation.

Specifically, the reference sample contained (estimated % viable cells): PBMC (60%, human), colon (30%, mouse), HEK293T (6%, RFP labelled human cell line), NIH3T3 (3%, GFP labelled mouse cells) and MDCK (1%, TurboFP650 labelled dog cells) (**Figure 1**). To reduce variability due to technical effects during library preparation, the reference sample was prepared in a single batch, distributed into aliquots of 250,000 cells, and cryopreserved. We have previously shown that cryopreservation is suitable for single-cell transcriptomics studies of these tissue types[21]. For cell capture and library preparation, the thawed samples underwent FACS separation to remove damaged cells and physical doublets (see below for detailed analysis of cell viability sorting).

A reference dataset for benchmarking experimental and computational protocols.

To obtain sufficient sensitivity to capture low-frequency cell types and subtle differences in cell state, we profiled ~3,000 cells with each scRNA-seq protocol. In total, we produced datasets for 5 microtiter plate-based methods and 7 microfluidics systems, including cell-capture technologies based on droplets (4), nanowells (1) and integrated fluidic circuits (IFC), to capture small (1) and medium (1) sized cells (**Figure 1** and **Supplementary Table 1**). We also included experiments to produce single-nucleus RNA sequencing (snRNA-seq) libraries (1), and an experimental variant that profiled >50,000 cells to produce a reference of our complex sample. The unified sample resource and standardized sample preparation (**Online Methods**) were designed to largely eliminate sampling effects, and allow the systematic comparison of scRNA-seq protocol performance.

To compare the different protocols, and to create a resource for the benchmarking and development of computational tools (e.g. batch effect correction, data integration and annotation), all datasets were processed in a uniform manner. Therefore, we designed a streamlined primary data processing pipeline tailored to the peculiarities of the reference sample (**Online Methods**). Briefly, raw sequencing reads were mapped to a joint human, mouse and canine reference genome and separately to their respective references to produce gene count matrices for subsequent analysis (GSE133549). Overall, we detected human, mouse and canine cell numbers consistent with the composition design of the reference sample (**Figure 1**). However, some protocols varied markedly from the expected frequencies in human (34-95%), mouse (4-66%) and canine (0-9%) cells. Although the reference sample was prepared in a standardized way, we cannot entirely exclude the introduction of composition variability during sample handling. Thus, the subsequent evaluation of protocol performance was performed on cell types and states common to all protocols.

Notably, we observed a higher fraction of mouse colon cells in unsorted (Chromium) and the single-nucleus sequencing dataset (Chromium (sn)). This likely results from damaging the more fragile colon cells during sample preparation, resulting in proportionally fewer colon cells when selecting for cell viability. To test if this composition bias in scRNA-seq can be avoided by skipping viability selection, we generated matched datasets selecting or not for intact cells. After quality control the detection of mouse colon cells increased proportionally without viability selection (51% vs 19%), with good-quality cells showing comparable library complexity in both libraries (e.g. numbers of detected genes; **Supplementary Figs. 1** and **2**). However, considerably more cells were removed during quality filtering (44% vs 15%), and this is a source of unwanted sequencing costs that must be taken into account, especially for tissues with high cell damage. Consequently, replacing viability staining with thorough *in silico* quality filtering in cell atlas experiments might better conserve the composition of the original tissue, but result in higher sequencing costs.

The canine cells, spiked-in at a low concentration, were detected by all protocols (1-9%) except gmcSCRB-seq. Furthermore, the different methods showed notable differences in mapping statistics between different genomic locations (**Figure 1**). As expected, due to the presence of unprocessed RNA in the nucleus, the snRNA-seq experiment detected the highest proportion of introns, although scRNA-seq protocols also showed high frequencies of intronic and intergenic mappings. The increased detection of unprocessed transcripts in CEL-seq2 may be due to a freezing step (-80ºC) after cell isolation and subsequent denaturation at high temperatures (95ºC), which could favor the accessibility of nuclear and chromatin-bound RNA molecules.

Molecule capture efficiency and library complexity

We produced reference datasets by analyzing 30,807 human and 19,749 mouse cells (Chromium V2; **Figure 2a-c**). The higher cell number allowed us to annotate the major cell types in our reference sample, and to extract population-specific markers (**Supplementary Table 2**). Noteworthy, the reference samples solely provided the basis to assign cell identities and gene marker sets and was not utilized to quantify the methods’ performance. This strategy ensured that the choice of technology for deriving the reference does not influence downstream analyses. Cell clustering and reference-based cell annotation showed high agreement (average 83%; **Supplementary Table 3**), and only cells with consistent annotations were used subsequently for comparative analysis at cell type level. The PBMCs (human) and colon cells (mouse) represented two largely different scenarios. While the differentiated PBMCs clearly separated into subpopulations (e.g. T/B-cells, monocytes, **Figure 2b** and **Supplementary Fig. 3a, 4a-d**), colon cells were ordered as a continuum of cell states that differentiate from intestinal stem cells into the main functional units of the colon (i.e. absorptive enterocytes and secretory cells, **Figure 2c** and **Supplementary Figs. 3b, 5a-d**). Notably, the subpopulation structure of our references was largely consistent with that of published datasets for human PBMC[18] and mouse colon[22] (**Supplementary Figs. 6** and **7**). After identifying major subpopulations and their respective markers in our reference sample, we clustered the cells of each sc/snRNA-seq protocol and annotated cell types using *matchSCore2* (**Online Methods**). This algorithm allows a gene marker-based projection of single cells (cell-by-cell) onto a reference sample and, thus, the identification of cell types in our datasets (**Supplementary Fig. 8** and **9**).

To compare the efficiency of mRNA capture between protocols we downsampled the sequencing reads per cell to a common depth and step-wise reduced fractions. Stochasticity introduced during downsampling did not affect the reproducibility of the results (**Supplementary Fig. 10**). Library complexity was determined separately for largely homogenous cell types with markedly different cell properties and function, namely human HEK293T cells, monocytes and B-cells (**Figure 2d,e**), and mouse colon secretory and transit-amplifying (TA) cells (**Supplementary Fig. 11a,b**). We observed large differences in the number of detected genes and molecules between the protocols, with consistent trends across cell types and gene quantification strategies (**Supplementary Fig. 11c,d**). Notably, some protocols, such as Smart-seq2 and Chromium V2, performed better with higher RNA quantities (HEK293T) compared to lower starting amounts (monocytes and B-cells), suggesting an input-sensitive optimum. Considering the different assay versions and application types of the Chromium system, a dedicated analysis showed increased detection of molecules and genes from nuclei to intact cells and towards the latest protocol versions (**Supplementary Fig. 12**). Consistent with the variable library complexity, the protocols presented large differences in drop-out probabilities (**Figure 2f**), with Quartz-seq2, Chromium V2 and CEL-seq2 showing consistently lower probability. Note that, despite the considerable differences between protocols, we observed a generally high technical reproducibility within the methods (**Supplementary Fig. 13**).

Technical effects and information content.

We further assessed the magnitude of technical biases, and the protocols' ability to describe cell populations. To quantify the technical variation within and across protocols, we selected highly variable genes (HVG) across all datasets, and plotted the variation in the main principle components (PC; **Figure 3a**). Using the downsampled data for HEK293T cells, monocytes and B-cells, we observed strong protocol-specific profiles, with the main source of variability being the number of genes detected per cell (**Figure 3b**). Data from snRNA-seq did not show notable outliers, indicating conserved representation of the transcriptome between the cytoplasm and nucleus. To quantify the protocol-related variance, we identified the PCs that correlated with the protocols' covariates in a linear model[23]. Indeed, the variance in the data was mainly explained by the protocols (HEK293T= 37.3%, Monocytes= 52.8% and B-cells=36.2%), a value that was reduced in HEK293T and monocytes when considering snRNA-seq as specific covariate (HEK293T= 9.7%, Monocytes= 22.2% and B-cells= 48.3%; **Online Methods**). The technical effects were also visible when using t-distributed stochastic neighbor embedding (tSNE) as non-linear dimensionality reduction method (**Supplementary Fig. 14**). By contrast, the methods largely mixed when the analysis was restricted to cell type-specific marker genes, suggesting a conserved cell identity profile across techniques (**Supplementary Fig. 15**).

Next, we quantified the similarities in information content of the protocols. Again, we used the downsampled datasets and commonly expressed genes and calculated the correlation between methods in average transcript counts across multiple cells, thus compensating for the sparsity of single-cell transcriptome data. For the three human cell types, we observed a broad spectrum of correlation between technologies, with generally lower correlation for smaller cell types (**Figure 3c**). While the transcriptome representation was generally conserved (**Figure 3a**), the snRNA-seq protocol resulted in a notable outlier when correlating the expression levels of common genes between protocols, possibly driven by decreased correlation of immature transcripts. Restricting the correlation analysis to population-specific marker genes, we observed less variation between protocols (Pearson's r, 0.5-0.7), which underlines that the expression of these markers is largely conserved between the methods (**Supplementary Fig. 16**).

To further test the suitability of protocols to describe cell types, we determined their sensitivity to detect population specific expression signatures, and found that they had remarkably variable power to detect marker genes. Specifically, population markers were detected with different accuracies (**Supplementary Fig. 17** and **18**), and the detection level varied substantially (**Figure 3d,e** and **Supplementary Table 4**). Quartz-seq2 and Smart-seq2 showed high expression levels for all cell type signatures, indicating that they have higher power for cell type identification. Since marker genes are particularly important for data interpretation (e.g. annotation), low marker detection levels could severely limit the interpretation of poorly explored tissues, or when trying to identify subtle differences between subpopulations. SnRNA-seq showed generally lower marker detection levels. However, gene markers were selected from intact cell experiments, which could lead to underestimating the performance of snRNA-seq to identify cell-type specific signatures in this analysis approach.

The protocols also detected vastly different total numbers of genes when accumulating transcript information over multiple cells, with strong positive outliers observed for the smaller cell types (**Figure 3f**). In particular, CEL-seq2 and Quartz-seq2 identified many more genes than other methods. Intriguingly, CEL-seq2 outperformed all other methods by detecting many weakly expressed genes; genes detected specifically by CEL-seq2 had significantly lower expression than the common genes detected by Quartz-seq2 ($p<2.2e-16$). The greater sensitivity to weakly expressed genes makes this protocol particularly suitable for describing cell populations in detail, an important prerequisite for creating a comprehensive cell atlas and functional interpretation.

Surprisingly, considering the increased library complexity of single-cell compared to single-nucleus RNA-seq, the latter protocol identified a similar number of genes when combining information across multiple cells and suggesting overall similar transcriptome complexity of the two compartments (**Supplementary Fig. 12**). ScRNA-seq detected additional genes enriched in biological processes such as organelle function, including many mitochondrial genes that were largely absent in the snRNA-seq datasets (**Supplementary Table 5**).

To further illustrate the power of the different protocols to chart the heterogeneity of complex samples, we clustered and plotted downsampled datasets in two-dimensional space (**Figure 4a**) and then calculated the cluster accuracy and Average Silhouette Width (ASW[24], **Figure 4b**), a commonly used measure for assessing the quality of data partitioning into communities. Consistent with the assumption that library complexity and sensitive marker detection provides greater power to describe complexity, methods that performed well for these two attributes showed better separation of subpopulations, greater ASW and cluster accuracy. This is illustrated in the monocytes, for which accurate clustering protocols separated the major subpopulations (CD14+ and FCGR3A+), while methods with low ASW did not distinguish between them. Similarly, several methods were able to distinguish between CD8+ and NK cells, while others were not.

Joint analysis across datasets

A common scenario for cell atlas projects is that data are produced at different sites using different scRNA-seq protocols. However, the final atlas is created from a combination of datasets, which requires that the technologies used are compatible. To assess how suitable it is to combine the results from our protocols into a joint analysis, we used downsampled human and mouse datasets to produce a joint quantification matrix for all techniques[25]. Importantly, single cells grouped themselves by cell type, suggesting that cell phenotypes are the main driver of heterogeneity in the joint datasets (**Figure 5a-d** and **Supplementary Fig. 19a,b** and **20**). Indeed, the combined data showed a clear separation of cell states (e.g. T-cell and enterocyte subpopulations) and rarer cell types, such as dendritic cells. However, within these populations, differences between the protocols pointed to the presence of technical effects that could not be entirely removed with downsampling to equal read depth and different merging tools (**Figure 5e-f** and **Supplementary Fig. 19c,d, 21a,b** and **22a,b**). To formally assess the capacity of the methods to be joined, we calculated the degree to which technologies mix in the merged datasets (**Figure 5g,h** and **Supplementary Figs. 21c,d** and **22c,d**). The protocols' suitability to be combined (mixability) was directly correlated with their power to discriminate between cell types (clustering accuracy). Thus, well-performing protocols result in high-resolution cellular maps and are suitable for consortium-driven projects that include different data sources. When integrating further downsampled datasets, we observed a drop in mixing ability (**Supplementary Fig. 19e**). Consequently, quality standard guidelines for consortia might define minimum coverage thresholds to ensure the subsequent option of data integration. A separate analysis of the single-nucleus and single-cell Chromium datasets resulted in well-integrated profiles, further supporting the potential to integrate cell atlases from cells and nuclei (**Supplementary Fig. 23** and **24**).

Cell atlas datasets will serve as a reference for annotating cell types and states in future experiments. Therefore, we assessed cells' ability to be projected onto our reference sample (**Figure 2b,c**). We used the population signature model defined by *matchSCore2* and evaluated the protocols based on their cell-by-cell mapping probability, which reflects the confidence of cell annotation (**Supplementary Fig. 25a-c**). Although there were some differences in the protocols' projection probabilities and a potential bias due to the selection of the reference protocol, a confident annotation was observed for most cells with inDrop and ddSEQ reporting the highest probabilities. Notably, high probability scores were also observed in further downsampled datasets (**Supplementary Fig. 25b**). This has practical consequences, as data derived from less well performing methods (from a cell atlas perspective) or from poorly sequenced experiments could be identifiable and thus suitable for specific analysis types, such as tissue composition profiling.

Discussion

Systematic benchmarking of available technologies is a crucial prerequisite for large-scale projects. Here, we evaluated scRNA-seq protocols for their power to produce a cellular map of complex tissues. Our reference sample simulated common scenarios in cell atlas projects, including differentiated cell types and dynamic cell states. We defined the strengths and weaknesses of key features that are relevant for cell atlas studies, such as comprehensiveness, integratability, and predictive value. The methods revealed a broad spectrum of performance, which should be considered when defining guidelines and standards for international consortia (**Figure 6**).

We expect that our results will guide informed decision-making processes for designing sc/snRNAseq studies. There are several features to consider when selecting protocols to produce a reproducible, integrative, and predictive reference cell atlas. At a given sequencing depth, the number and complexity of detected RNA molecules defines the power to describe cell phenotypes and infer their function. There are also additional essential features for cell atlas projects and their interpretation, such as population marker identification. Improved versions of plate-based methods, including Quartz-seq2, CEL-seq2, and Smart-seq2 generate such high-resolution transcriptome profiles. Also, microfluidic systems showed excellent performance in our comparison, particularly the Chromium system. While the scale of plate-based experiments is limited by the lower throughput of their individual processing units, microfluidic systems, especially droplet-based methods, can be easily applied to thousands of cells simultaneously. Protocol modification scales up throughput even further, and allows more cost-effective experiments[26–29]. Generally, late multiplexing methods, such as Smart-seq2, are more costly, but costs can be reduced by minituarization[30] and using non-commercial enzymes[31]. Custom droplet-based protocols have lower costs than their commercialized counterparts, but the optimized chemistry in commercial systems resulted in improved performance in this comparison. Nevertheless, existing platforms are undergoing continued development in both the private (**Supplementary Fig. 12**) and academic sector, so updated protocol versions promise to improve performance further. For consortium-driven projects, it is important to consider the integratability of data. We have shown that several protocols, including those with reduced library complexity and snRNA-seq, were readily integratable with other methods.

The use of PBMCs is ideal for multi-centre benchmarking efforts; blood cells are easy to isolate and show a high recovery rate after freezing. We also included mouse colon, a solid tissue requiring dissociation before scRNA-seq. Tissue digestion and cryopreservation of colon cells present additional challenges (e.g., increased rate of damaged cells), which we addressed by focusing on commonly detected cell types. Although we observed differences in the frequencies of cells from mouse and human, the composition of cell subtypes within tissues was conserved, reassuring the consistent capture of major cell types across all methods. Accordingly, subsequent analyses could be stratified by cell type, avoiding the need for a ground truth in sample composition. Further, viability sorting with minimal mechanic forces (low speed and wide nozzle size) was applied to remove damaged cells and to benchmark protocols with high-quality samples. This work standardized sample processing in order to limit technical variance to the library preparation steps, a crucial requisite for the multi-center benchmarking design. Nevertheless, on-site differences introduced during sample thawing or viability sorting could not entirely be excluded. However, our analysis also showed that viable cells selected by sorting or through thorough data quality control generate highly similar library complexity, suggesting potential differences in sample processing to have minor impacts on the data quality and supporting the robustness our results. Processing time presents another variable related to sample and data quality. While cells are directly sorted into their respective reaction volumes for plate-based methods, processing times can vary between microfluidic systems. However, this was considered as inherent feature of the protocols’ library preparation workflow which contributes to the overall performance.

Across sample origins and cell types, all tested features pointed to consistent protocol performance. In addition to the differences in protocol performance, it was the cells' RNA content and complexity that dominated the molecule and gene detection rates, which we have seen through the stratified analysis of vastly different cells types. As such, we expect the conclusions to be valid beyond the here tested human and mouse tissues.

Several additional steps are crucial for the success of single-cell projects, especially sample preparation. Optimizing sample procurement and tissue processing conditions is of crucial importance to avoid composition biases and gene expression artifacts[32–35] that could limit the value of a cell atlas. Therefore, dedicated studies are required to define optimal conditions for tissue and organ preparation in healthy and disease contexts.

From a technical perspective, multiple steps of a protocol are critical for generating complex sequencing libraries. All sc/snRNA-seq methods require multi-step whole-transcriptome amplification, including reverse-transcription (RT), conversion to amplifiable cDNA, and amplification[1]. Theoretically, the multiplicative reaction-efficiency of respective steps determines a method's power to detect RNA molecules, and in this sense Quartz-Seq2 was particularly efficient. We specifically tested for potential advantages of the Quartz-seq2 column- over bead-based purification, but did not detect differences in cDNA yield (**Supplementary Fig. 26**). However, we observed that bead concentration critically affected the yield of amplified cDNA. Moreover, performance was more stable for purification with columns compared to beads, which should be taken into account when implementing existing or developing novel sc/snRNA-seq methods.

A further essential step towards complex libraries is the conversion of first-strand cDNA to amplifiable cDNA. Three main strategies are used for this conversion, a) template-switching, b) RNaseH-DNA polymerase I mediated second-strand synthesis for *in vitro* transcription, and c) poly-A tagging[1]. Improvement of the three strategies led to better quantitative performance of scRNA-seq[36–39]. For Quartz-Seq2[37], improved poly-A tagging was most important to increase the amplified cDNA yield compared to Quartz-Seq[40], and probably explains the excellent result in this benchmarking exercise. However, optimization of the cDNA conversion still has potential to improve scRNA-seq methods.

Within the cDNA amplification step, increased PCR cycle numbers lead to PCR biases within the sequencing libraries. Early pooling increases the amount of cDNA molecules in the amplification step and reduces PCR bias. This especially favors early pooling methods at low sequencing depth (as performed in this work), as previously shown for bulk RNA-seq[41]. Similarly, *in vitro* transcription linearly amplifies cDNA with less biases than PCR-based methods, and partly explains the good performance of CEL-seq2. Further, early multiplexing of different cell numbers leads to different PCR cycle requirements (Quartz-Seq2: 768 cells and 10 cycles vs gmcSCRB-seq: 96 cells and 19 cycles, using the same DNA polymerase for amplification). The number of cells per amplification pool depends on the amount of amplifiable cDNA, implying that the good performance of Quartz-Seq2 was mainly due to efficient conversion of amplifiable cDNA from RNA with poly-A tagging.

It is equally important to benchmark computational pipelines for data analysis and interpretation[23,42–44]. We envision that the datasets provided by our study will serve as a valuable resource for the single-cell community to develop and evaluate novel strategies towards an informative and interpretable cell atlas. Moreover, the multi-center benchmarking framework presented here can readily be transferred to other organs where common tissue/cell types are analyzed using different scRNA-seq protocols (e.g. brain atlas projects).

Acknowledgements

This project has been made possible in part by grant number 2018-182827 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation. HH is a Miguel Servet (CP14/00229) researcher funded by the Spanish Institute of Health Carlos III (ISCIII). CM is supported by an AECC postdoctoral fellowship. This work has received funding from the European Union’s Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No. H2020-MSCA-ITN-2015-675752 (Singek), and the Ministerio de Ciencia, Innovación y Universidades (SAF2017-89109-P; AEI/FEDER, UE). S is supported by the DFG (GR4980) the Behrens-Weise-Foundation. DG and S are supported by the Max Planck Society. CZ is supported by EMBO through long-term fellowship ALTF 673-2017. The single-nucleus RNA-Seq data were generated with support from the National Institute of Allergy and Infectious Diseases grant U24AI118672, the Manton Foundation, and the Klarman Cell Observatory (AR). IN was supported by JST CREST grant number JPMJCR16G3; Japan and the Projects for Technological Development, Research Center Network for Realization of Regenerative Medicine by Japan, the Japan Agency for Medical Research and Development (AMED). AJ, LEW, JWB, and WE have been supported by funding from the DFG (EN 1093/2-1 and SFB1243 TP A14). We thank ThePaperMill for critical reading and scientific editing services and the Eukaryotic Single Cell Genomics Facility at Scilifelab (Stockholm, Sweden) for support. This publication is part of a project (BCLLATLAS) that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810287). Core funding is from the ISCIII and the Generalitat de Catalunya.

Author Contributions

HH designed the study. EM and AL performed all data analyses. CM, AAV and EB prepared the reference sample. CZ, DJM, SP and OS supported the data analysis. MG and IG provided technical and sequencing support. S, DG, JKL, SCB, CS, AO, RCJ, KK, CB, YT, YS, KT, TH, CB, CF, SS, TT, CC, XA, LTN, AR, JZL, AJ, LEW, JWB, WE, RS, IN provided sequencing-ready single-cell libraries or sequencing raw data. HH, EM and AL wrote the manuscript with contributions from the co-authors. All authors read and approved the final manuscript.

Conflicts of Interest

AR is a co-founder and equity holder of Celsius Therapeutics, and an SAB member of ThermoFisher Scientific and Syros Pharmaceuticals. AR is a co-inventor on patent applications to numerous advances in single-cell genomics, including droplet based sequencing technologies, as in PCT/US2015/0949178, and methods for expression and analysis, as in PCT/US2016/059233 and PCT/US2016/059239. KK, CB and YT are employed by Bio-Rad Laboratories. JKL and SCB are employees and shareholders at 10x Genomics, Inc.. SCB is a former employee and shareholder of Fluidigm Corporation. CS and AO are employed by Fluidigm. All other authors declare no conflicts of interest associated with this manuscript.

References

1. Lafzi, A., Moutinho, C., Picelli, S. & Heyn, H. Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Nat. Protoc.* **13**, 2742–2757 (2018).
2. Prakadan, S. M., Shalek, A. K. & Weitz, D. A. Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices. *Nat. Rev. Genet.* **18**, 345–361 (2017).
3. Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).
4. Montoro, D. T. *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* **560**, 319–324 (2018).
5. Plasschaert, L. W. *et al.* A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377–381 (2018).
6. Aizarani, N. *et al.* A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019).
7. Karaiskos, N. *et al.* The Drosophila embryo at single-cell transcriptome resolution. *Science* **358**, 194–199 (2017).
8. Wagner, D. E. *et al.* Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
9. Regev, A. *et al.* Science Forum: The Human Cell Atlas. *eLife* **6**, e27041 (2017).
10. Cao, J. *et al.* Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017).
11. Plass, M. *et al.* Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science* eaaq1723 (2018) doi:10.1126/science.aaq1723.
12. Moffitt, J. R. *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11046–11051 (2016).

13. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell *in situ* RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014).
14. Alioto, T. S. *et al.* A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing. *Nat. Commun.* **6**, 10001 (2015).
15. Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol. Cell* **65**, 631-643.e4 (2017).
16. Svensson, V. *et al.* Power analysis of single-cell RNA-sequencing experiments. *Nat. Methods* **14**, 381–387 (2017).
17. Tung, P.-Y. *et al.* Batch effects and the effective design of single-cell gene expression studies. *Sci. Rep.* **7**, 39921 (2017).
18. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
19. Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
20. Grün, D. *et al.* Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015).
21. Guillaumet-Adkins, A. *et al.* Single-cell transcriptome conservation in cryopreserved cells and tissues. *Genome Biol.* **18**, 45 (2017).
22. Consortium, T. T. M., Quake, S. R., Wyss-Coray, T. & Darmanis, S. Transcriptomic characterization of 20 organs and tissues from mouse at single cell resolution creates a Tabula Muris. *bioRxiv* 237446 (2017) doi:10.1101/237446.
23. Büttner, M., Miao, Z., Wolf, F. A., Teichmann, S. A. & Theis, F. J. A test metric for assessing single-cell RNA-seq batch correction. *Nat. Methods* **16**, 43–49 (2019).
24. Azuaje, F. A cluster validity framework for genome expression data. *Bioinforma. Oxf. Engl.* **18**, 319–320 (2002).
25. Lin, Y. *et al.* scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. *Proc. Natl. Acad. Sci.* 201820006 (2019) doi:10.1073/pnas.1820006116.

26. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).

27. Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).

28. McGinnis, C. S. *et al.* MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).

29. Gaublomme, J. T. *et al.* Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat. Commun.* **10**, 1–8 (2019).

30. Mora-Castilla, S. *et al.* Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing. *J. Lab. Autom.* **21**, 557–567 (2016).

31. Picelli, S. *et al.* Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).

32. Brink, S. C. van den *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).

33. Wohnhaas, C. T. *et al.* DMSO cryopreservation is the method of choice to preserve cells for droplet-based single-cell RNA sequencing. *Sci. Rep.* **9**, 1–14 (2019).

34. Tosti, L. *et al.* Single nucleus RNA sequencing maps acinar cell states in a human pancreas cell atlas. *bioRxiv* 733964 (2019) doi:10.1101/733964.

35. Massoni-Badosa, R. *et al.* Sampling artifacts in single-cell genomics cohort studies. *bioRxiv* 2020.01.15.897066 (2020) doi:10.1101/2020.01.15.897066.

36. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).

37. Sasagawa, Y. *et al.* Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biol.* **19**, (2018).

38. Hashimshony, T. *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77 (2016).

39. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun.* **9**, 2937 (2018).

40. Sasagawa, Y. *et al.* Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol.* **14**, 3097 (2013).
41. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. The impact of amplification on differential expression analyses by RNA-seq. *Sci. Rep.* **6**, 25533 (2016).
42. Soneson, C. & Robinson, M. D. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* **15**, 255–261 (2018).
43. Saelens, W., Cannoodt, R., Todorov, H. & Saeys, Y. A comparison of single-cell trajectory inference methods: towards more accurate and robust tools. *bioRxiv* 276907 (2018) doi:10.1101/276907.
44. Holland, C. H. *et al.* Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. *Genome Biol.* **21**, 36 (2020).
45. Kharchenko, P. V., Silberstein, L. & Scadden, D. T. Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11**, 740–742 (2014).
46. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
47. Korsunsky, I. *et al.* Fast, sensitive, and accurate integration of single cell data with Harmony. *bioRxiv* 461954 (2018) doi:10.1101/461954.

Online Methods

Ethical Statement

This study was approved by the Parc de Salut MAR Research Ethics Committee (reference number: 2017/7585/I) to Dr. Holger Heyn. We adhered to ethical and legal protection guidelines for human subjects, including informed consent.

Reference sample

Cell Lines

NIH3T3-GFP, MDCK-TurboFP650 and HEK293-RFP were cultured at 37ºC in an atmosphere of 5% (v/v) carbon dioxide in Dulbecco's Modified Eagle's Medium, supplemented with 10% (w/v) fetal bovine serum, 100 U penicillin, and 100 µg/L streptomycin (Invitrogen). On the reference sample preparation day, the culture medium was removed and the cells washed with 1X PBS. Afterwards, cells were trypsinized (trypsin 100X), pelleted at 800 x g for 5min, washed in 1X PBS, re-suspended in PBS-EDTA (2mM) and stored on ice.

Mouse Colon Tissue

The colon from eleven mice (7x*LGR5/GFP* and 4WT) was dissected and removed. For single-cell separation the colons were treated separately. The colon was sliced, opened and washed twice in cold 1X HBSS. It was then placed on a petri plate on ice and minced with razor blades until disintegration. The minced tissue was transferred to a 15 ml tube containing 5 ml of 1X HBSS and 83 µl of collagenase IV (final concentration 166 U/ml). The solution was incubated for 15 min at 37ºC (vortexed for 10 sec every 5 min). To inactivate the collagenase IV, 1 ml of FBS was added and vortexed for 10 seconds. The solution was filtered through a 70 µm nylon mesh (changed when clogged). Finally, all samples were combined, cells pelleted for 5 min at

400 g at 4ºC. The supernatant was removed, and the cells resuspended in 20 ml of 1X HBSS and stored on ice.

Isolation of peripheral blood mononuclear cells (PBMC)

Whole blood was obtained from four donors (two female, two male). The extracted blood was collected in Heparin tubes (GP supplies) and processed immediately. For each donor, PBMCs were isolated according to the manufacturer's instructions for FICOLL extraction (pluriSelect). Briefly, blood from two Heparin tubes (approximately 8 ml) was combined, diluted in 1X PBS and carefully added to a 50 ml tube containing 15 ml FICOLL. The tubes were centrifuged for 30 min at 500 g (minimum acceleration and deceleration). The interphase was carefully collected and diluted with 1X PBS + EDTA (2mM). Following a second centrifugation, the supernatant was discarded and the pellet resuspended in 2 ml of 1X PBS + EDTA (2mM) and stored on ice.

Preparation of the reference sample

Cell counting was performed using an automated cell counter (TC20™ Automated Cell Counter, Bio-Rad Laboratories). The reference sample was calculated to include human PBMC (60%), mouse colon (30%), HEK293T (6%, RFP labelled human cell line), NIH3T3 (3%, GFP labelled mouse cells) and MDCK (1%, TurboFP650 labelled dog cells). To adjust for cell integrity loss during sample processing, we measured the viability during cell counting, and accounted for an expected viability loss after cryopreservation (10% for cell lines and PBMC; 50% for colon[21]). All single cell solutions were combined in the proportions mentioned above and diluted to 250,000 viable cells per 0.5 ml. For cryopreservation, 0.5 ml of cell suspension was aliquoted into cryotubes and gently mixed with a freezing solution (final concentration 10% DMSO; 10% heat-inactivated FBS). Cells were then frozen by gradually decreasing the temperature (1ºC/min) to –80ºC (cryopreserved), and stored in liquid nitrogen. MARS-Seq and Smart-Seq2 experiments were performed to validate sample quality and composition before distributing aliquots to the partners.

Sample processing

Samples were stored at -80ºC upon arrival. Before processing, samples were de-frozen in a water bath (37ºC) with continuous agitation until the material was almost thawed. The entire volume was transferred to a 15 ml Falcon using a 1000 ul tip (wide-bored or cut tip) without mixing by pipetting. 1000 ul of pre-warmed (37ºC) Hibernate-A were added drop-wise while gently swirling the sample. Sample was rested for 1 min. An additional 2000 ul of pre-warmed (37ºC) Hibernate-A were added drop-wise while gently swirling the sample. Sample was rested for 1 min. Another 2000 ul of pre-warmed (37ºC) Hibernate-A were added drop-wise while gently swirling the sample. Sample was rested for 1 min. Then, 3000 ul of pre-warmed (37ºC) Hibernate-A were added drop-wise and the Falcon inverted 6 times. Sample was rested for 1 min. An additional 5000 ul of pre-warmed (37ºC) Hibernate-A were added drop-wise and the Falcon inverted 6 times. Sample was rested for 1 min. The sample was centrifuged at 400 x g for 5 min at 4ºC (pellet clearly visible). The supernatant was removed until 500 ul remained in the tube. The pellet was resuspended by gentle pipetting. 3500 ul of 1X PBS + 2mM EDTA were added and the sample stored on ice until processing. Before FACS isolation, cells were filtered through a nylon mesh and 3 ul of DAPI were added before gentle mixing. During FACS isolation, DAPI-positive cells were excluded to remove dead and damaged cells. Further, the exclusion of GFP-positive cells simulated the removal of a cell type from a complex sample. **Supplementary Fig. 27** shows representative FACS plots and gating strategies.

Single-cell RNA sequencing library preparation

For a detailed sample processing description see **Supplementary Notes**.

Data analysis

For Primary data preprocessing, Clustering, Sample deconvolution and annotation and

Reference datasets see **Supplementary Notes**.

MatchSCore2

To systematically assign cell identities to unannotated cells coming from different protocols, we used *matchSCore2*, a mathematical framework for classifying cell types based on reference data (https://github.com/elimereu/matchSCore2). The reference data consists of a matrix of gene expression counts in individual cells whose identity is known. The main steps of the *matchSCore2* annotation are the following:

1. *Normalization of the reference data*

   Gene expression counts are log-normalized for each cell using the natural logarithm of 1 + counts per 10,000. Genes are then scaled and centered using the ScaleData function in the Seurat package.

2. *Definition of signatures and their relative scores*

   For each of the cell types in the reference data, positive markers were computed using the Wilcoxon Rank sum test. The top 100 ranked markers in each cell type were used as the signature for that type. To each cell, we assigned a vector $x=(x_1,..,x_n)$ of signature scores, where n is the number of cell types in the reference data. The *i*-th signature score for the *k*-th cell is computed as follows:

   $\text{Score}_k = \sum_{j \ in \ J} z_{jk}$

   where *J* is the set of genes in signature *i*, and z_{jk} represents the z-score of gene *j in the k*-th cell.

3. *Training of the probabilistic model on the reference data.*

   We proposed a supervised multinomial logistic regression model, which uses the enrichment of the signature of each reference cell type in each cell to assign identity to that cell. In other words, for each cell *k* and signature *i*, we calculate the *i*-th cell type signature score X_i in the *k*-th cell as described in point 2. The distribution of the signature scores is preserved, independent of which protocol is used (**Supplementary Fig. 28** and **29**). More specifically, we defined the variables $X_{1,\ldots,}X_n$, where X_i is the vector in which the scores for signature *i* of all cells are contained. Then we used X_i as the predictors of a multinomial logistic regression.

The model assumes that the number of cells from each type in the training reference data $T_{1,}\, T_{2,\ldots,}T_n$, are random variables and that the variable $T=(T_{1,}\, T_{2,\ldots,}T_n)$ follows a multinomial distribution $M(N, \pi=(\pi_{1,\ldots,}\, \pi_n))$, where π_i is the proportion of the *i*-th cell type and N is the total number of cells.

To test the performance of the model, training and test sets were created by subsampling the reference into two datasets, maintaining the original proportions of cell types in both sets. The model was trained by using the *multinom* function from the *nnet* R package (*decay*=1e-04, *maxit*=500). To improve the convergence of the model function, X_i variables were scaled to the interval [0,1].

Cell Classification

For each cell, model predictions consisted of a set of probability values per identity class, and the highest probability was used to annotate the cell if it was >0.5; otherwise the cell resulted unclassified.

Model accuracy

To evaluate the fitted model using our reference datasets, we assessed the prediction accuracy in the test set, which was around 0.9 for human and 0.85 for mouse reference. We further assessed *matchSCore2* classifications in datasets from other sequencing methods by looking at the agreement between clusters and classification. Notably, the resulting average agreement was of

80% (range: from 58% in gmcSCRB-seq to 92% in Quartz-Seq2), while the rate of unclassified cells was less than 2%.

Downsampling

To decide on a common downsampling threshold for sequencing depth per cell, we inspected the distribution of the total number of reads per cell for each technique, and chose the lowest first quartile (fixed to 20,000 reads/cell). We then performed stepwise downsampling (25%, 50% and 75%) using the zUMIs downsampling function. We omitted cells that did not achieve the required minimum depth (**Supplementary Table 6**). Notably, stochasticity introduced during downsampling did not affect the results of this study, as exemplified by the consistent numbers of detected molecules across different downsampling iterations (**Supplementary Fig. 10**).

Estimation of dropout probabilities

We investigated the impact of dropout events in HEK293T, monocytes and B-cells extracted for each technique on downsampled data (20,000 reads/cell). For datasets with >50 cells from the selected populations, we randomly sampled 50 cells to eliminate the effect of differing cell number. The dropout probability was computed using SCDE R package[45]. SCDE models the measurements of each cell as a mixture of a negative binomial process to account for the correlation between amplification and detection of a transcript and its abundance, and a Poisson process to account for the background signal. We then used estimated individual error models for each cell as a function of expression magnitude to compute dropout probabilities using SCDE’s scde.failure.probability function. Next, we calculated the average estimated dropout probability for each cell type and technique. To integrate dropout measures into the final benchmarking score, we calculated the Area Under the Curve (AUC) of the expression prior and failure probabilities (**Figure 2f** and **Supplementary Table 7**). We expected that protocols that result in fewer dropouts would have lower AUC.

Quantification of variance introduced by batches

In order to quantify the amount of variance that is introduced by batches (protocols, processing units or experiments), we used the top 20 PCs and the standard deviation of each PC, previously calculated on HVG. Next, using pcRegression function of kBET R package[23], we regressed the batch covariate (protocols/processing units/experiments as categories defined in the kBET model) and each PC to get the coefficient of determination as an approximation of the variance explained by batches and the proportions of explained variance in each PC. We either reported the percentage of the variance that is significantly correlating with batch in first 20 PCs, or R-squared measures of the model for each PC.

Cumulative number of genes

The cumulative number of detected genes in downsampled data was calculated separately for each cell type. For cell types with >50 cells annotated, we randomly selected 50 cells and calculated the average number of detected genes per cell after 50 permutations over n sampled cells, where n is an increasing sequence of integers from 1 to 50.

GO enrichment analysis

In order to compare functional gene sets between single-cell and single-nucleus datasets, we performed GO enrichment analysis on the set of protocol-specific genes using *simpleGO* (https://github.com/iaconogi/simpleGO). For each cell type (HEK293T, monocytes, B-cells), we selected two gene sets extracted from the cumulated genes and using the maximum number of detected cells common to all three Chromium versions: i) Genes that were uniquely detected in the intersection of Chromium (V2) and (V3), but not in Chromium (sn) and ii) genes that were uniquely identified with Chromium (sn). For each of the gene sets we identified the union over cell types before applying *simpleGO*.

Correlation analysis.

Pearson correlations across protocols were computed independently for B-cells, Monocytes and HEK293T cells. For each cell type, cells were downsampled to the maximum common number of cells across all protocols. Gene counts of commonly expressed genes (from datasets downsampled to 20K reads) were averaged across cells before computing their Pearson's correlations. The *corplot* library was then used to plot resulting correlations. Protocols were ordered by agglomerative hierarchical clustering.

Silhouette scores

To measure the strength of the clusters, we calculated the Average Silhouette Width (ASW)[24]. The downsampled data (20,000 reads/cell) were clustered by Seurat[46], using graph based clustering with the first 8 principle components and resolution of 0.6. We then computed an average Silhouette width for the clusters using a Euclidean distance matrix (based on principle components 1 to 8). We report the average Silhouette width for each technique separately.

Dataset merging

Dataset integration across protocols is challenging and we applied different tools to assess the integratability of the sc/snRNA-seq methods, while conserving biological variability. To integrate datasets, we used Seurat[46], harmony[47] and scMerge[25], evaluated the results separately and averaged the integration capacity of the protocols into a joint score. We combined downsampled count matrices using the sce_cbind function in scMerge, which includes the union of genes from different batches. While both harmony and Seurat integration apply similar preprocessing steps (log-normalization, scaling and highly variable genes identification) as implemented in the Seurat tool, scMerge uses a set of genes with stable expression levels across different cell types and then creates pseudo-replicates across datasets, allowing the estimation and correction for undesired sources of variability. However, for all three alignment methods, Seurat was applied to perform clustering and UMAP after the protocol correction, in order to minimize the variability related to the downstream analysis. The clustering accuracy metric was used together with the mixability score to quantify the success of the integration. Omitting the cell integration step before visualizing the datasets together in a single tSNE/UMAP resulted in a protocol-specific distribution with cell types scattered to multiple clusters (**Supplementary Fig. 30**).

Clustering accuracy

To determine the clusterability of methods to identify cell types, we measured the probability of cells to be clustered with cells of the same type. Let C_k , $k \in \{1, \dots, N\}$ represent the cluster of cells corresponding to a unique cell type (based on the highest agreement between clusters and cell types), and T_j , $j \in \{1, \dots, S\}$ the set of different cell types, where $C \subseteq T$. For each cell type T_j, we compute the proportion p_{jk} of T_j cells that cluster in their correct cluster C_k. We define the cell-type separation accuracy as the average of these proportions.

Mixability

To account for the level of mixing of each technology, we used kBet[23] to quantify batch effects by measuring the rejection rate of a Pearson's χ^2 test for random neighborhoods. To make a fair comparison, kBet was applied to the common cell types separately by subsampling batches to the minimum number of cells in each cell type. Due to the reduced number of cells, the option heuristic was set to 'False', and the testSize was increased to ensure a minimum number of cells.

Mixability was calculated by averaging cell type specific rejection rates.

Benchmarking score

To create an overall benchmarking score with which to compare technologies, we considered six key metrics: gene detection, overall level of expression in transcriptional signatures, cluster accuracy, classification probability, cluster accuracy after integration, and mixability. Each metric was scaled to the interval [0,1], then in order to equalize the weight of each metric score, the harmonic mean across these metrics was calculated to obtain the final Benchmarking scores.

Gene detection, overall expression in cell type signatures, and classification probabilities were computed separately for B-cells, HEK293T cells and monocytes, and then aggregated by the arithmetic mean across cell types. Notably, the choice of protocol to create the reference dataset (Chromium) for initial cell annotation had no impact on the outcome of the study (**Supplementary Fig. 31**).

Data Availability

All raw sequencing data and processed gene expression files are freely available through the Gene Expression Omnibus (GEO; GSE133549).

Code availability

All code for the analysis is provided as supplementary material. All code is also available under https://github.com/ati-lz/HCA_Benchmarking and https://github.com/elimereu/matchSCore2**.**

Figure legends

Figure 1. Overview of the experimental design and data processing.

The reference sample consists of human PBMC (60%) and HEK293T (6%), mouse colon (30%) and NIH3T3 (3%) and dog MDCK (1%). The sample was prepared in one single batch, cryopreserved and sequenced by 13 different sc/snRNA-seq methods. Sequences were uniformly mapped to a joint human, mouse and canine reference and then separately to produce gene expression counts for each sequencing method.

Figure 2. Comparison of 13 sc/snRNA-seq methods.

a. Color legend of sc/snRNA-seq protocols. **b.** UMAP of 30,807 cells from the human reference sample (Chromium) colored by cell type annotation. **c.** UMAP of 19,749 cells from the mouse reference (Chromium) colored by cell type annotation. **d.** Boxplots displaying the minimum, 1st, 2nd, 3rd quantiles and the maximum number of genes detected across protocols, in downsampled (20K) HEK293T cells, monocytes and B-cells. Cell identities were defined by combining the clustering of each dataset and cell projection onto the reference. **e.** Number of detected genes at step-wise downsampled sequencing depths. Points represent the average number of detected genes as a fraction of all cells of the corresponding cell type at the corresponding sequencing depth. **f.** Dropout probabilities as a function of expression magnitude, for each protocol and cell type, calculated on downsampled data (20K) for 50 randomly selected cells.

Figure 3. Similarity measures of sc/snRNA sequencing methods.

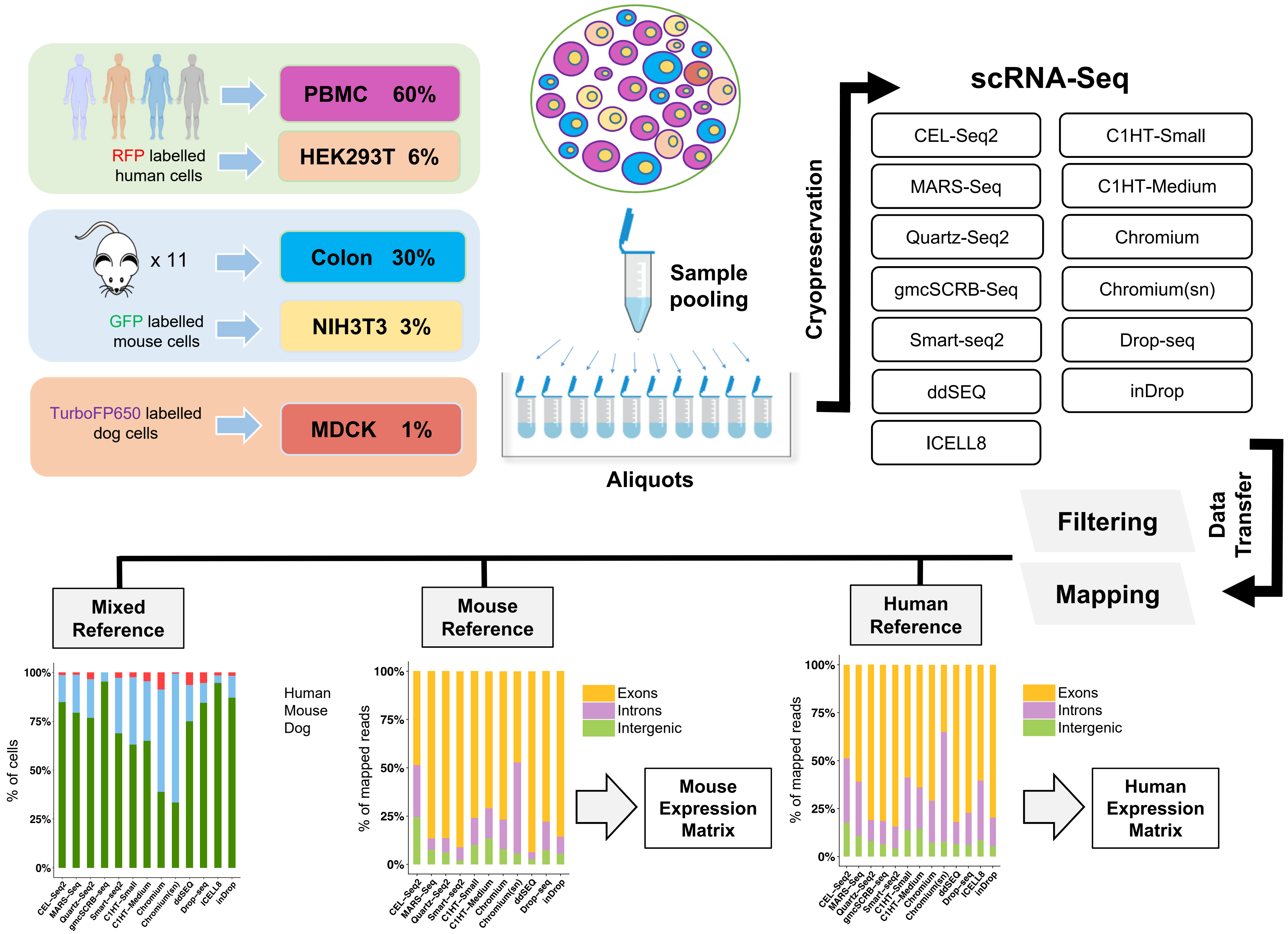
a,b. PCA analysis on downsampled data (20K) using highly variable genes between protocols, separated into HEK293T cells, monocytes and B-cells, and color-coded by protocol (**a**) and number of detected genes per cell (**b**). **c.** Pearson correlation plots across protocols using expression of common genes. For a fair comparison, cells were downsampled to the same number for each method (B cells=32, Monocytes = 57, HEK293T= 55). Protocols are ordered by agglomerative hierarchical clustering. **d.** Average log expression values of cell-type-specific reference markers for downsampled (20K) HEK293T cells, monocytes, and B-cells. **e.** Log expression values of reference markers on downsampled data (20K) for HEK293T cells, monocytes and B-cells (maximum 50 random cells per technique). **f.** Cumulative gene counts per protocol as the average of 100 randomly sampled HEK293T cells, monocytes and B-cells separately on downsampled data (20K).

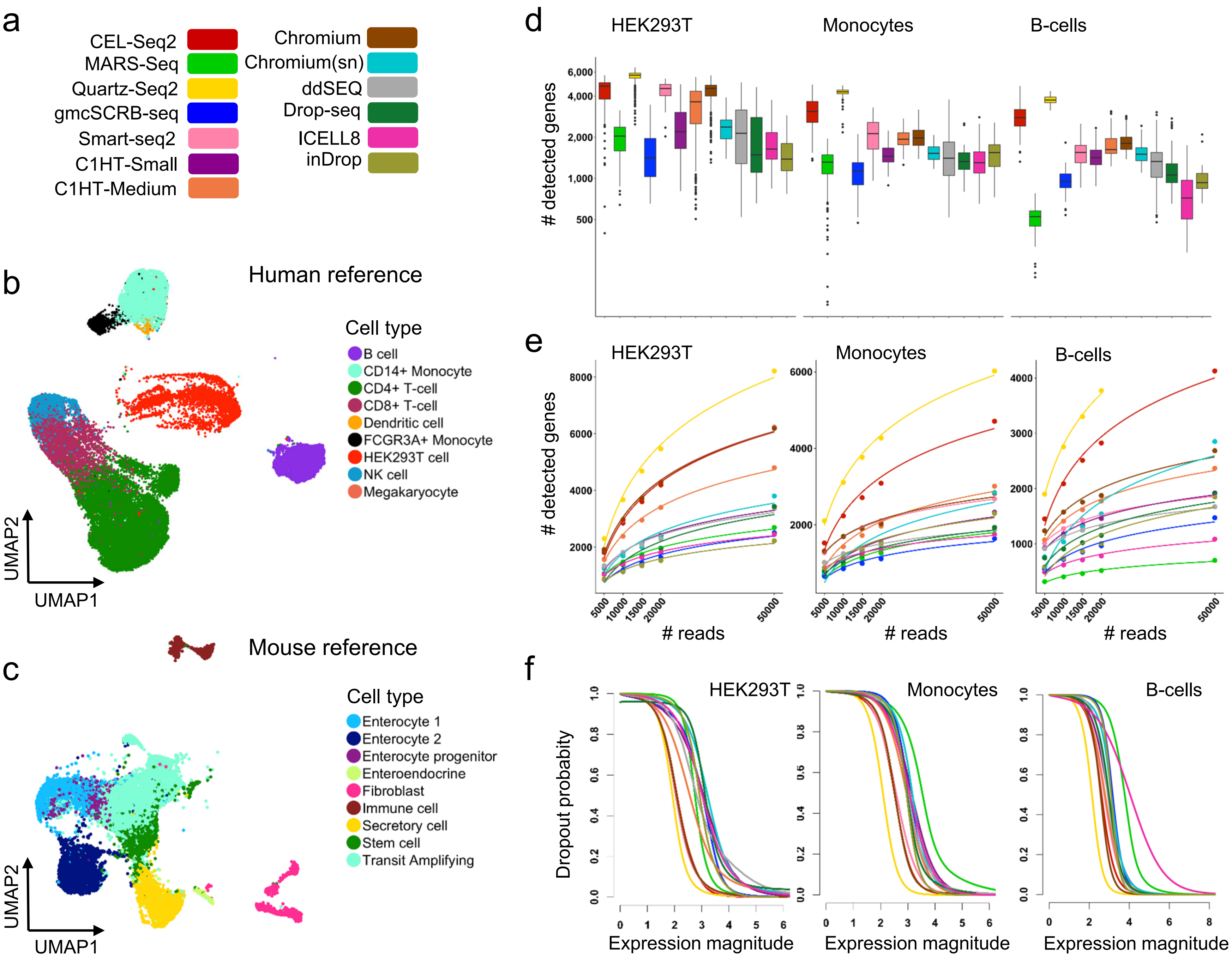
Figure 4. Clustering analysis of 13 sc/snRNA-seq methods on downsampled datasets (20K). a. T-SNE visualizations of unsupervised clustering in human samples from 13 different methods. Each dataset was analyzed separately after downsampling to 20K reads per cell. Cells are colored by cell type inferred by *matchSCore2* before downsampling. Cells that did not achieve a probability score of 0.5 for any cell type were considered unclassified. **b.** Clustering accuracy and Average Silhouette Width for clusters in each protocol.

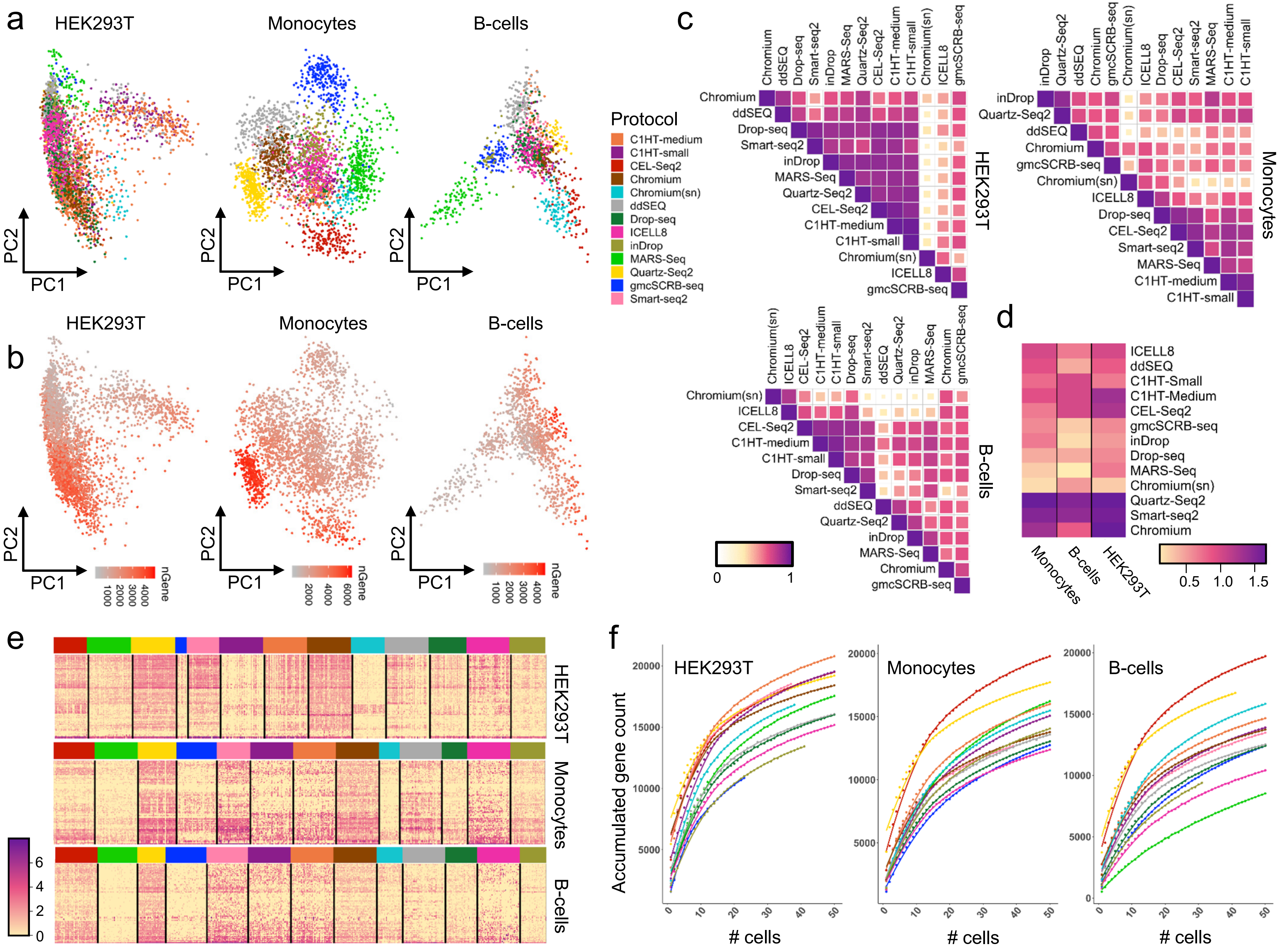
Figure 5. Integration of sc/snRNA-seq methods. UMAP visualization of cells after integrating technologies for 18,034 human (**a,b**) and 7,902 mouse (**c,d**) cells. Cells are colored by cell type (**a,c**) and sc/snRNA-seq protocol (**b,d**). **e,f.** Barplots showing normalized and method-corrected (integrated) expression scores of cell-type-specific signatures for human HEK293T cells, monocytes, B-cells (**e**) and mouse secretory and transit-amplifying cells (**f**). Bars represent cells and colors methods. **g,h.** Evaluation of method integratability in human (**g**) and mouse (**h**). Protocols are compared according to their ability to group cell types into clusters (after integration) and mix with other technologies within the same clusters. Points are colored by sequencing method.

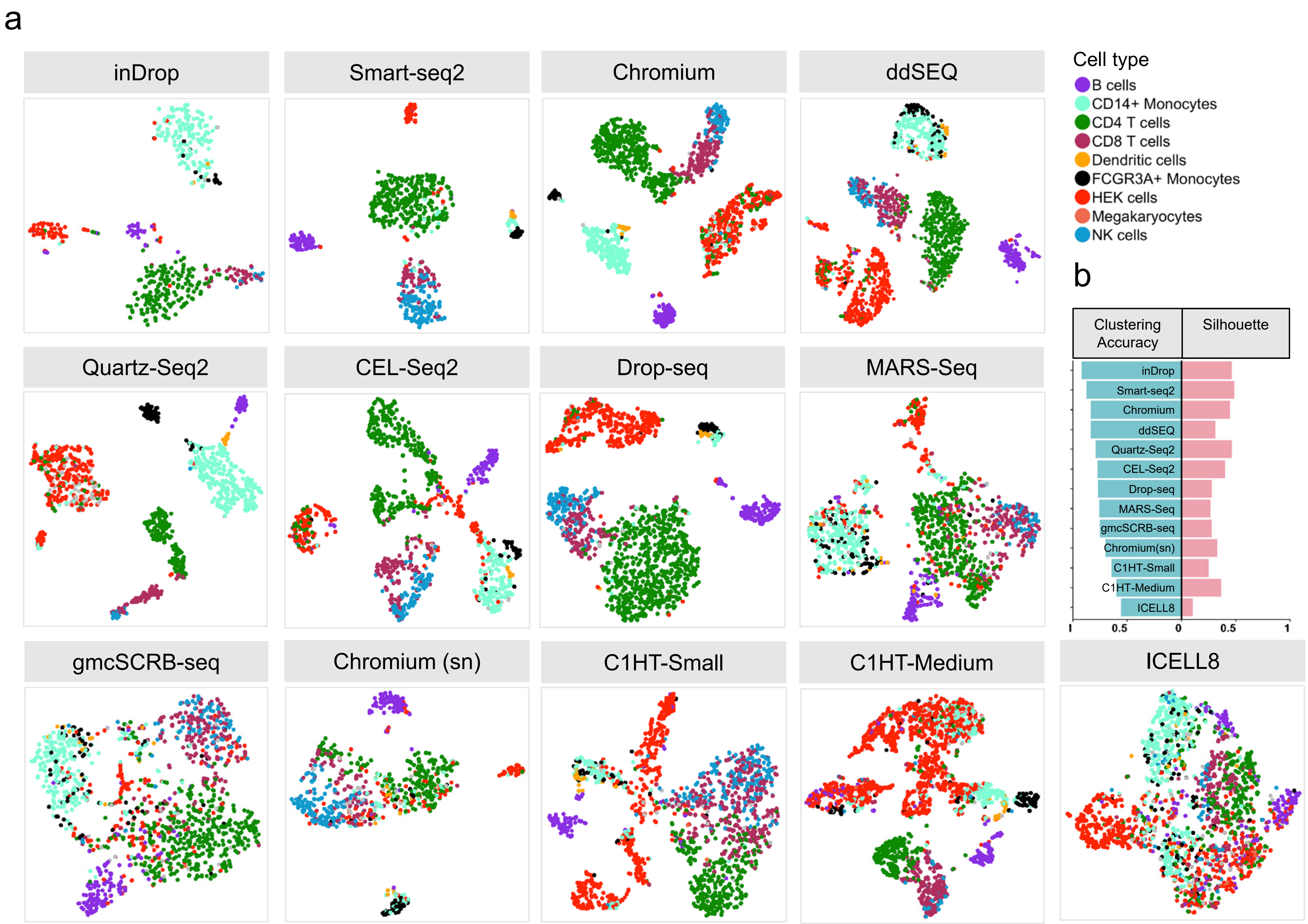
Figure 6. Benchmarking summary of 13 sc/snRNA-seq methods. Methods are scored by key analytical metrics, characterizing protocols according to their ability to recapitulate the original structure of complex tissues, and their suitability for cell atlas projects. The methods are ordered

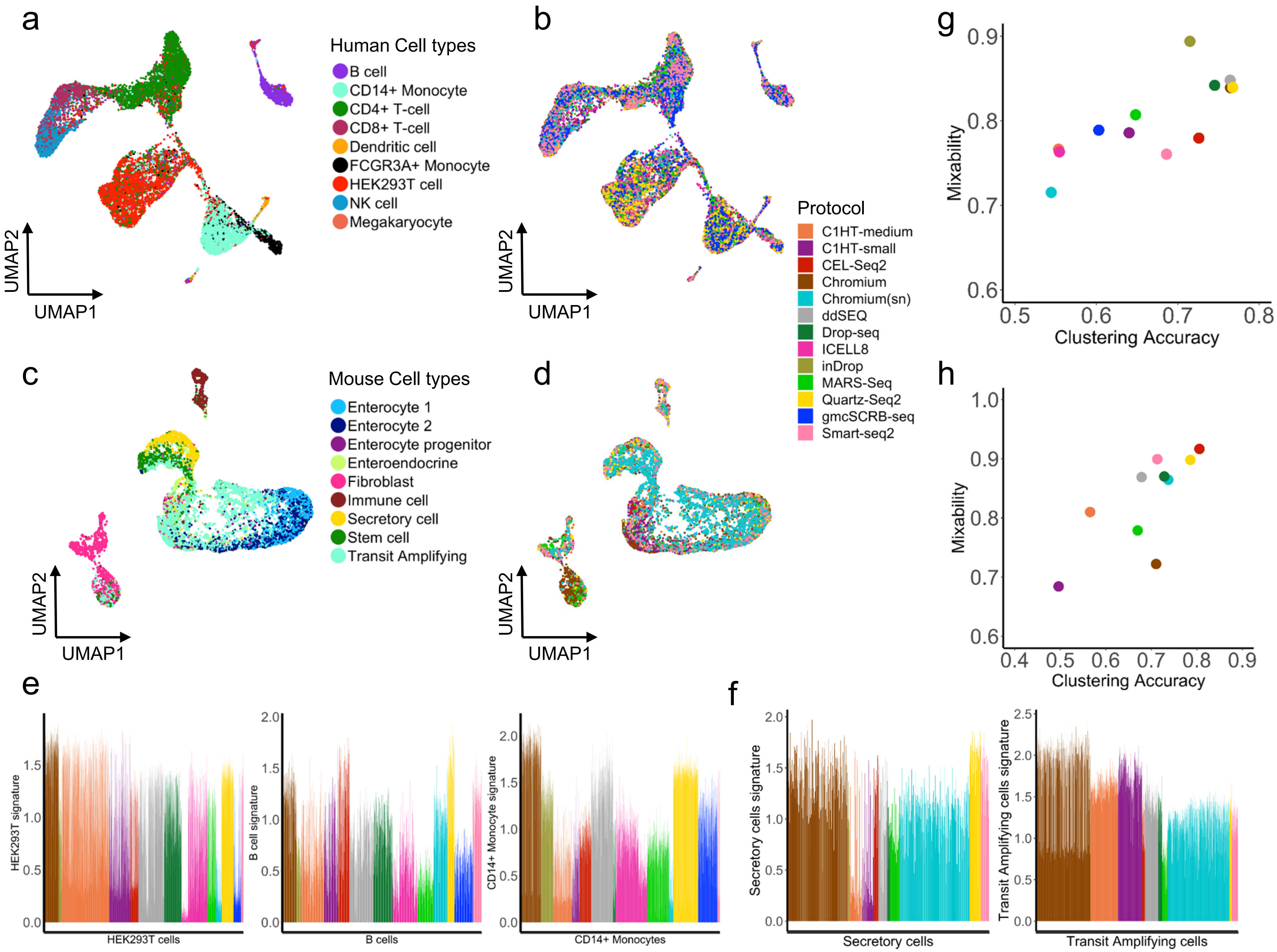
by their overall benchmarking score, which is computed by averaging the scores across metrics assessed from the human datasets.

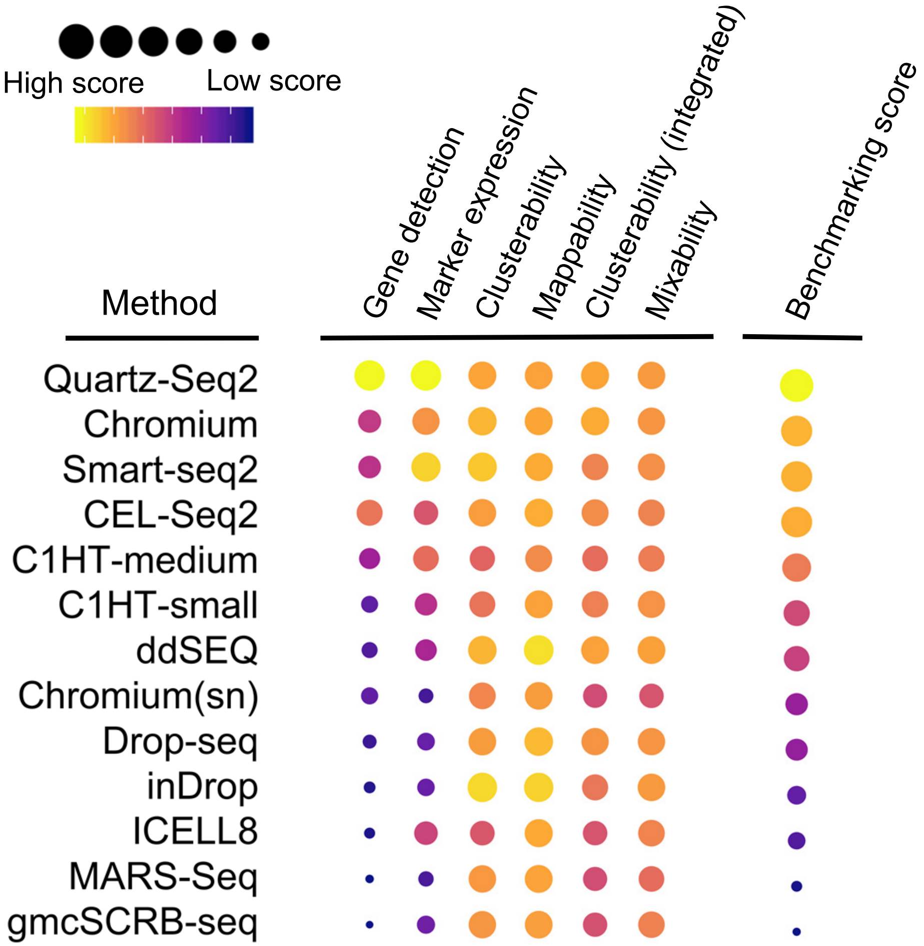

RFP labelled human cells
PBMC 60%
HEK293T 6%
x 11
GFP labelled mouse cells
Colon 30%
NIH3T3 3%
TurboFP650 labelled dog cells
MDCK 1%
Sample pooling
Aliquots
Cryopreservation
scRNA-Seq
CEL-Seq2
MARS-Seq
Quartz-Seq2
gmcSCRB-Seq
Smart-seq2
ddSEQ
ICELL8
C1HT-Small
C1HT-Medium
Chromium
Chromium(sn)
Drop-seq
inDrop
Data Transfer
Filtering
Mapping
Mixed Reference
Mouse Reference
Human Reference
% of cells
100%
75%
50%
25%
0%
Human
Mouse
Dog
% of mapped reads
Exons
Introns
Intergenic
CEL-Seq2
MARS-Seq
Quartz-Seq2
gmcSCRB-seq
Smart-seq2
C1HT-Small
C1HT-Medium
Chromium
Chromium(sn)
ddSEQ
Drop-seq
ICELL8
inDrop
Mouse Expression Matrix
Human Expression Matrix


a
CEL-Seq2
MARS-Seq
Quartz-Seq2
gmcSCRB-seq
Smart-seq2
C1HT-Small
C1HT-Medium
Chromium
Chromium(sn)
ddSEQ
Drop-seq
ICELL8
inDrop
b
Human reference
Cell type
B cell
CD14+ Monocyte
CD4+ T-cell
CD8+ T-cell
Dendritic cell
FCGR3A+ Monocyte
HEK293T cell
NK cell
Megakaryocyte
UMAP2
UMAP1
c
Mouse reference
Cell type
Enterocyte 1
Enterocyte 2
Enterocyte progenitor
Enteroendocrine
Fibroblast
Immune cell
Secretory cell
Stem cell
Transit Amplifying
UMAP2
UMAP1
d
HEK293T
Monocytes
B-cells
detected genes
6,000
4,000
2,000
1,000
500
e
HEK293T
Monocytes
B-cells
detected genes
8000
6000
4000
2000
6000
4000
2000
4000
3000
2000
1000
5000
10000
15000
20000
50000
reads
f
HEK293T
Monocytes
B-cells
Dropout probaity
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
6
0
2
4
6
8
Expression magnitude

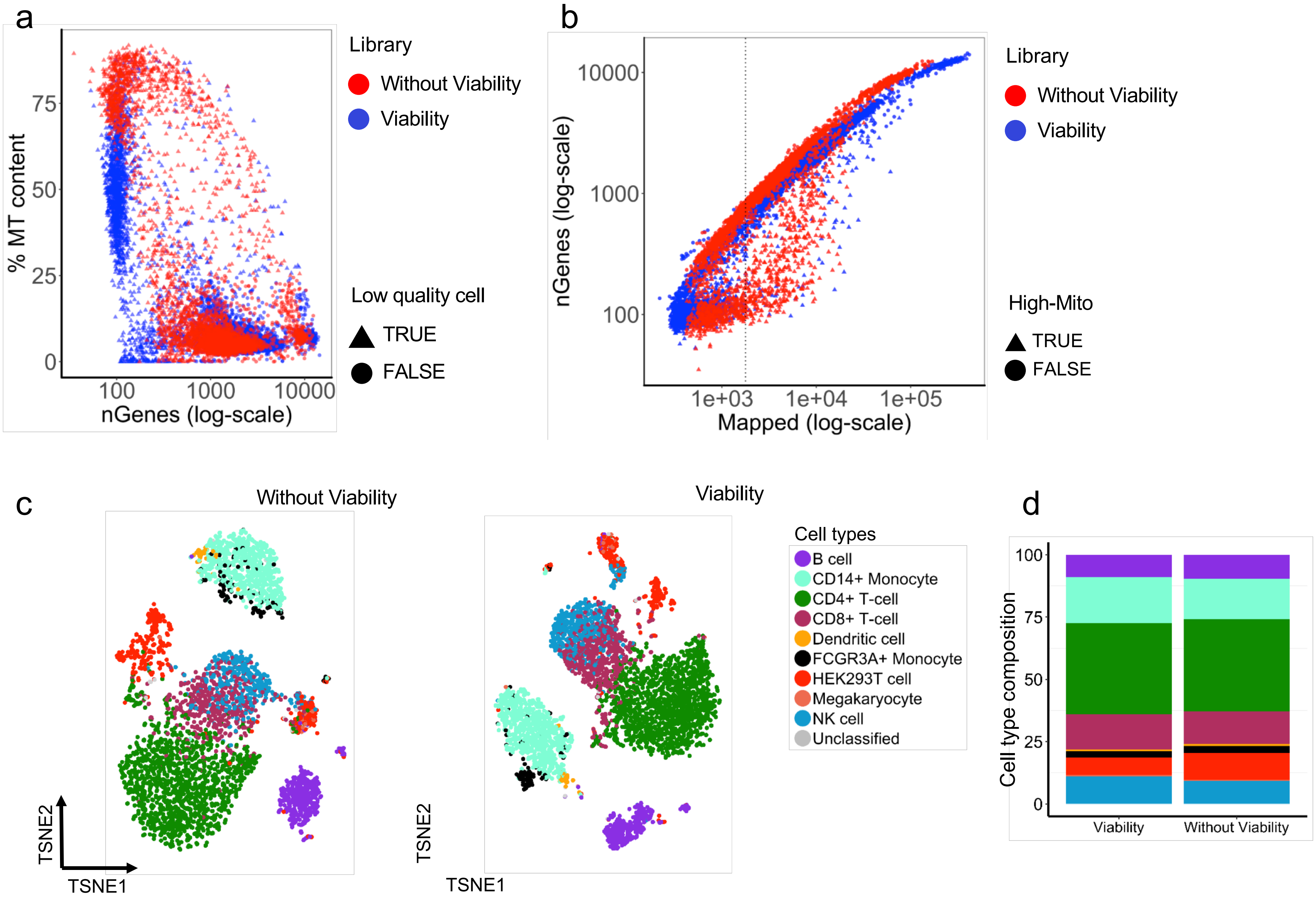

a
HEK293T
Monocytes
B-cells
PC2
PC1
Protocol
C1HT-medium
C1HT-small
CEL-Seq2
Chromium
Chromium(sn)
ddSEQ
Drop-seq
ICELL8
inDrop
MARS-Seq
Quartz-Seq2
gmcSCRB-seq
Smart-seq2
b
HEK293T
Monocytes
B-cells
PC2
PC1
nGene
1000 2000 3000 4000
2000 4000 6000
1000 2000 3000 4000
c
HEK293T
Chromium
ddSEQ
Drop-seq
Smart-seq2
inDrop
MARS-Seq
Quartz-Seq2
CEL-Seq2
C1HT-medium
C1HT-small
Chromium(sn)
ICELL8
gmcSCRB-seq
Monocytes
inDrop
Quartz-Seq2
ddSEQ
Chromium
gmcSCRB-seq
Chromium(sn)
ICELL8
Drop-seq
CEL-Seq2
Smart-seq2
MARS-Seq
C1HT-medium
C1HT-small
B-cells
Chromium(sn)
ICELL8
CEL-Seq2
C1HT-medium
C1HT-small
Drop-seq
Smart-seq2
ddSEQ
Quartz-Seq2
inDrop
MARS-Seq
Chromium
gmcSCRB-seq
0
1
d
ICELL8
ddSEQ
C1HT-Small
C1HT-Medium
CEL-Seq2
gmcSCRB-seq
inDrop
Drop-seq
MARS-Seq
Chromium(sn)
Quartz-Seq2
Smart-seq2
Chromium
Monocytes
B-cells
HEK293T
0.5 1.0 1.5
e
HEK293T
Monocytes
B-cells
0 2 4 6
f
Accumulated gene count
HEK293T
Monocytes
B-cells
5000 10000 15000 20000
0 5000 10000 15000 20000
0 10 20 30 40 50
cells


a
inDrop
Smart-seq2
Chromium
ddSEQ
Quartz-Seq2
CEL-Seq2
Drop-seq
MARS-Seq
gmcSCRB-seq
Chromium (sn)
C1HT-Small
C1HT-Medium
ICELL8
Cell type
B cells
CD14+ Monocytes
CD4 T cells
CD8 T cells
Dendritic cells
FCGR3A+ Monocytes
HEK cells
Megakaryocytes
NK cells
b
Clustering Accuracy
Silhouette
inDrop
Smart-seq2
Chromium
ddSEQ
Quartz-Seq2
CEL-Seq2
Drop-seq
MARS-Seq
gmcSCRB-seq
Chromium(sn)
C1HT-Small
C1HT-Medium
ICELL8
1
0.5
0
0.5
1

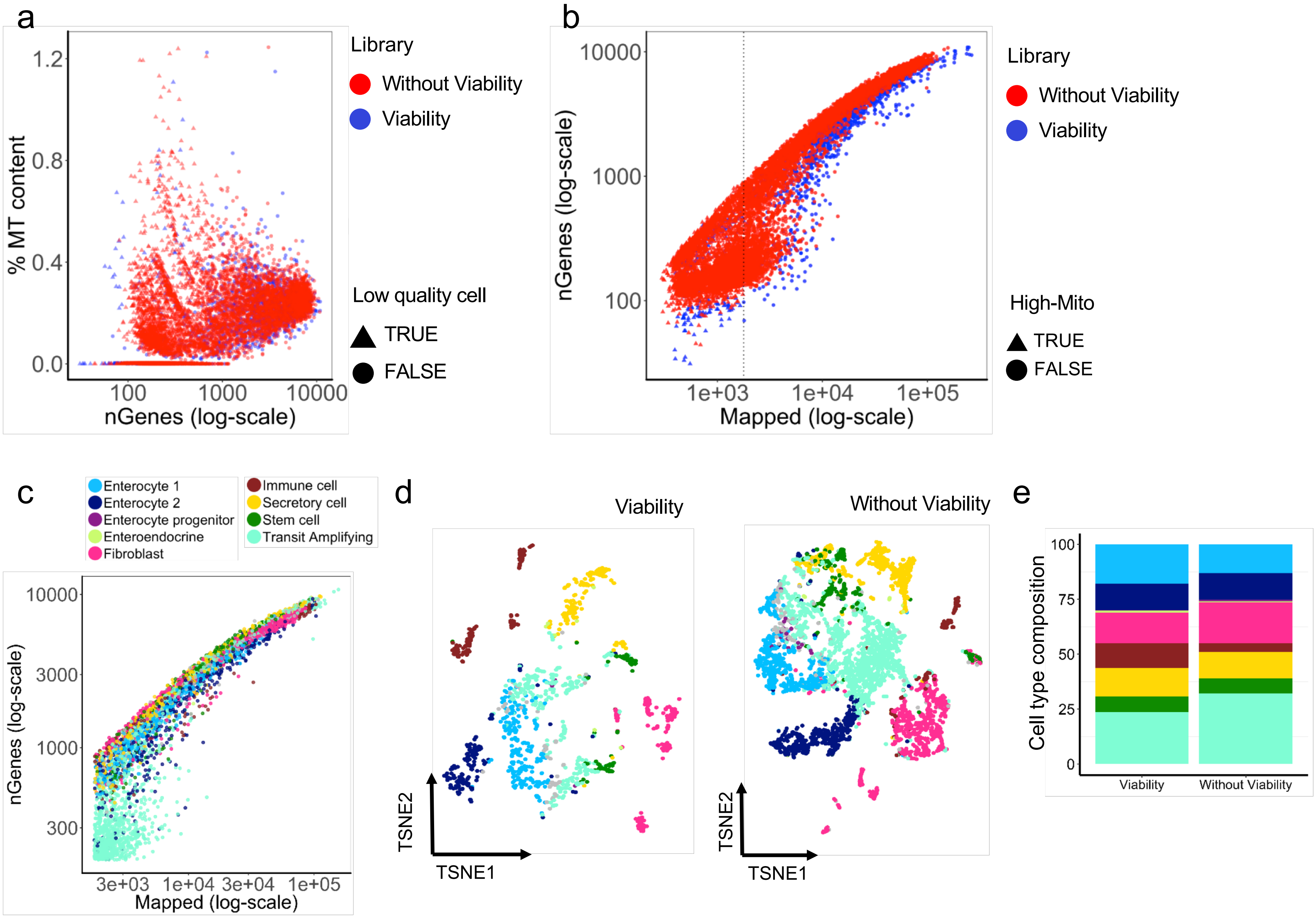

a
Human Cell types
B cell
CD14+ Monocyte
CD4+ T-cell
CD8+ T-cell
Dendritic cell
FCGR3A+ Monocyte
HEK293T cell
NK cell
Megakaryocyte
UMAP2
UMAP1
b
UMAP2
UMAP1
Protocol
C1HT-medium
C1HT-small
CEL-Seq2
Chromium
Chromium(sn)
ddSEQ
Drop-seq
ICELL8
inDrop
MARS-Seq
Quartz-Seq2
gmcSCRB-seq
Smart-seq2
g
Mixability
0.9
0.8
0.7
0.6
0.5
0.6
0.7
0.8
Clustering Accuracy
c
Mouse Cell types
Enterocyte 1
Enterocyte 2
Enterocyte progenitor
Enteroendocrine
Fibroblast
Immune cell
Secretory cell
Stem cell
Transit Amplifying
UMAP2
UMAP1
d
UMAP2
UMAP1
h
Mixability
1.0
0.9
0.8
0.7
0.6
0.4
0.5
0.6
0.7
0.8
0.9
Clustering Accuracy
e
HEK293T signature
HEK293T cells
B cell signature
B cells
CD14+ Monocyte signature
CD14+ Monocytes
0.0
0.5
1.0
1.5
2.0
f
Secretory cells signature
Secretory cells
Transit Amplifying cells signature
Transit Amplifying cells
2.5


High score
Low score
Method
Gene detection
Marker expression
Clusterability
Mappability
Clusterability (integrated)
Mixability
Benchmarking score
Quartz-Seq2
Chromium
Smart-seq2
CEL-Seq2
C1HT-medium
C1HT-small
ddSEQ
Chromium(sn)
Drop-seq
inDrop
ICELL8
MARS-Seq
gmcSCRB-seq

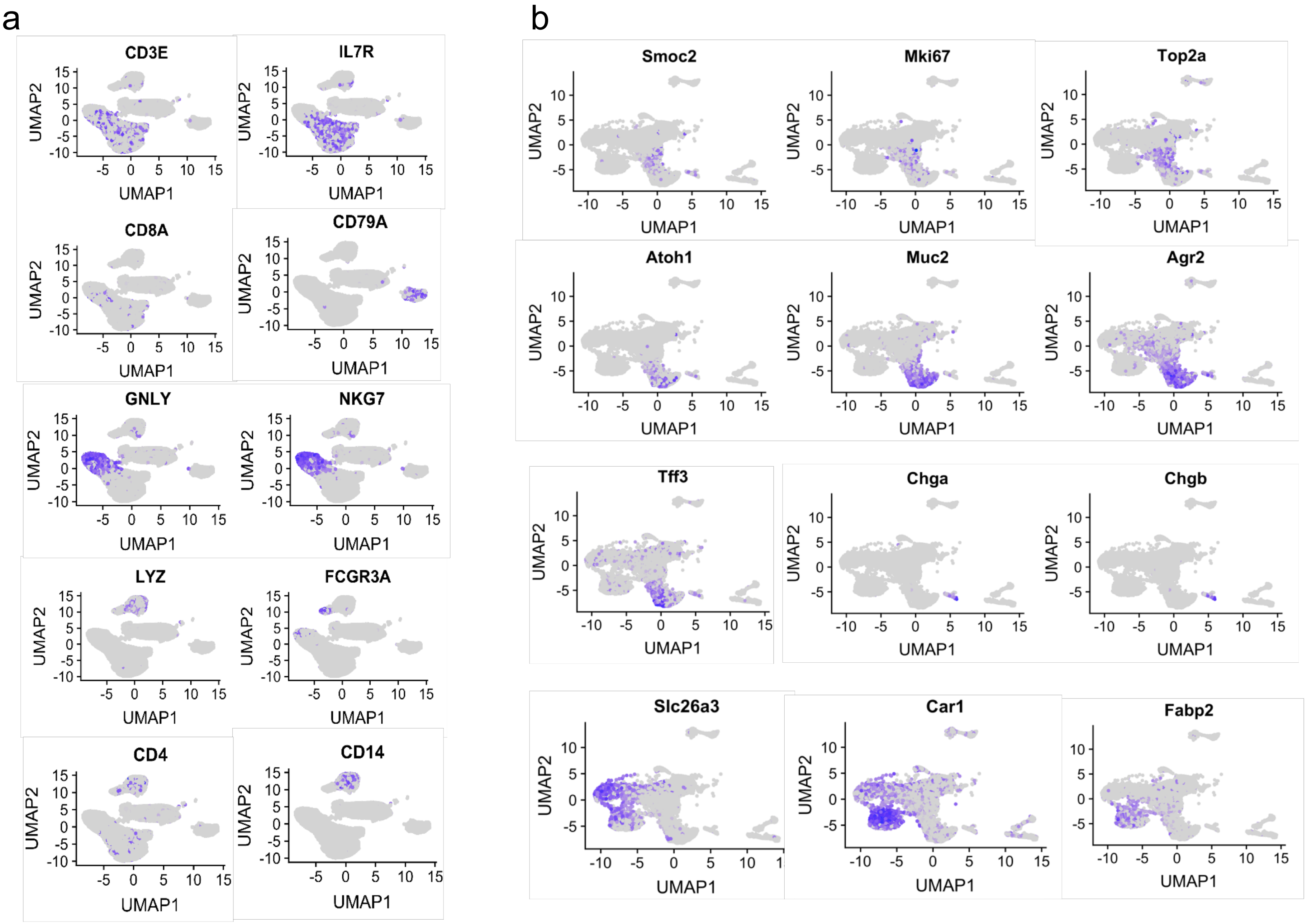
Supplementary Figure 1



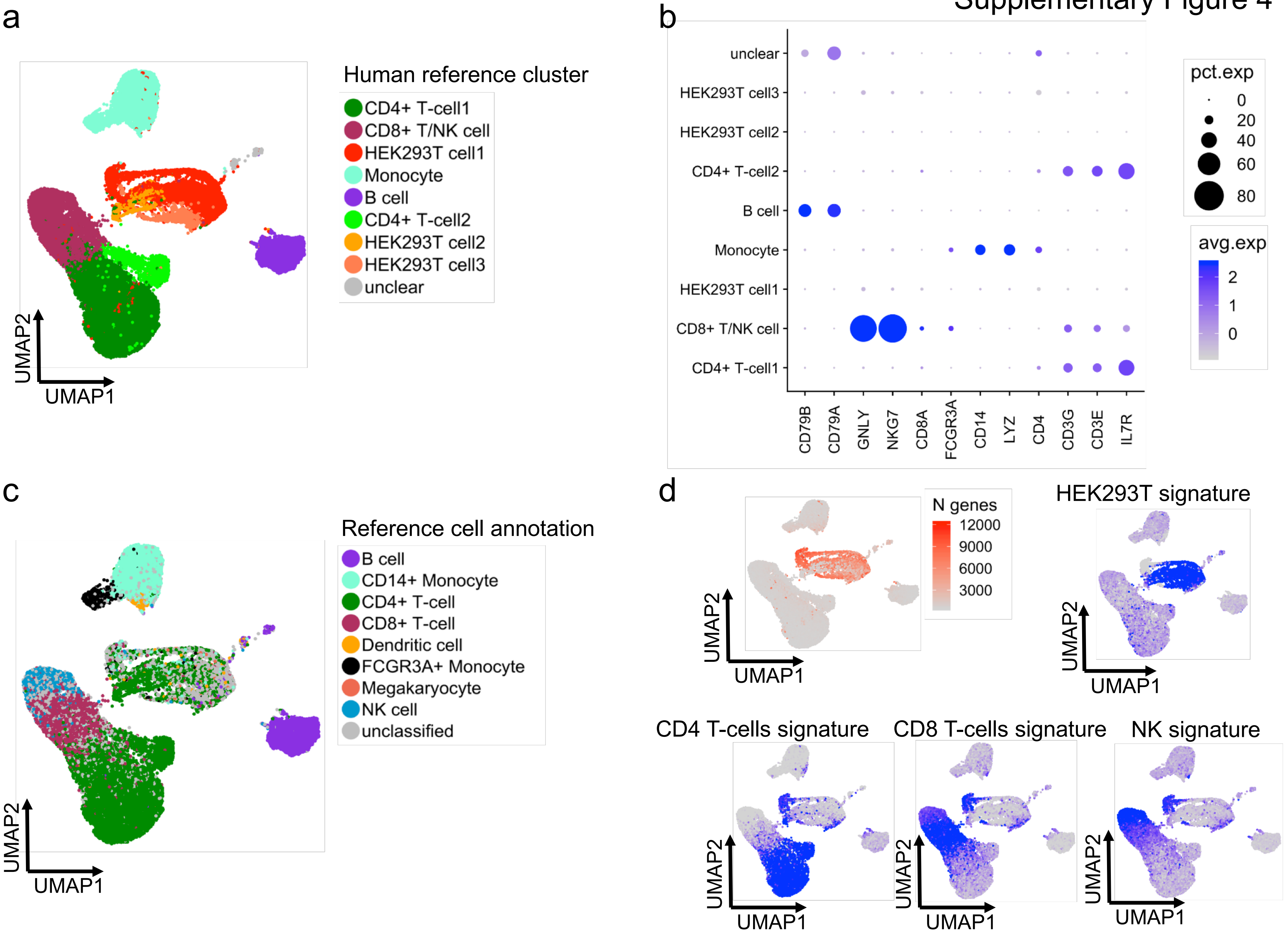
Supplementary Figure 2



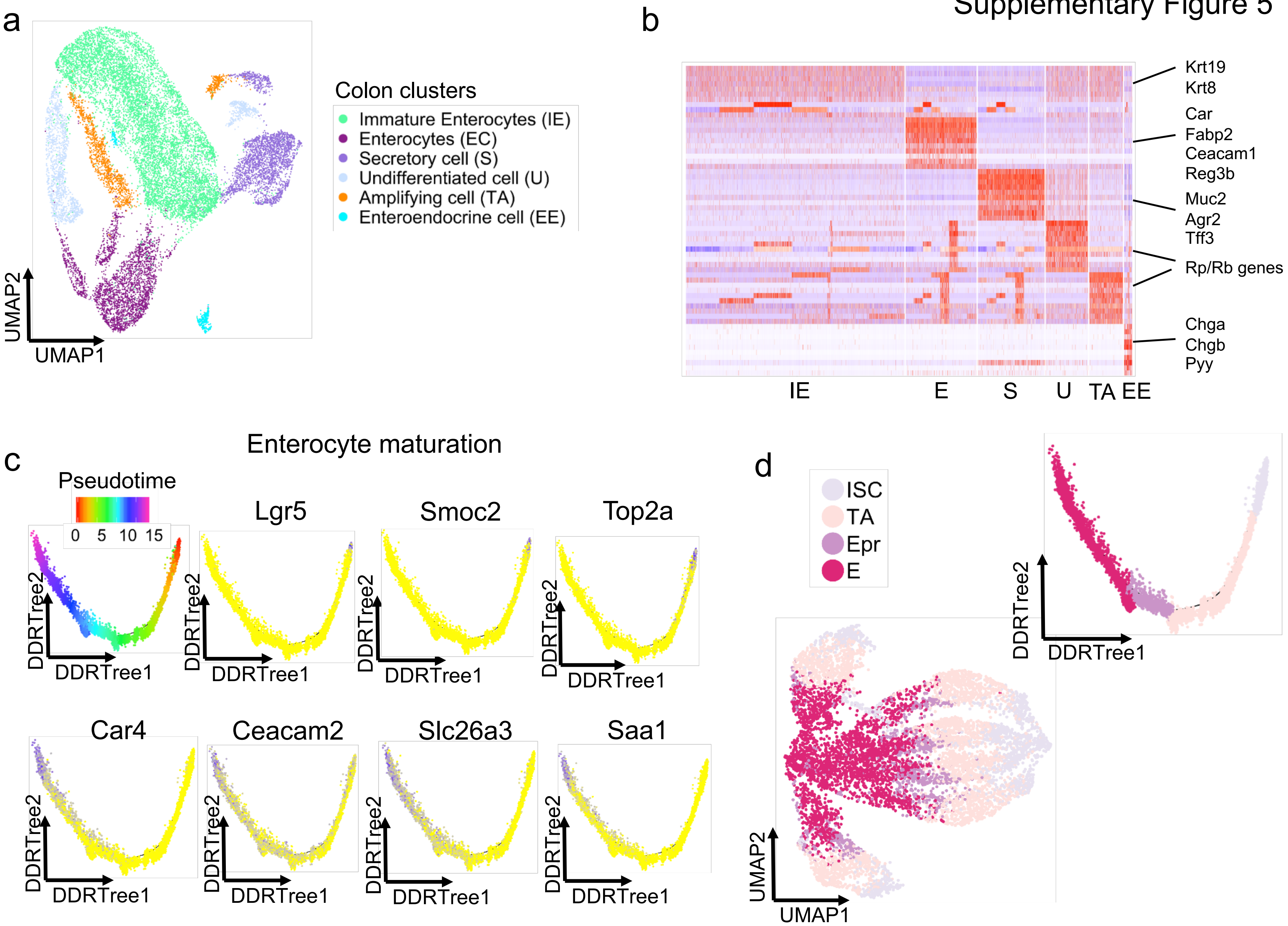
Supplementary Figure 3



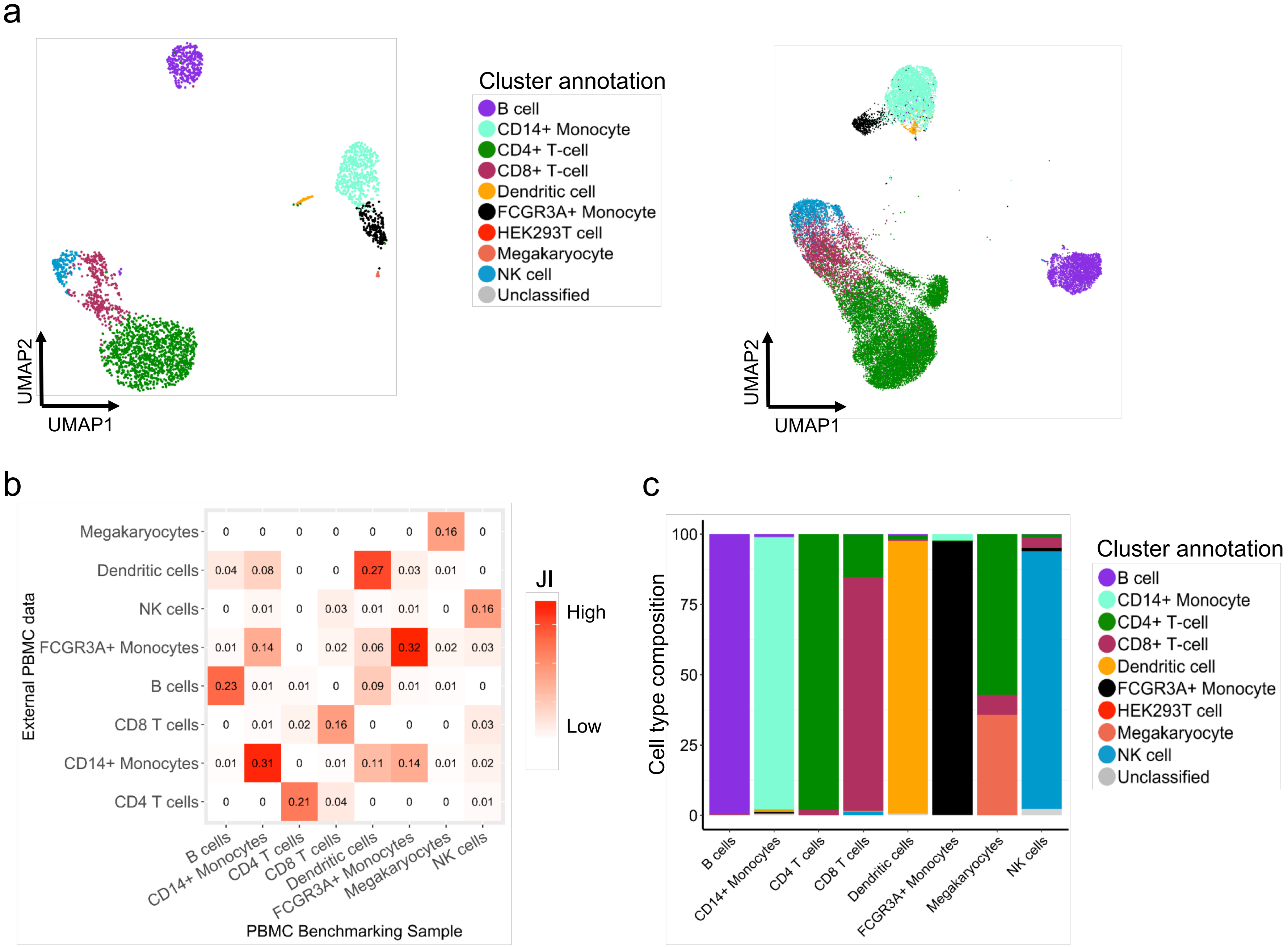
Supplementary Figure 4



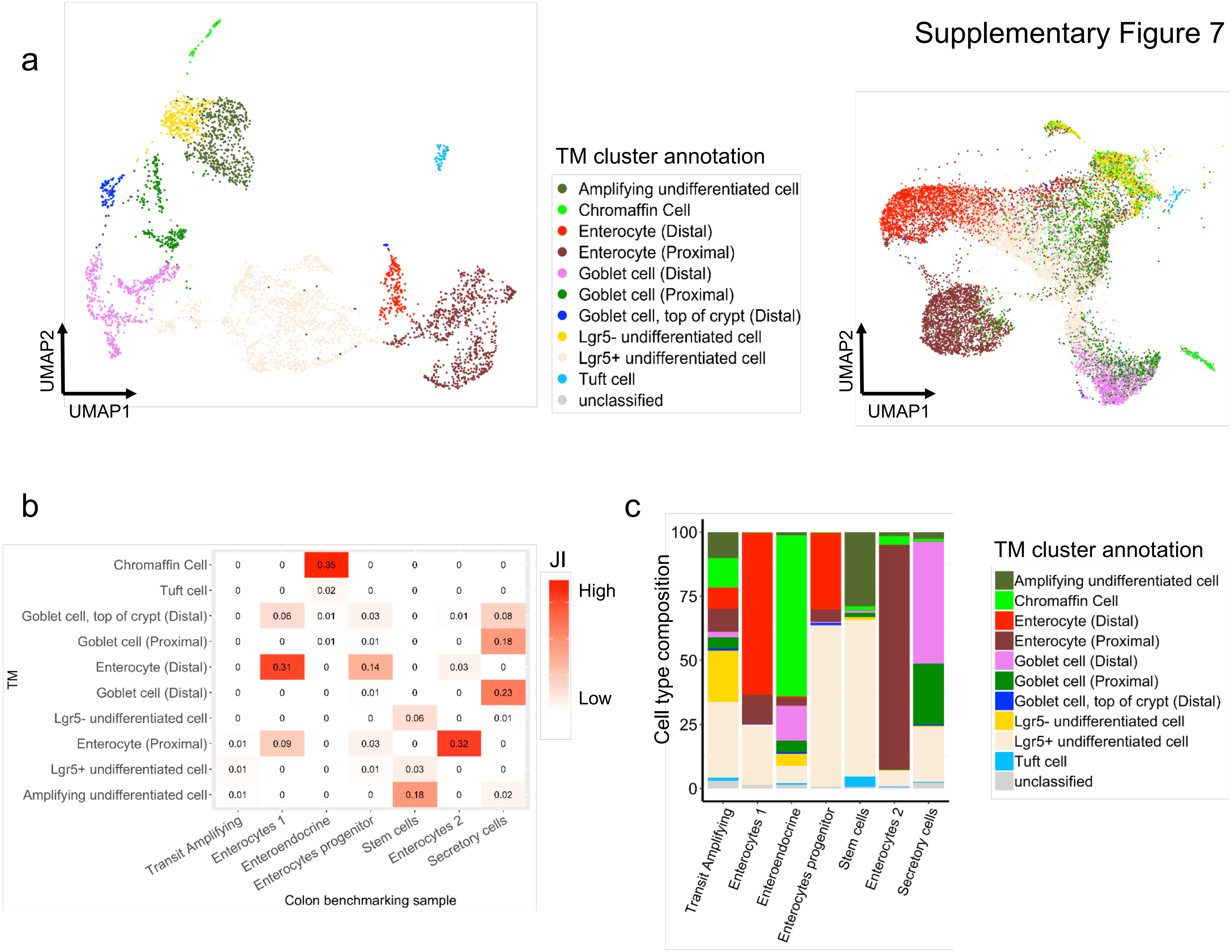
Supplementary Figure 5



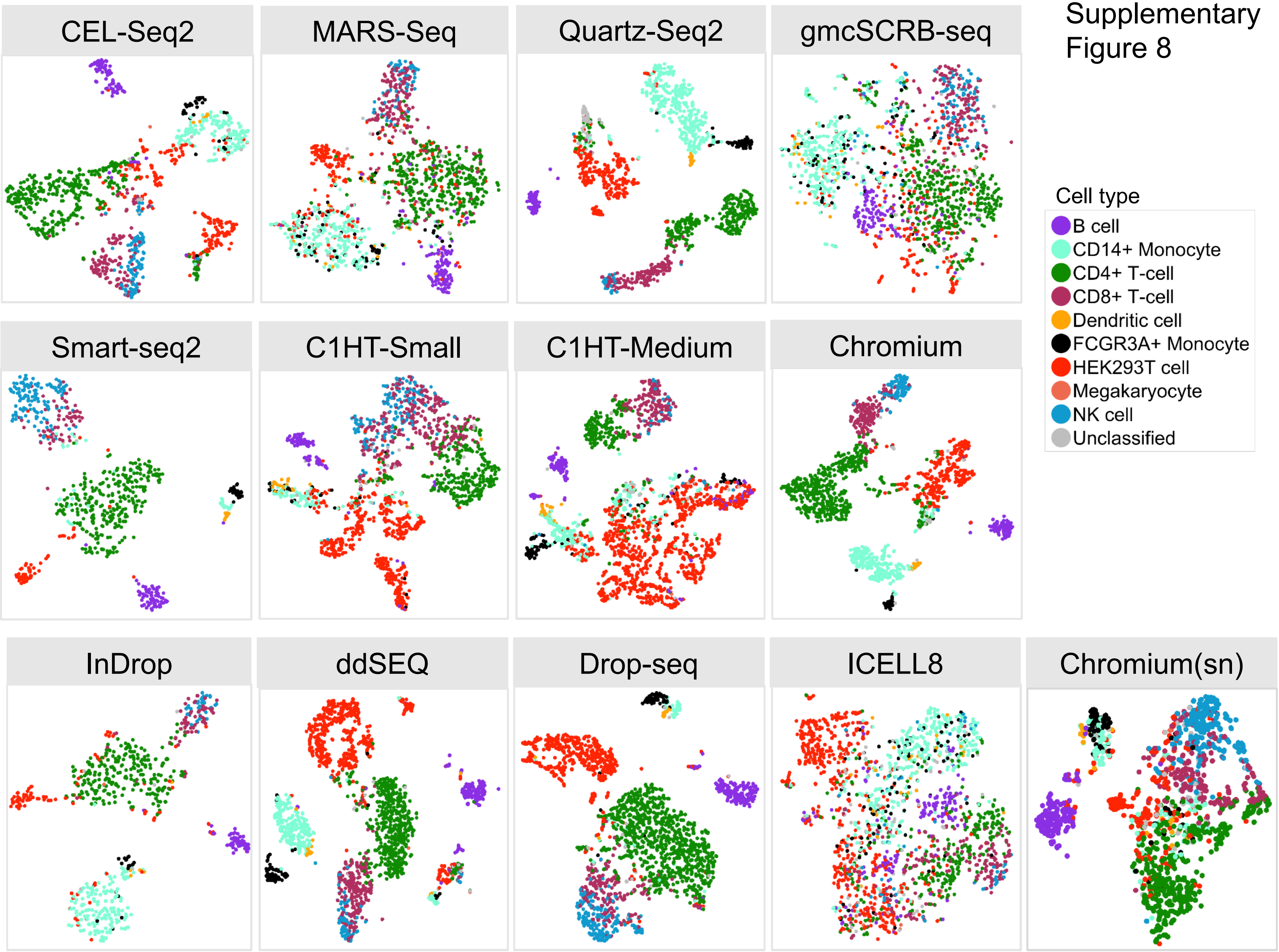
Supplementary Figure 6



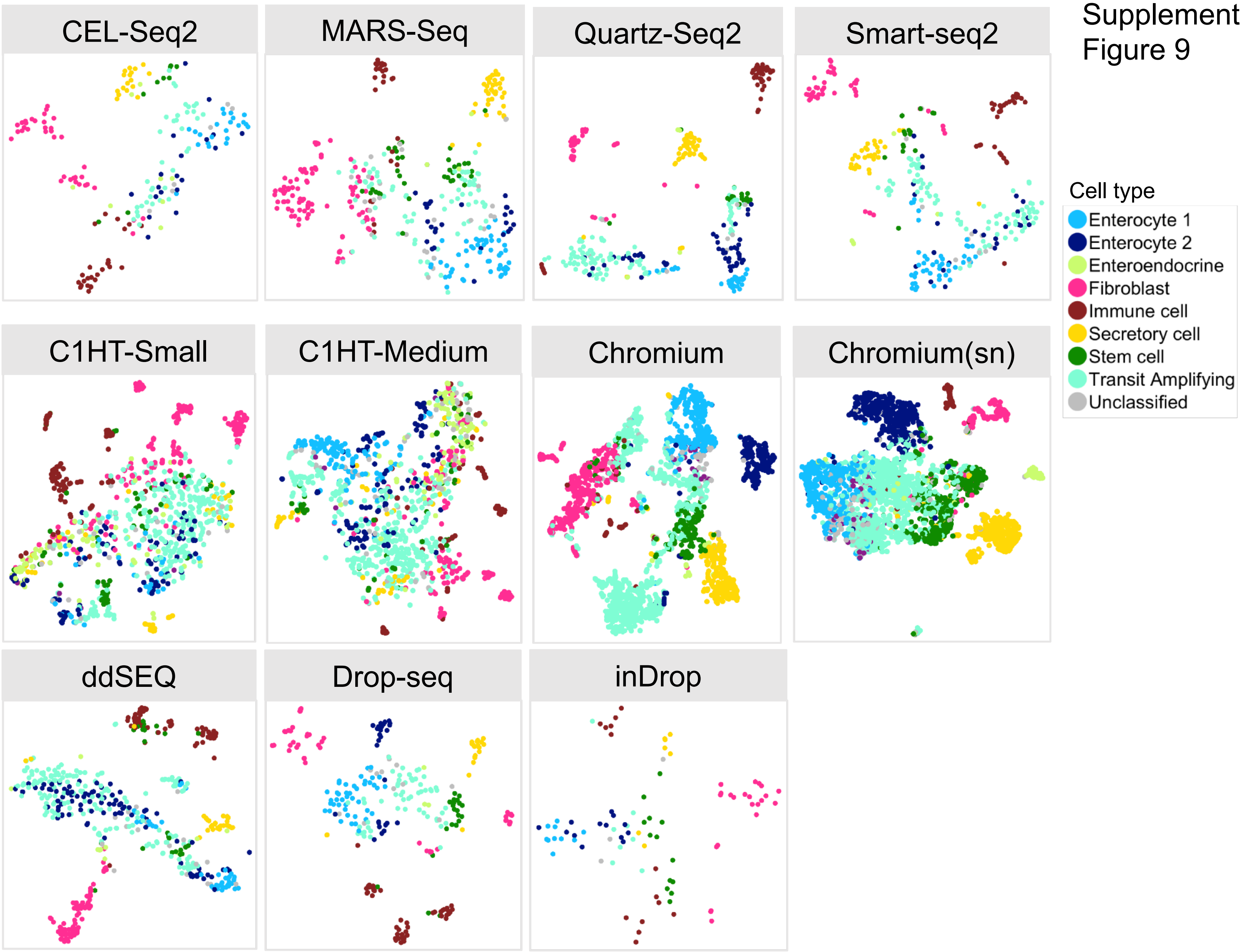
Supplementary Figure 7



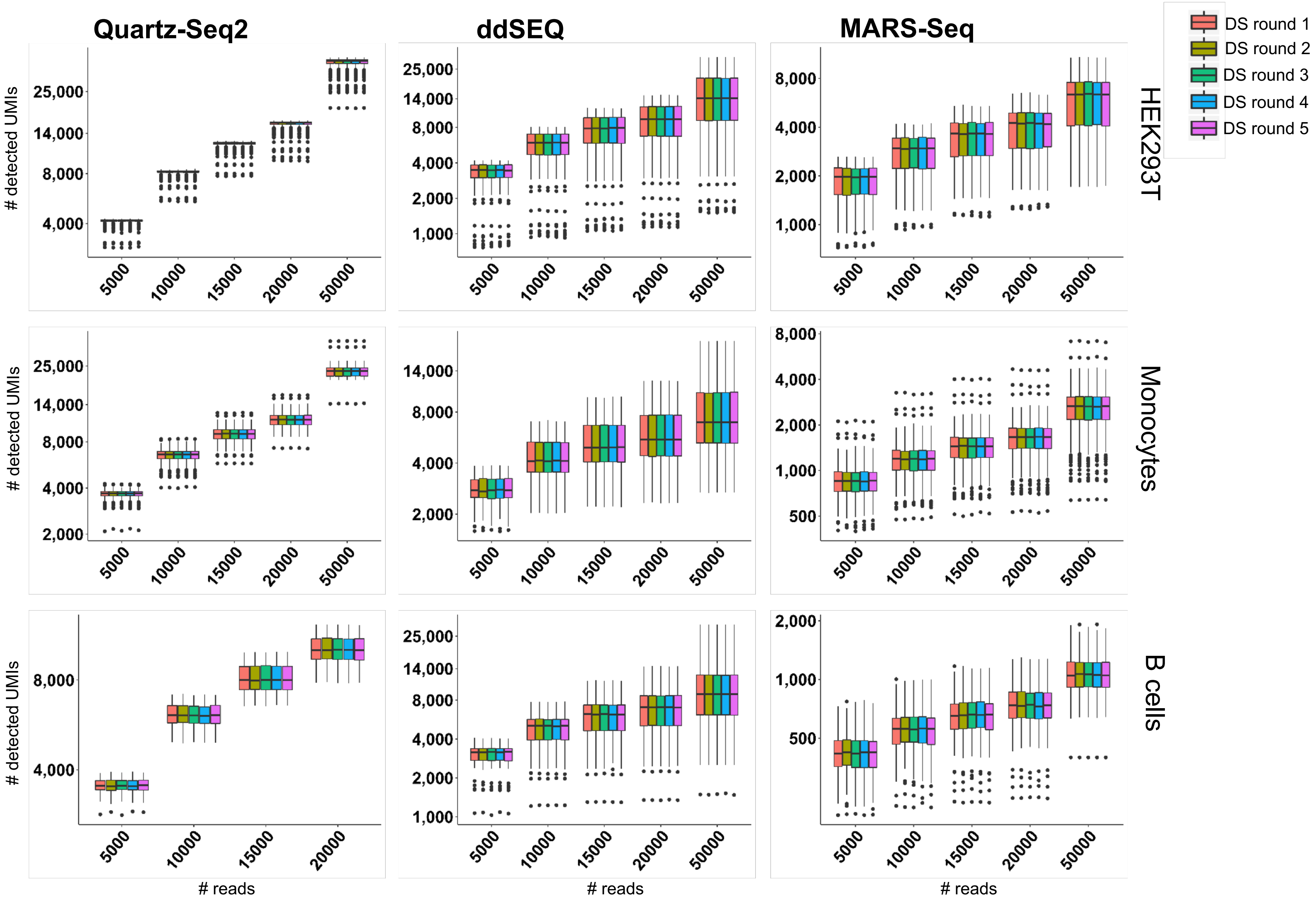
Supplementary Figure 8



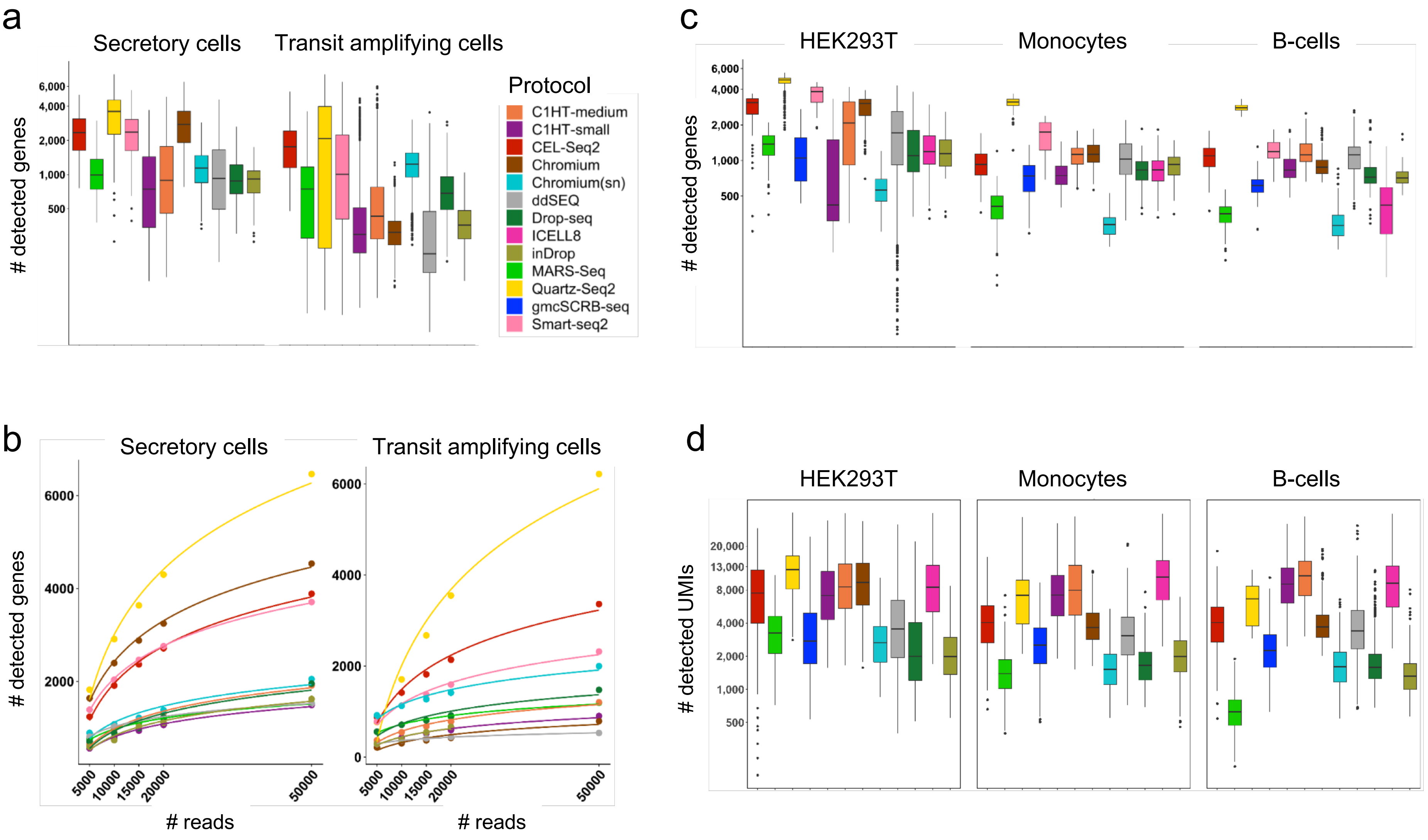
Supplementary Figure 9



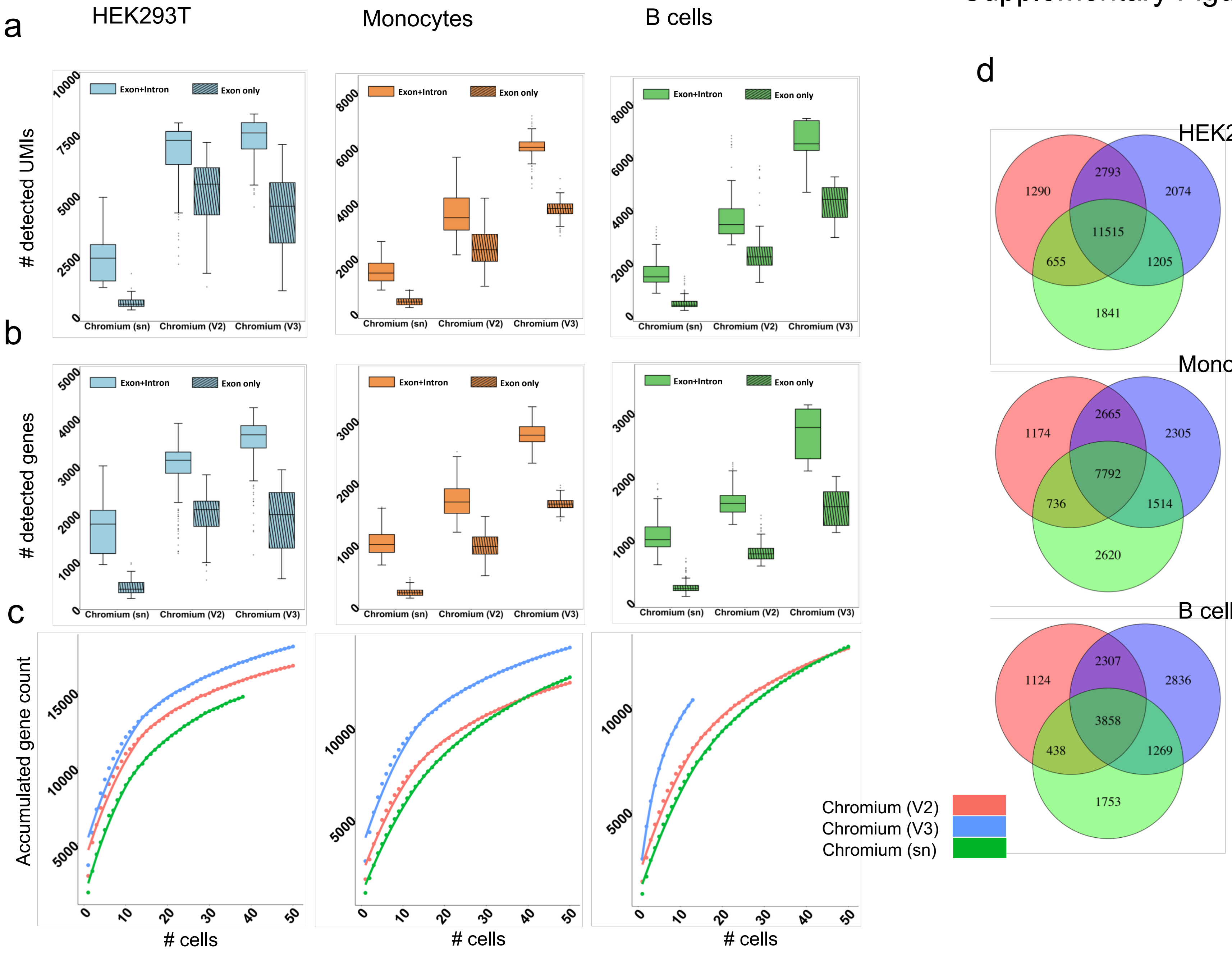
Supplementary Figure 10



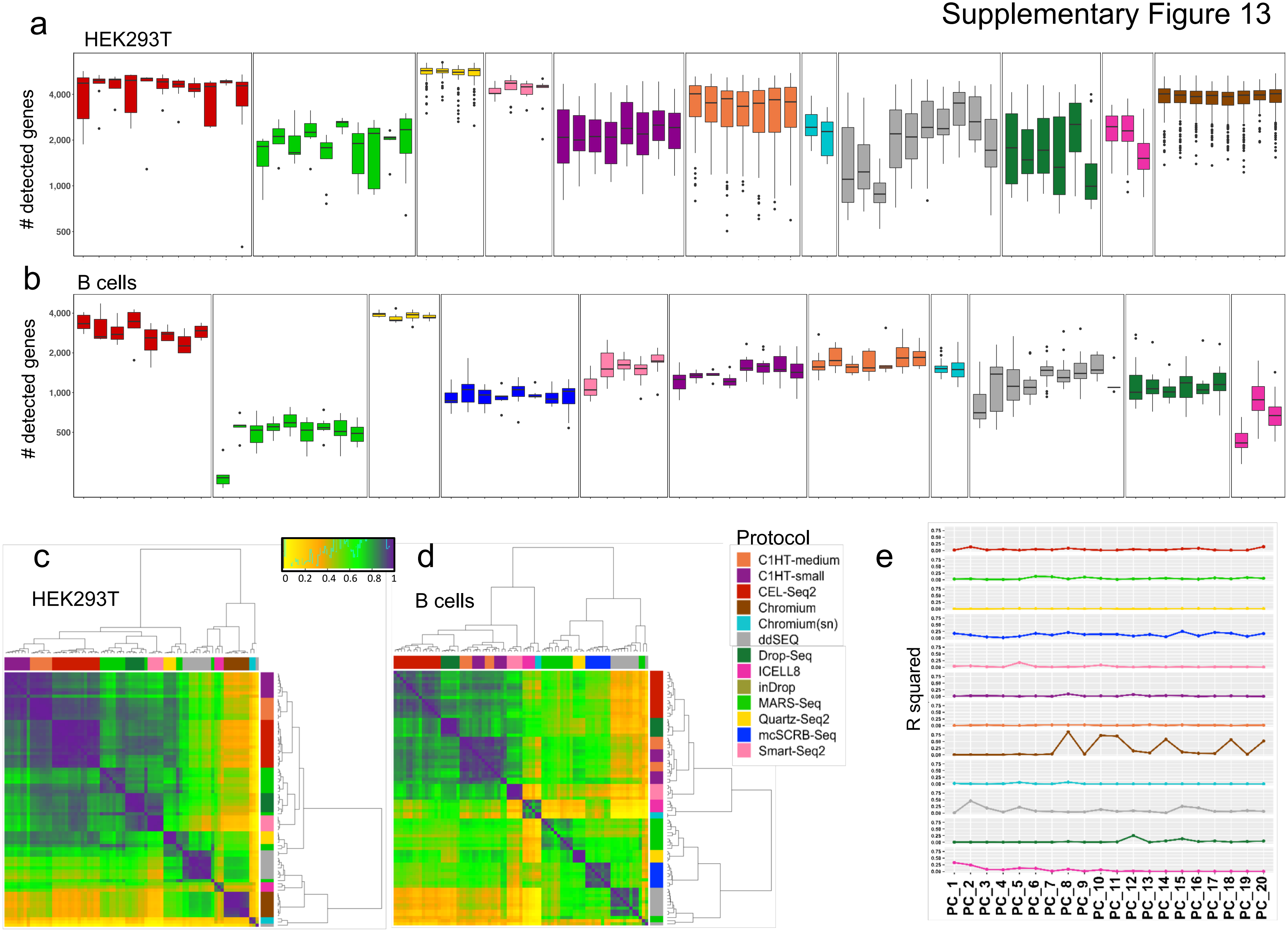
Supplementary Figure 11



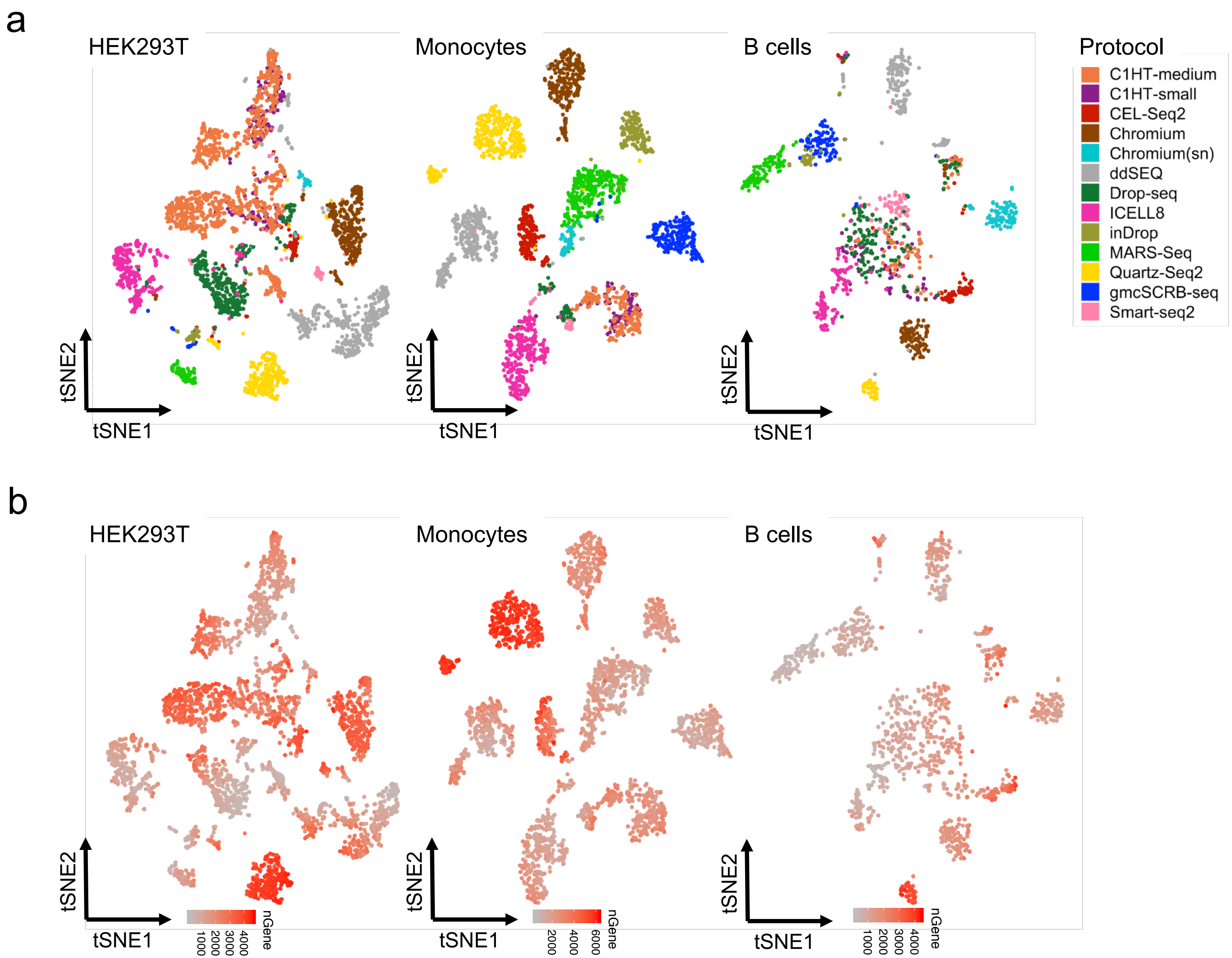
Supplementary Figure 12



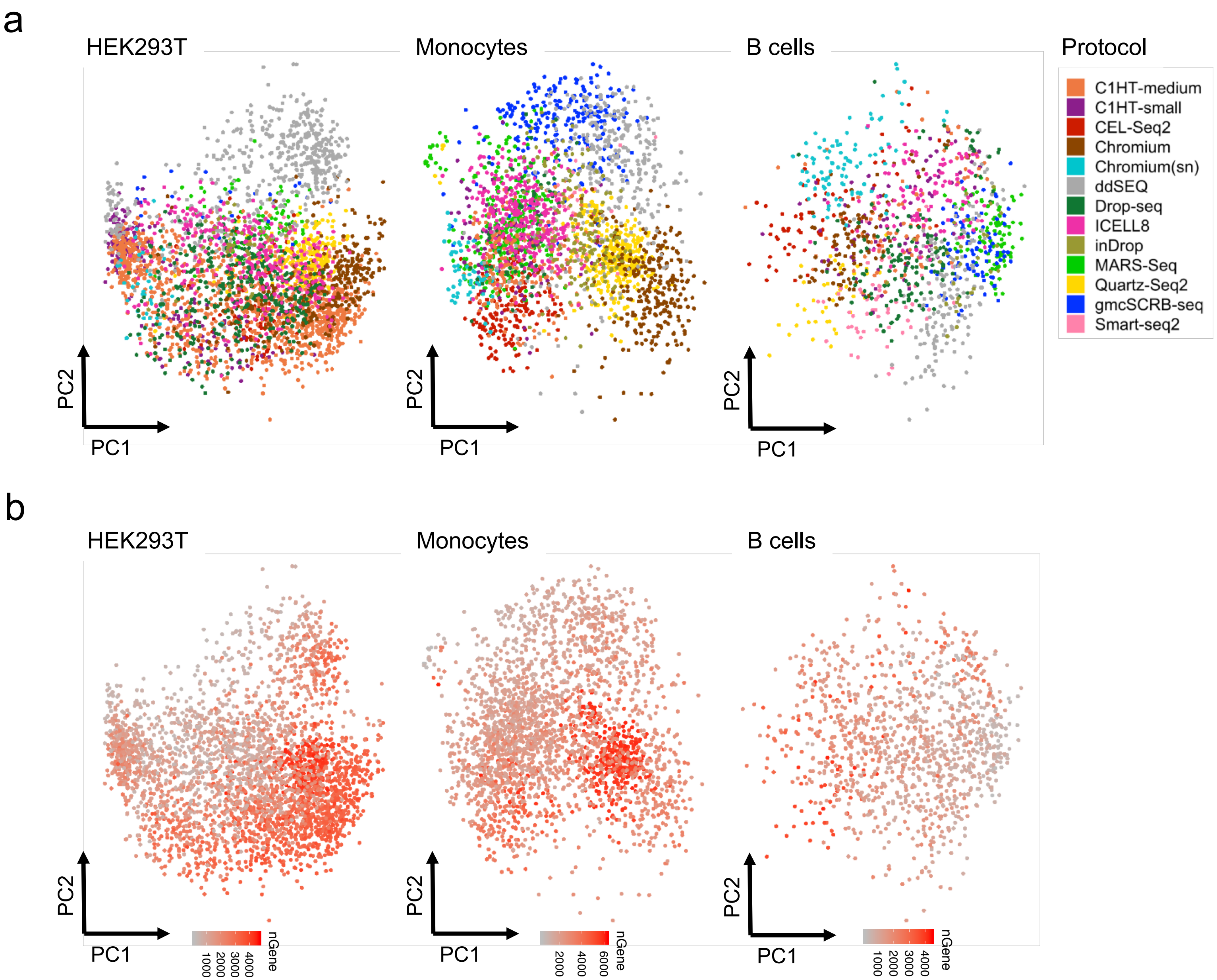
Supplementary Figure 13



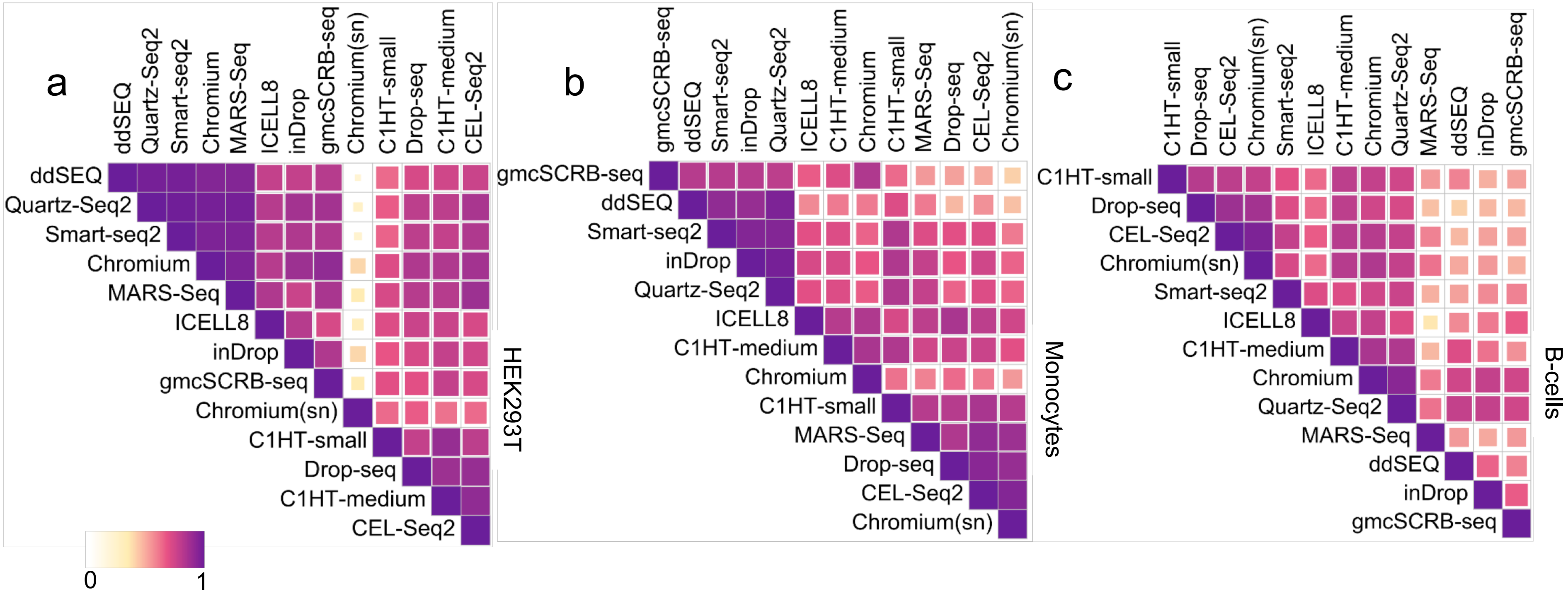
Supplementary Figure 14



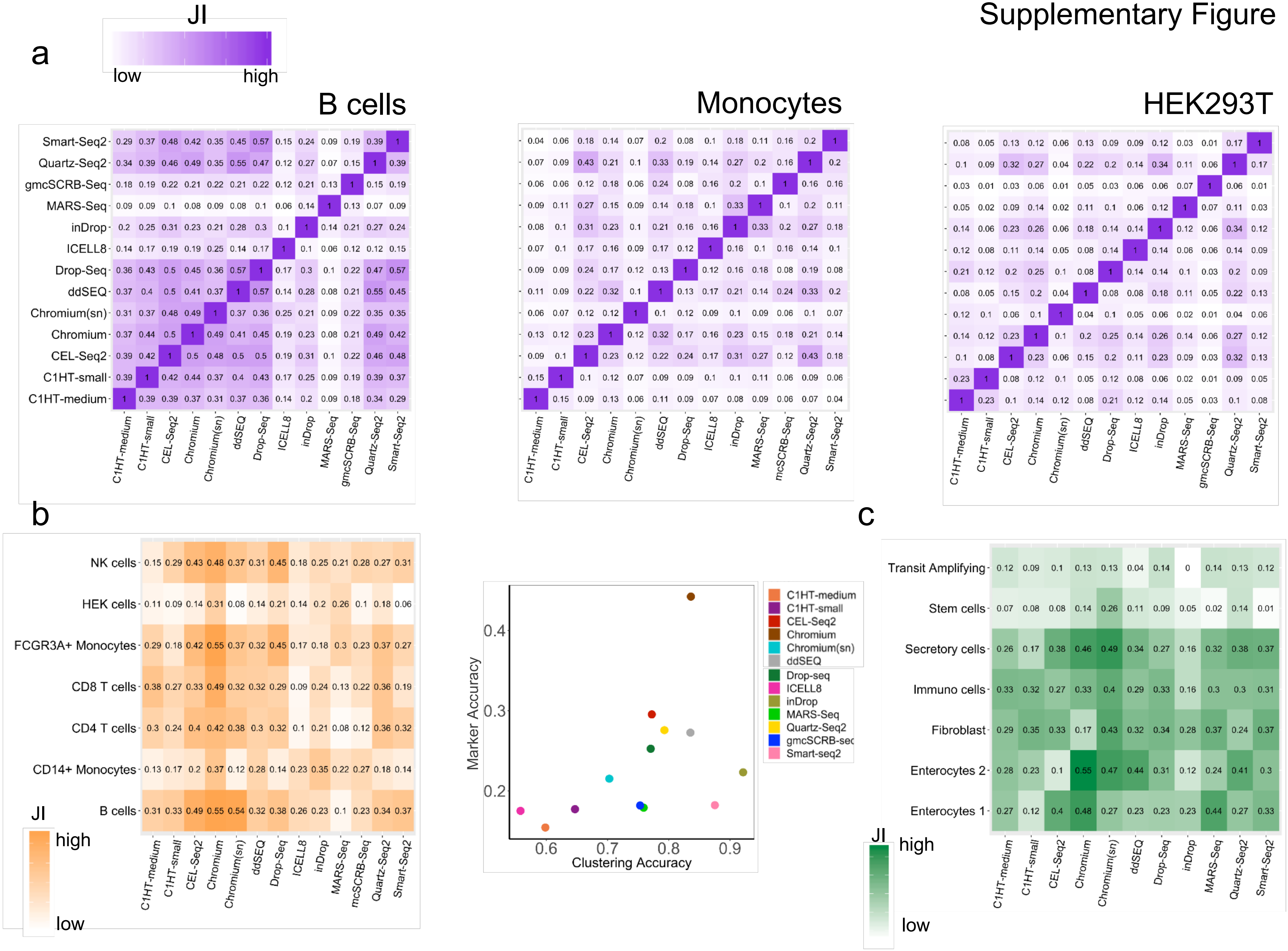
Supplementary Figure 15



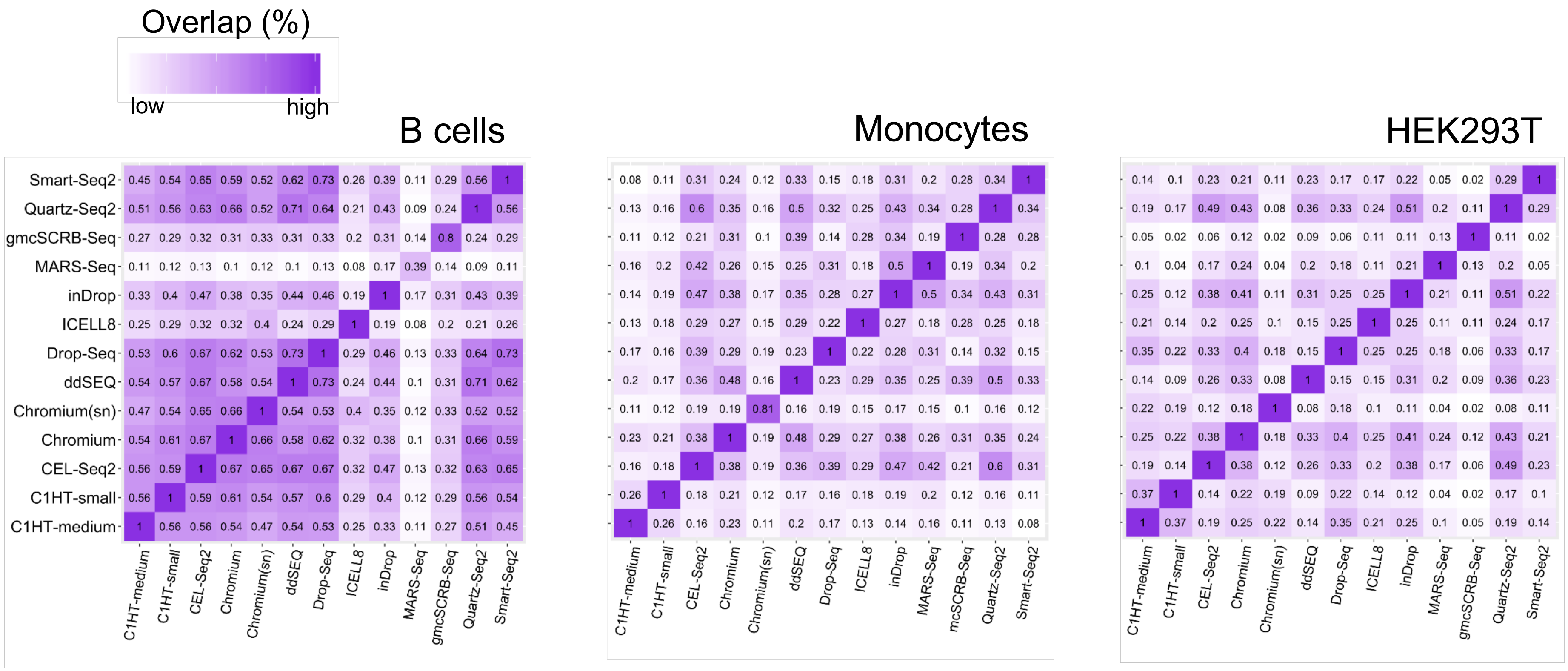
Supplementary Figure 16



Supplementary Figure 17



Supplementary Figure 18



B cells

	C1HT-medium	C1HT-small	CEL-Seq2	Chromium	Chromium(sn)	ddSEQ	Drop-Seq	ICELL8	inDrop	MARS-Seq	gmcSCRB-Seq	Quartz-Seq2	Smart-Seq2
Smart-Seq2	0.45	0.54	0.65	0.59	0.52	0.62	0.73	0.26	0.39	0.11	0.29	0.56	1
Quartz-Seq2	0.51	0.56	0.63	0.66	0.52	0.71	0.64	0.21	0.43	0.09	0.24	1	0.56
gmcSCRB-Seq	0.27	0.29	0.32	0.31	0.33	0.31	0.33	0.2	0.31	0.14	0.8	0.24	0.29
MARS-Seq	0.11	0.12	0.13	0.1	0.12	0.1	0.13	0.08	0.17	0.39	0.14	0.09	0.11
inDrop	0.33	0.4	0.47	0.38	0.35	0.44	0.46	0.19	1	0.17	0.31	0.43	0.39
ICELL8	0.25	0.29	0.32	0.32	0.4	0.24	0.29	1	0.19	0.08	0.2	0.21	0.26
Drop-Seq	0.53	0.6	0.67	0.62	0.53	0.73	1	0.29	0.46	0.13	0.33	0.64	0.73
ddSEQ	0.54	0.57	0.67	0.58	0.54	1	0.73	0.24	0.44	0.1	0.31	0.71	0.62
Chromium(sn)	0.47	0.54	0.65	0.66	1	0.54	0.53	0.4	0.35	0.12	0.33	0.52	0.52
Chromium	0.54	0.61	0.67	1	0.66	0.58	0.62	0.32	0.38	0.1	0.31	0.66	0.59
CEL-Seq2	0.56	0.59	1	0.67	0.65	0.67	0.67	0.32	0.47	0.13	0.32	0.63	0.65
C1HT-small	0.56	1	0.59	0.61	0.54	0.57	0.6	0.29	0.4	0.12	0.29	0.56	0.54
C1HT-medium	1	0.56	0.56	0.54	0.47	0.54	0.53	0.25	0.33	0.11	0.27	0.51	0.45

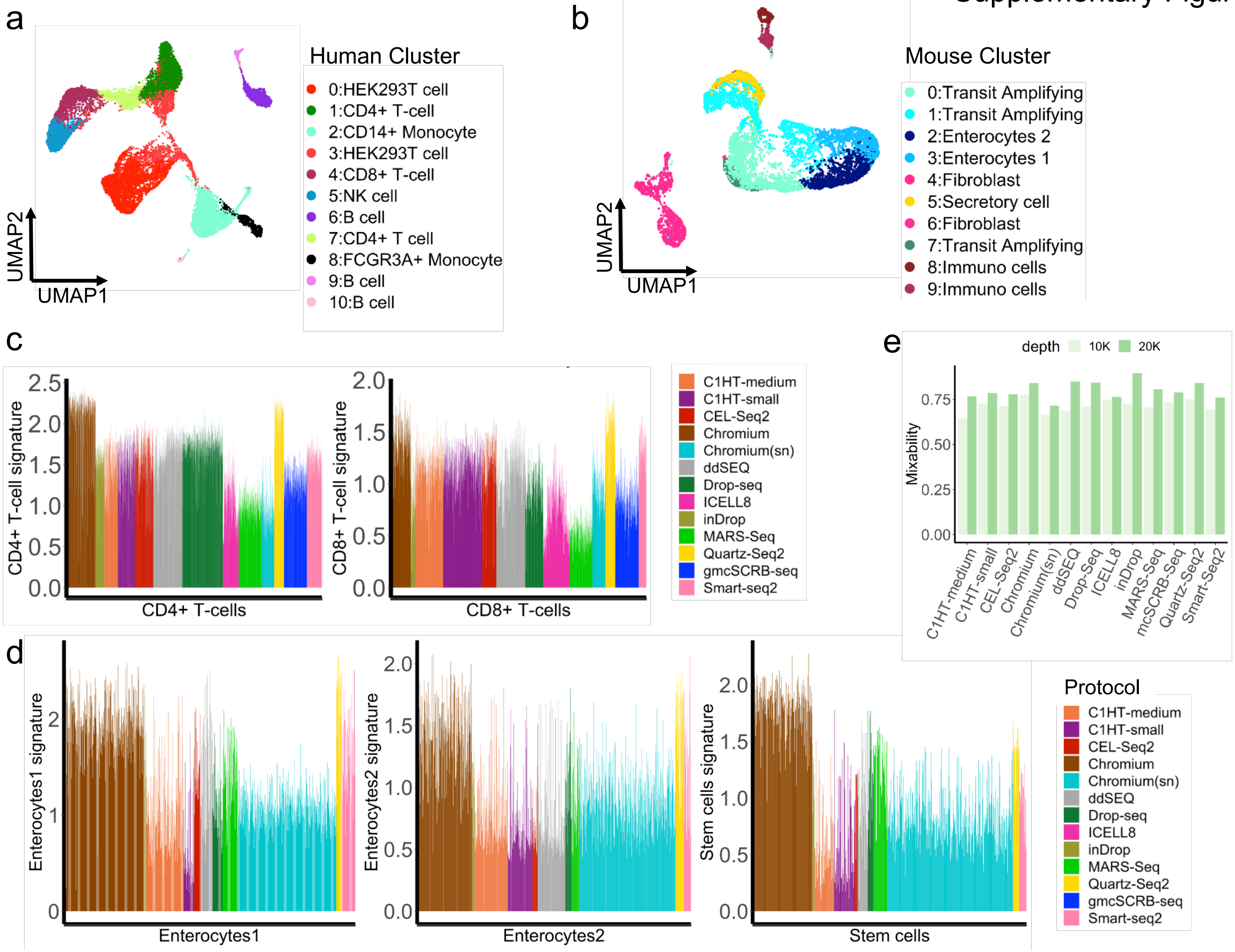
Monocytes

C1HT-medium	C1HT-small	CEL-Seq2	Chromium	Chromium(sn)	ddSEQ	Drop-Seq	ICELL8	inDrop	MARS-Seq	mcSCRB-Seq	Quartz-Seq2	Smart-Seq2
0.08	0.11	0.31	0.24	0.12	0.33	0.15	0.18	0.31	0.2	0.28	0.34	1
0.13	0.16	0.6	0.35	0.16	0.5	0.32	0.25	0.43	0.34	0.28	1	0.34
0.11	0.12	0.21	0.31	0.1	0.39	0.14	0.28	0.34	0.19	1	0.28	0.28
0.16	0.2	0.42	0.26	0.15	0.25	0.31	0.18	0.5	1	0.19	0.34	0.2
0.14	0.19	0.47	0.38	0.17	0.35	0.28	0.27	1	0.5	0.34	0.43	0.31
0.13	0.18	0.29	0.27	0.15	0.29	0.22	1	0.27	0.18	0.28	0.25	0.18
0.17	0.16	0.39	0.29	0.19	0.23	1	0.22	0.28	0.31	0.14	0.32	0.15
0.2	0.17	0.36	0.48	0.16	1	0.23	0.29	0.35	0.25	0.39	0.5	0.33
0.11	0.12	0.19	0.19	0.81	0.16	0.19	0.15	0.17	0.15	0.1	0.16	0.12
0.23	0.21	0.38	1	0.19	0.48	0.29	0.27	0.38	0.26	0.31	0.35	0.24
0.16	0.18	1	0.38	0.19	0.36	0.39	0.29	0.47	0.42	0.21	0.6	0.31
0.26	1	0.18	0.21	0.12	0.17	0.16	0.18	0.19	0.2	0.12	0.16	0.11
1	0.26	0.16	0.23	0.11	0.2	0.17	0.13	0.14	0.16	0.11	0.13	0.08

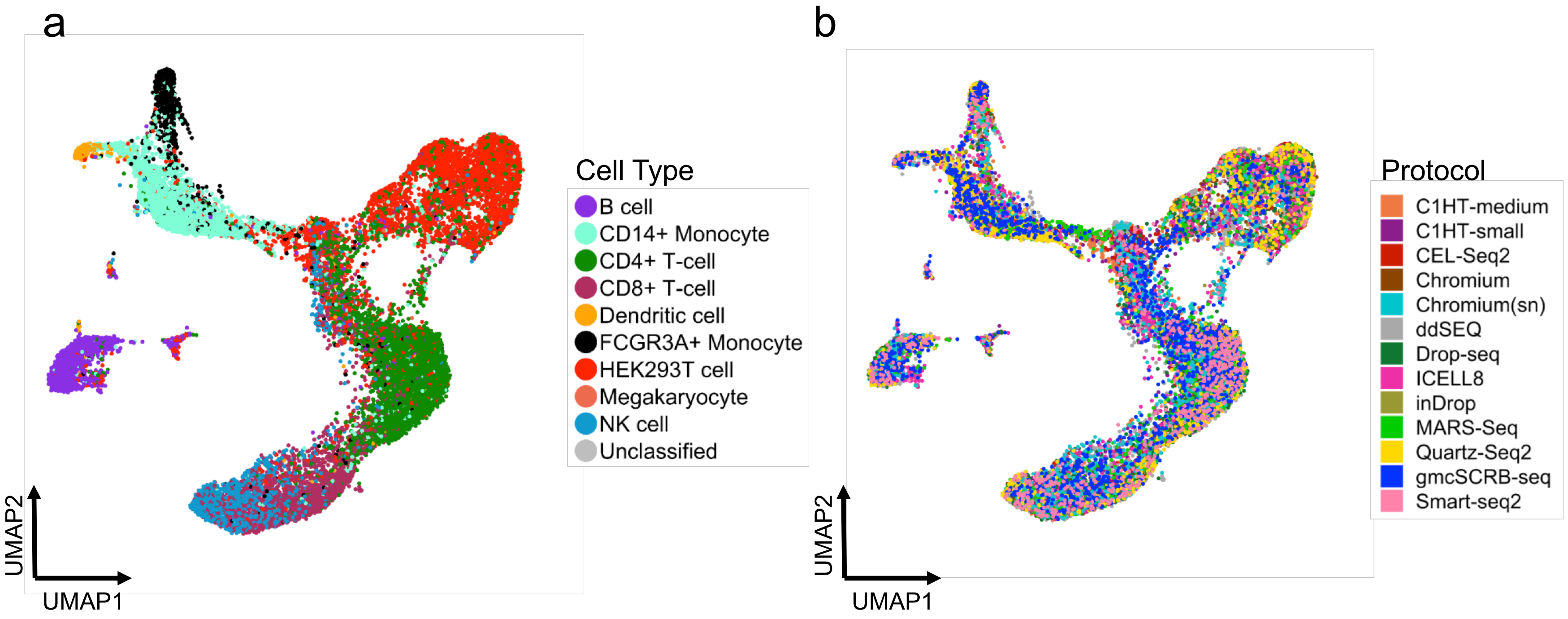
HEK293T

C1HT-medium	C1HT-small	CEL-Seq2	Chromium	Chromium(sn)	ddSEQ	Drop-Seq	ICELL8	inDrop	MARS-Seq	gmcSCRB-Seq	Quartz-Seq2	Smart-Seq2
0.14	0.1	0.23	0.21	0.11	0.23	0.17	0.17	0.22	0.05	0.02	0.29	1
0.19	0.17	0.49	0.43	0.08	0.36	0.33	0.24	0.51	0.2	0.11	1	0.29
0.05	0.02	0.06	0.12	0.02	0.09	0.06	0.11	0.11	0.13	1	0.11	0.02
0.1	0.04	0.17	0.24	0.04	0.2	0.18	0.11	0.21	1	0.13	0.2	0.05
0.25	0.12	0.38	0.41	0.11	0.31	0.25	0.25	1	0.21	0.11	0.51	0.22
0.21	0.14	0.2	0.25	0.1	0.15	0.25	1	0.25	0.11	0.11	0.24	0.17
0.35	0.22	0.33	0.4	0.18	0.15	1	0.25	0.25	0.18	0.06	0.33	0.17
0.14	0.09	0.26	0.33	0.08	1	0.15	0.15	0.31	0.2	0.09	0.36	0.23
0.22	0.19	0.12	0.18	1	0.08	0.18	0.1	0.11	0.04	0.02	0.08	0.11
0.25	0.22	0.38	1	0.18	0.33	0.4	0.25	0.41	0.24	0.12	0.43	0.21
0.19	0.14	1	0.38	0.12	0.26	0.33	0.2	0.38	0.17	0.06	0.49	0.23
0.37	1	0.14	0.22	0.19	0.09	0.22	0.14	0.12	0.04	0.02	0.17	0.1
1	0.37	0.19	0.25	0.22	0.14	0.35	0.21	0.25	0.1	0.05	0.19	0.14

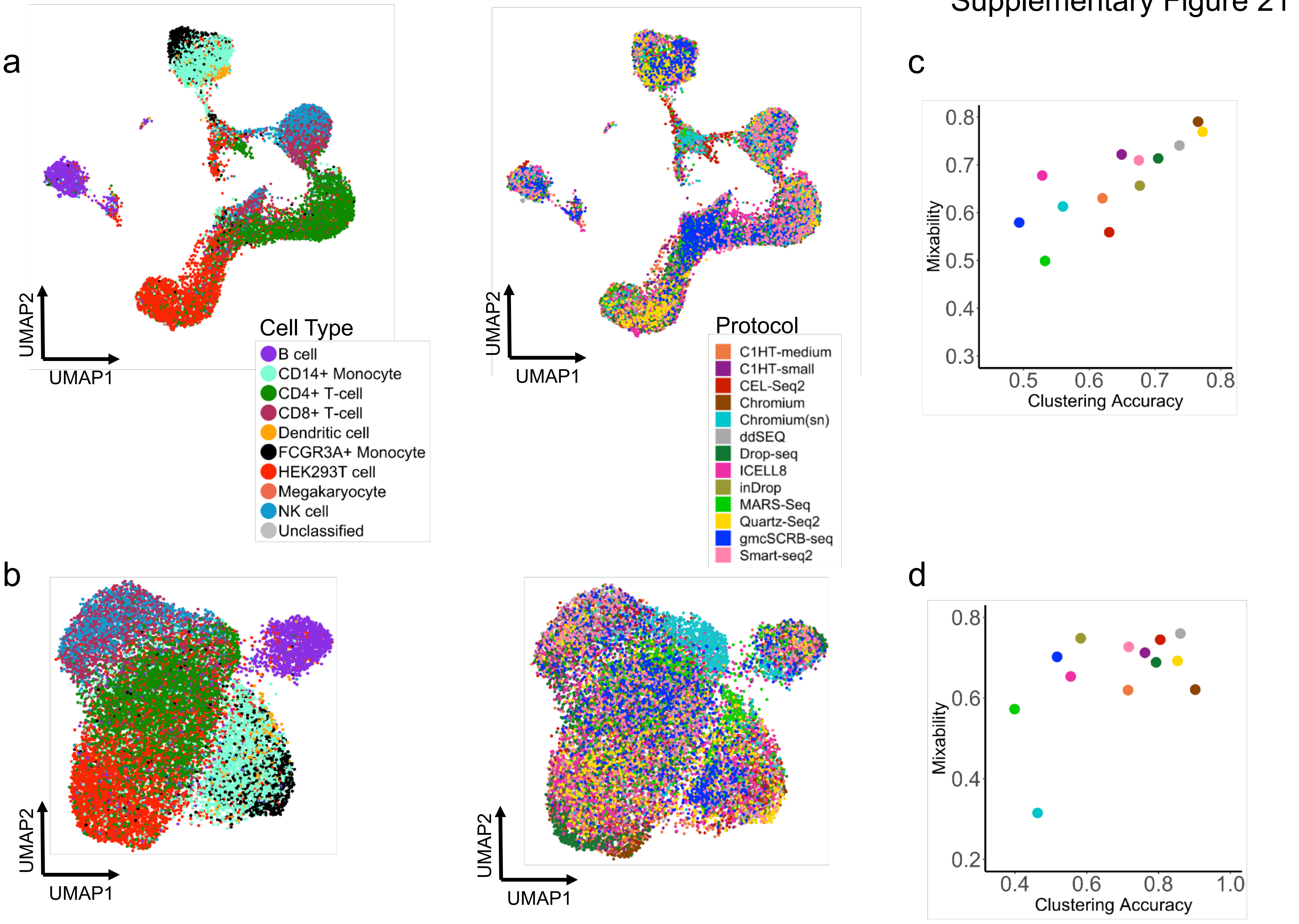
Supplementary Figure 19



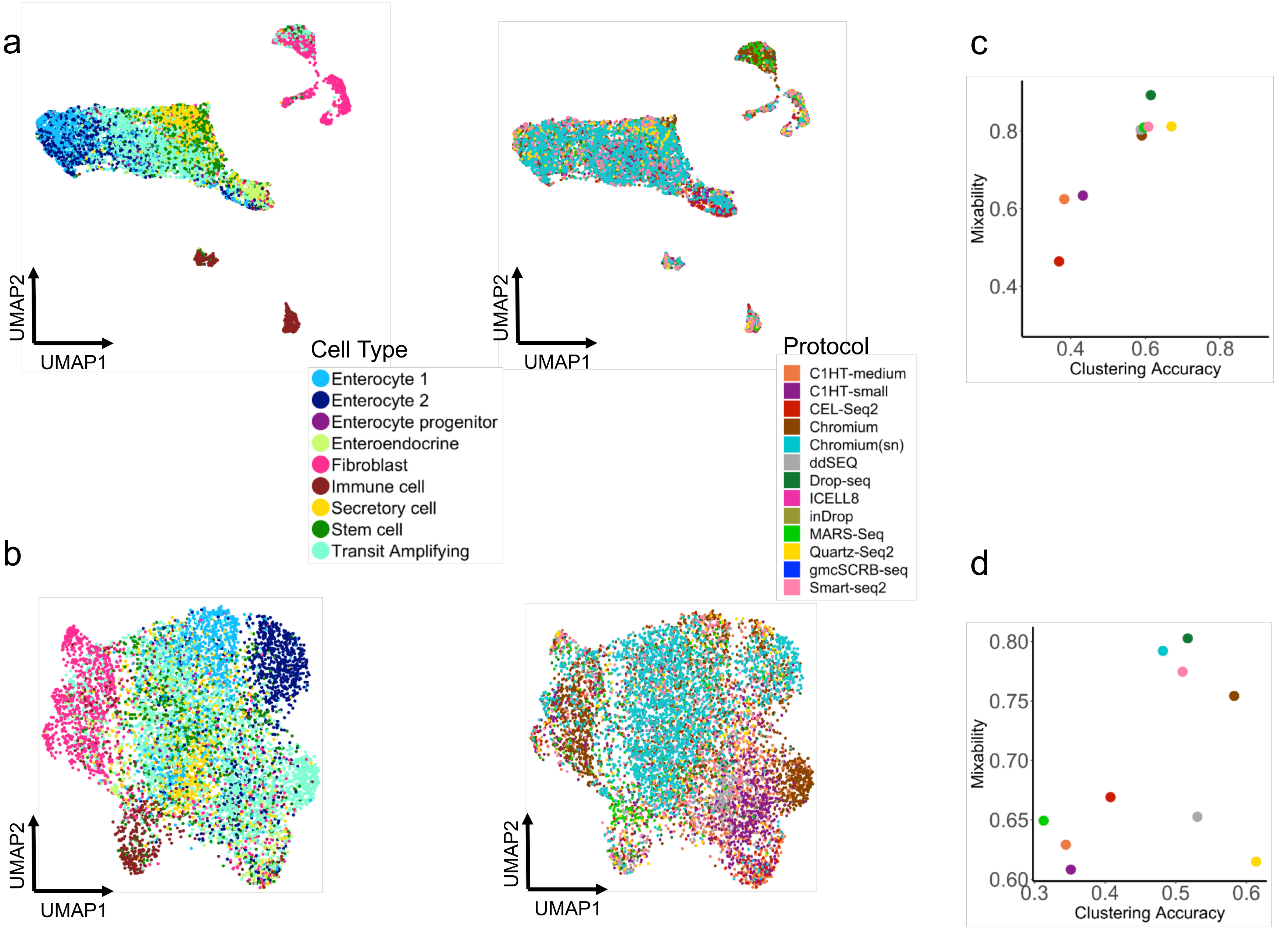
Supplementary Figure 20



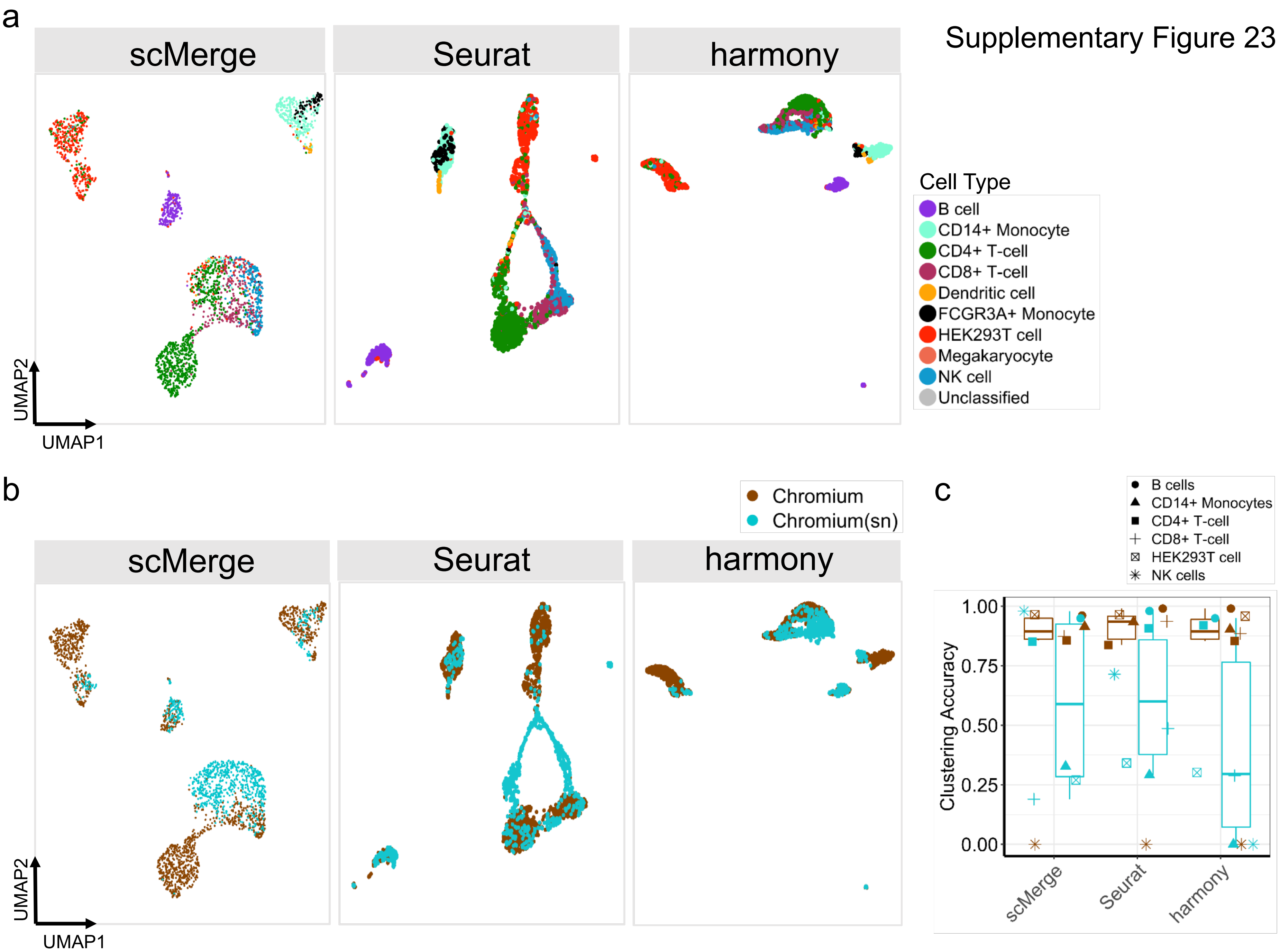
Supplementary Figure 21



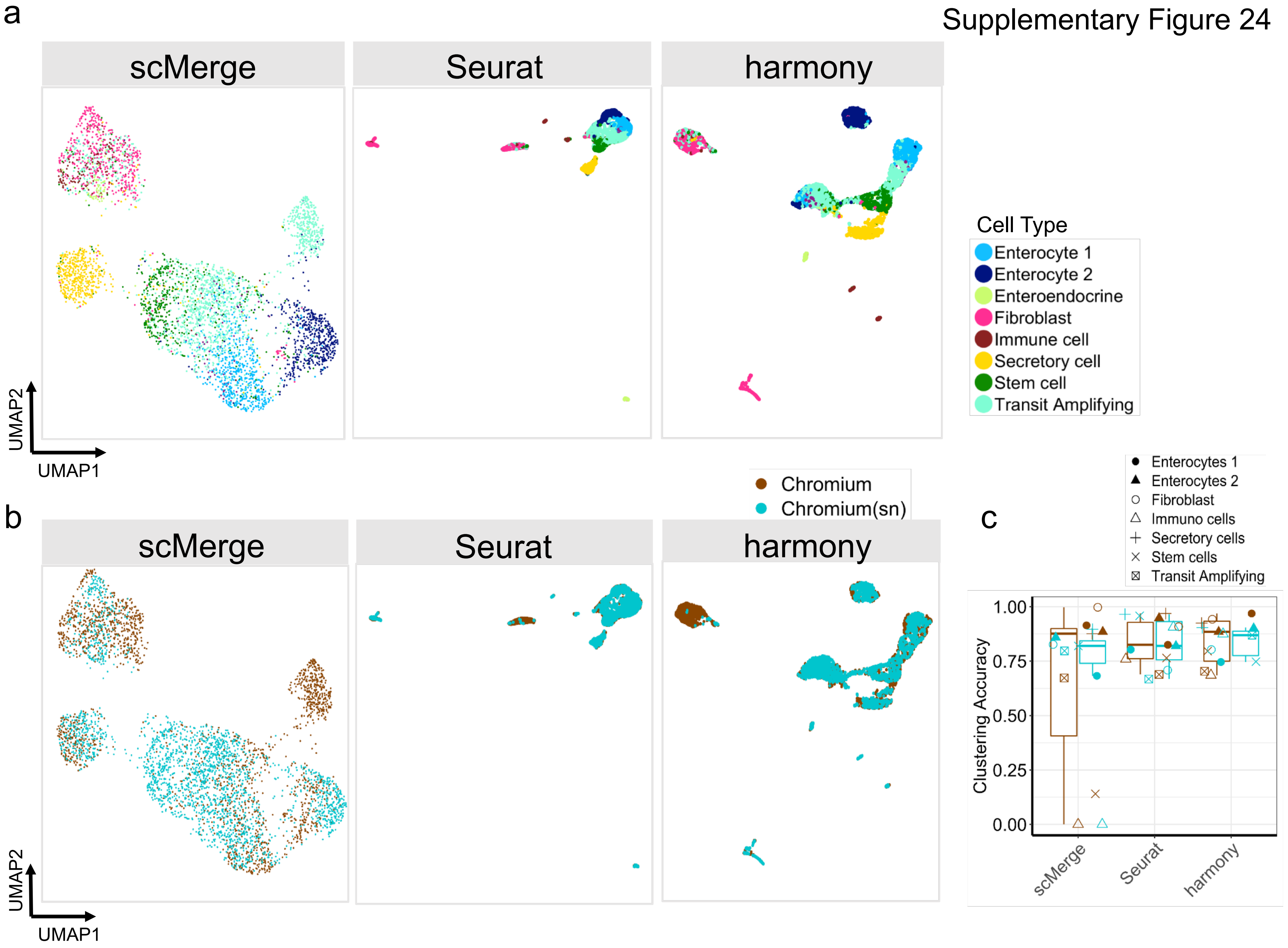
Supplementary Figure 22



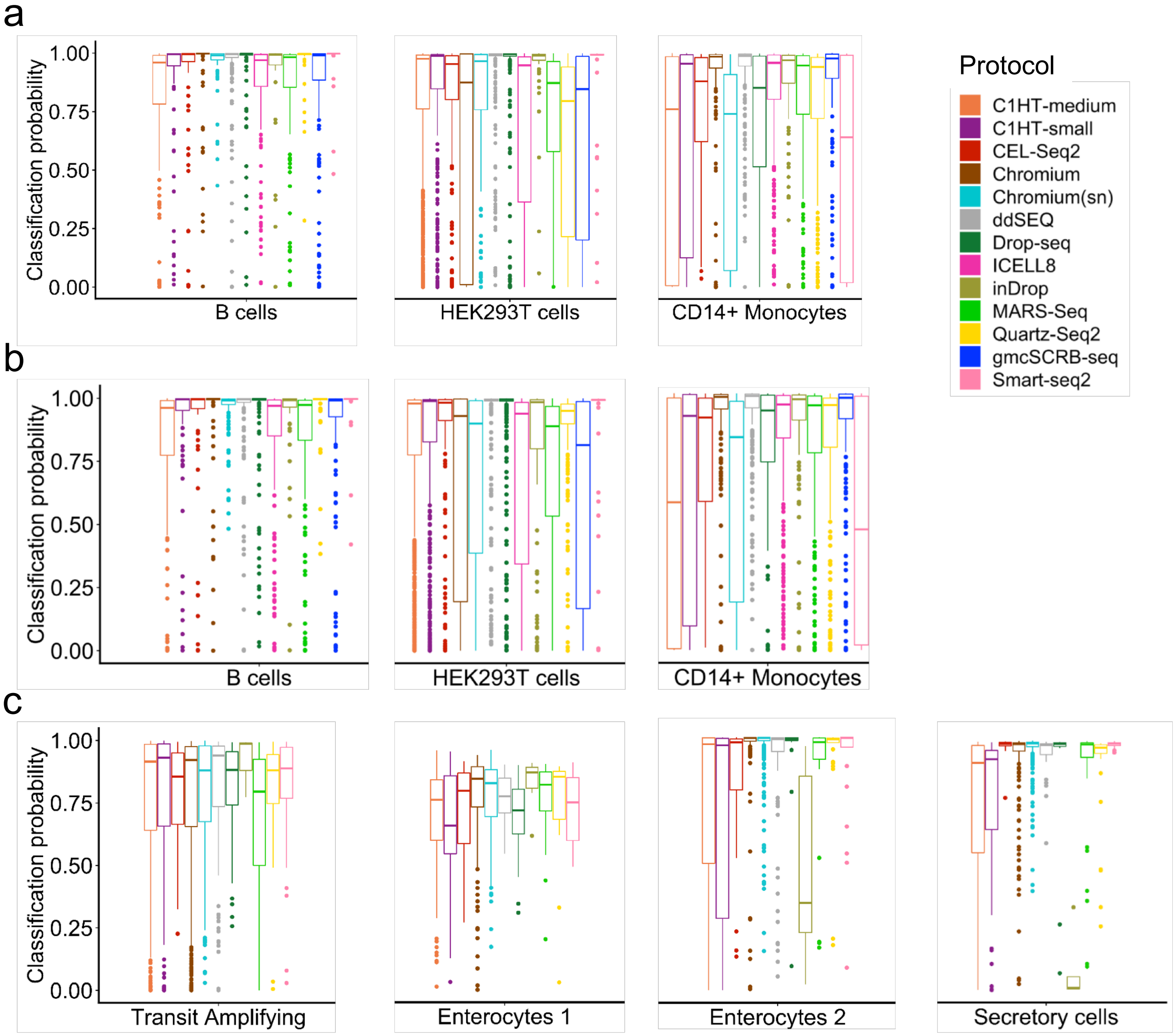
Supplementary Figure 23



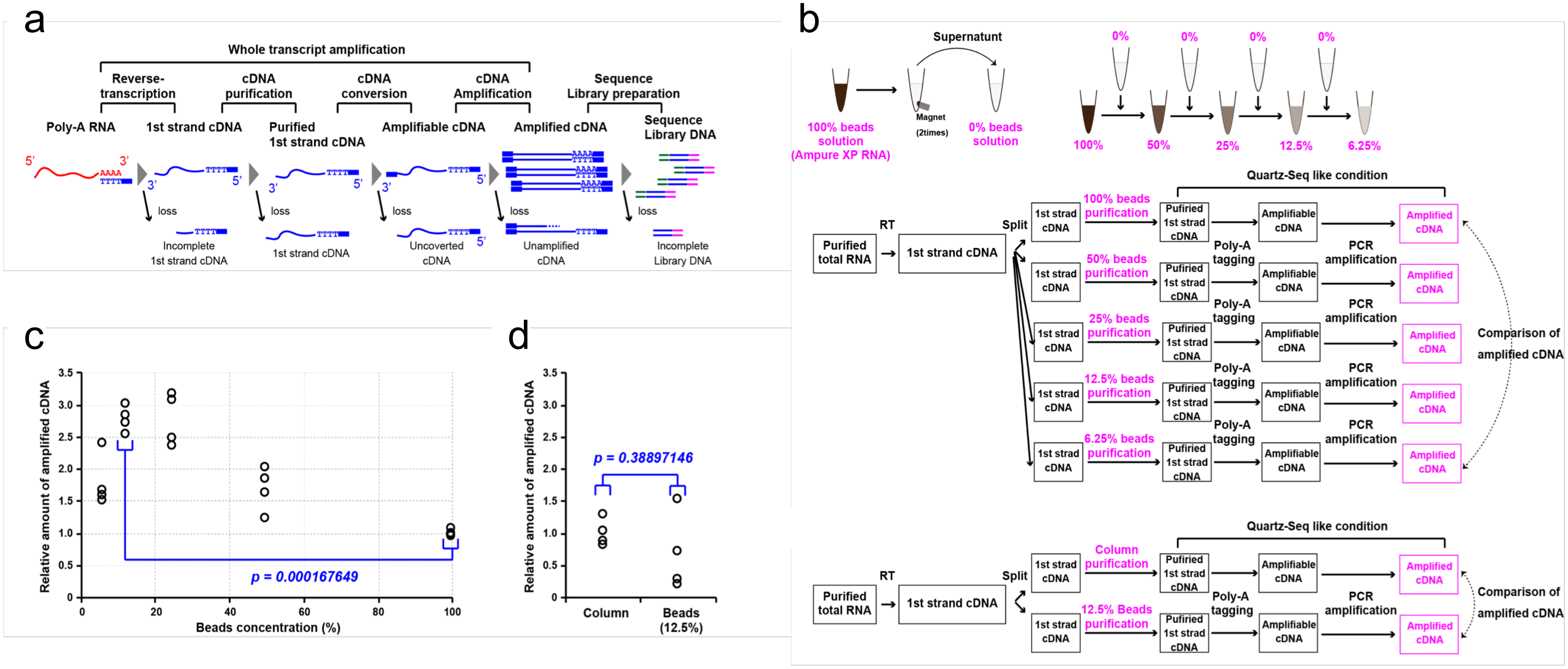
Supplementary Figure 24



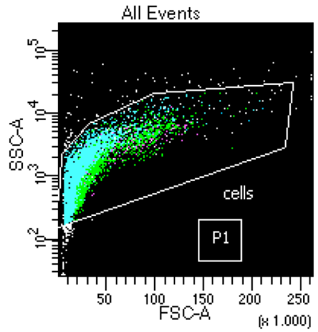
Supplementary Figure 25

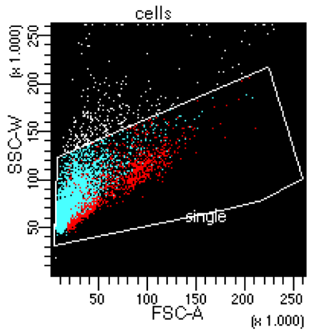


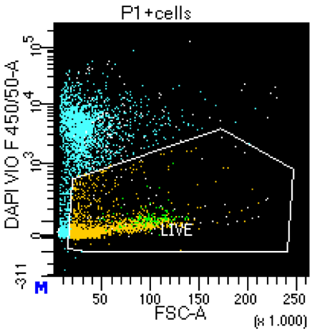
Supplementary Figure 26

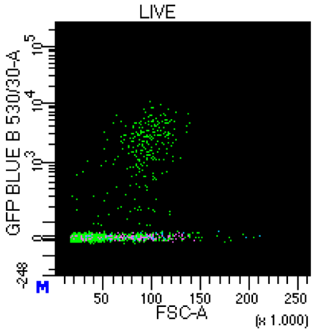


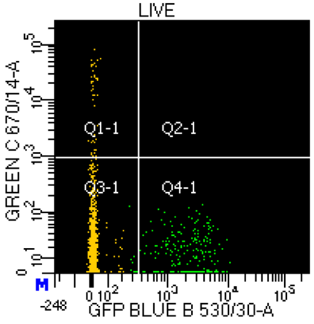
Supplementary Figure 27

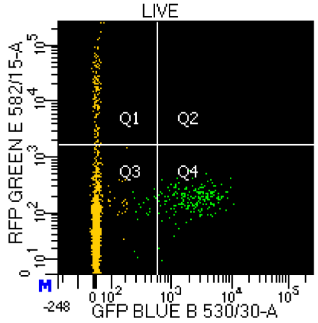


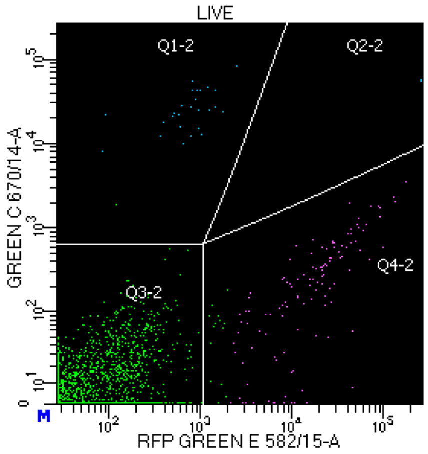


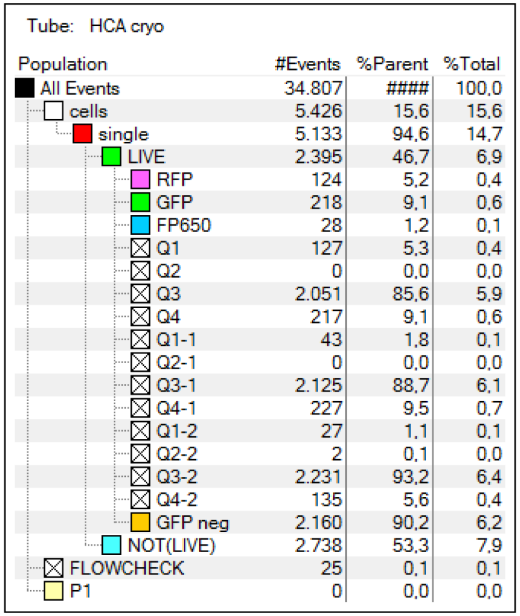








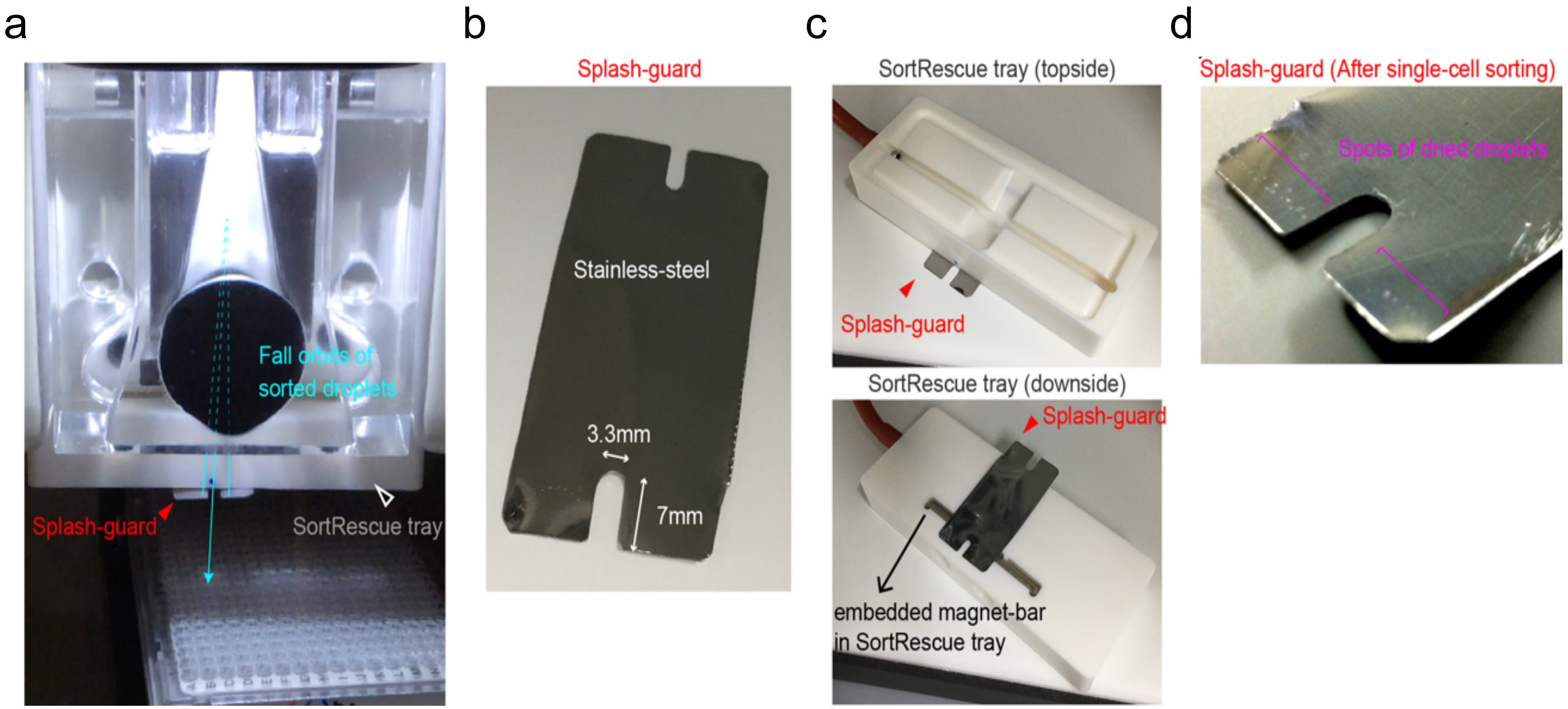




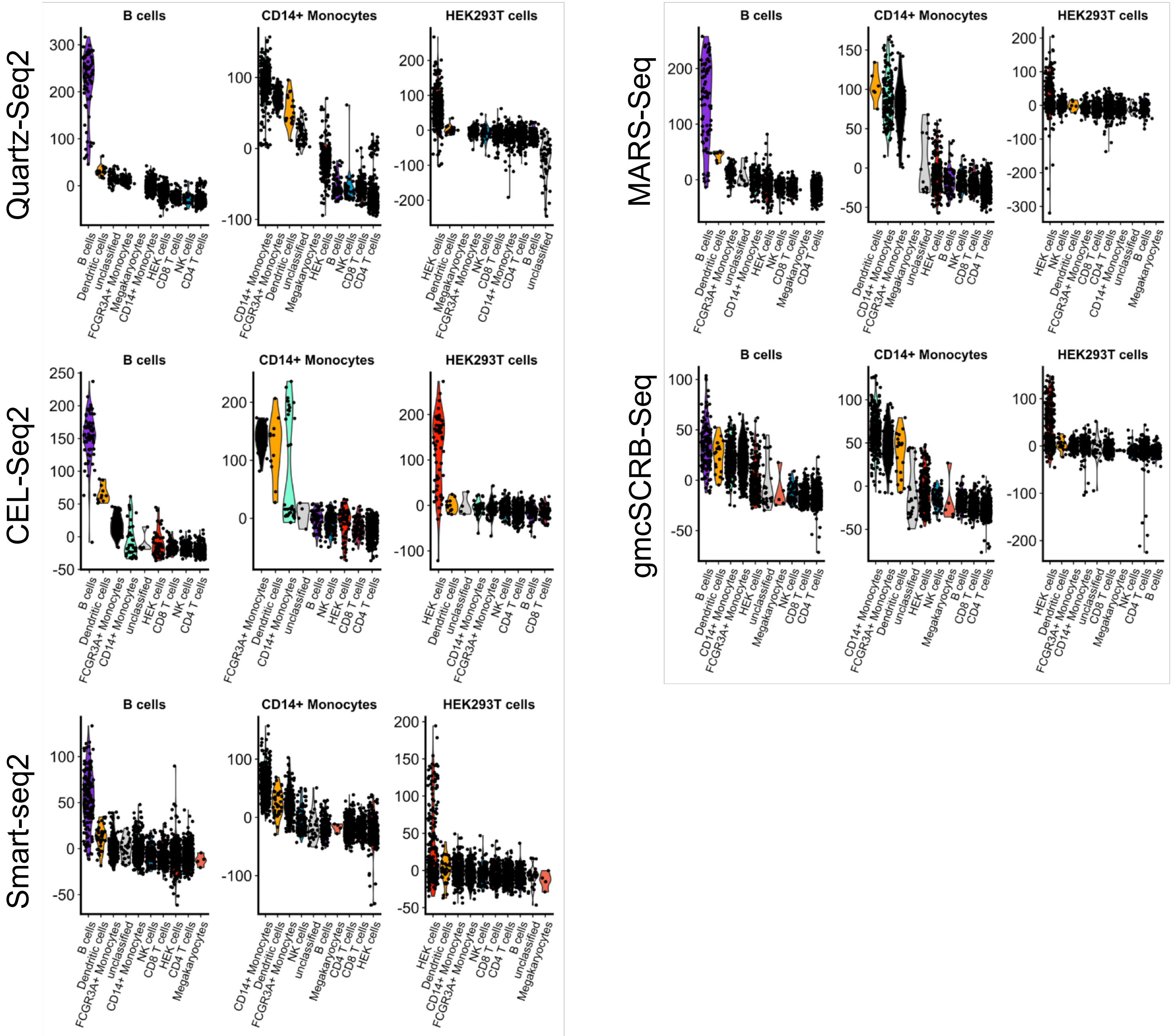
Tube: HCA cryo

Population	#Events	%Parent	%Total
⬛ All Events	34.807	####	100,0
☐ cells	5.426	15,6	15,6
🟥 single	5.133	94,6	14,7
🟩 LIVE	2.395	46,7	6,9
🟪 RFP	124	5,2	0,4
🟩 GFP	218	9,1	0,6
🟦 FP650	28	1,2	0,1
☒ Q1	127	5,3	0,4
☒ Q2	0	0,0	0,0
☒ Q3	2.051	85,6	5,9
☒ Q4	217	9,1	0,6
☒ Q1-1	43	1,8	0,1
☒ Q2-1	0	0,0	0,0
☒ Q3-1	2.125	88,7	6,1
☒ Q4-1	227	9,5	0,7
☒ Q1-2	27	1,1	0,1
☒ Q2-2	2	0,1	0,0
☒ Q3-2	2.231	93,2	6,4
☒ Q4-2	135	5,6	0,4
🟨 GFP neg	2.160	90,2	6,2
🟦 NOT(LIVE)	2.738	53,3	7,9
☒ FLOWCHECK	25	0,1	0,1
🟨 P1	0	0,0	0,0

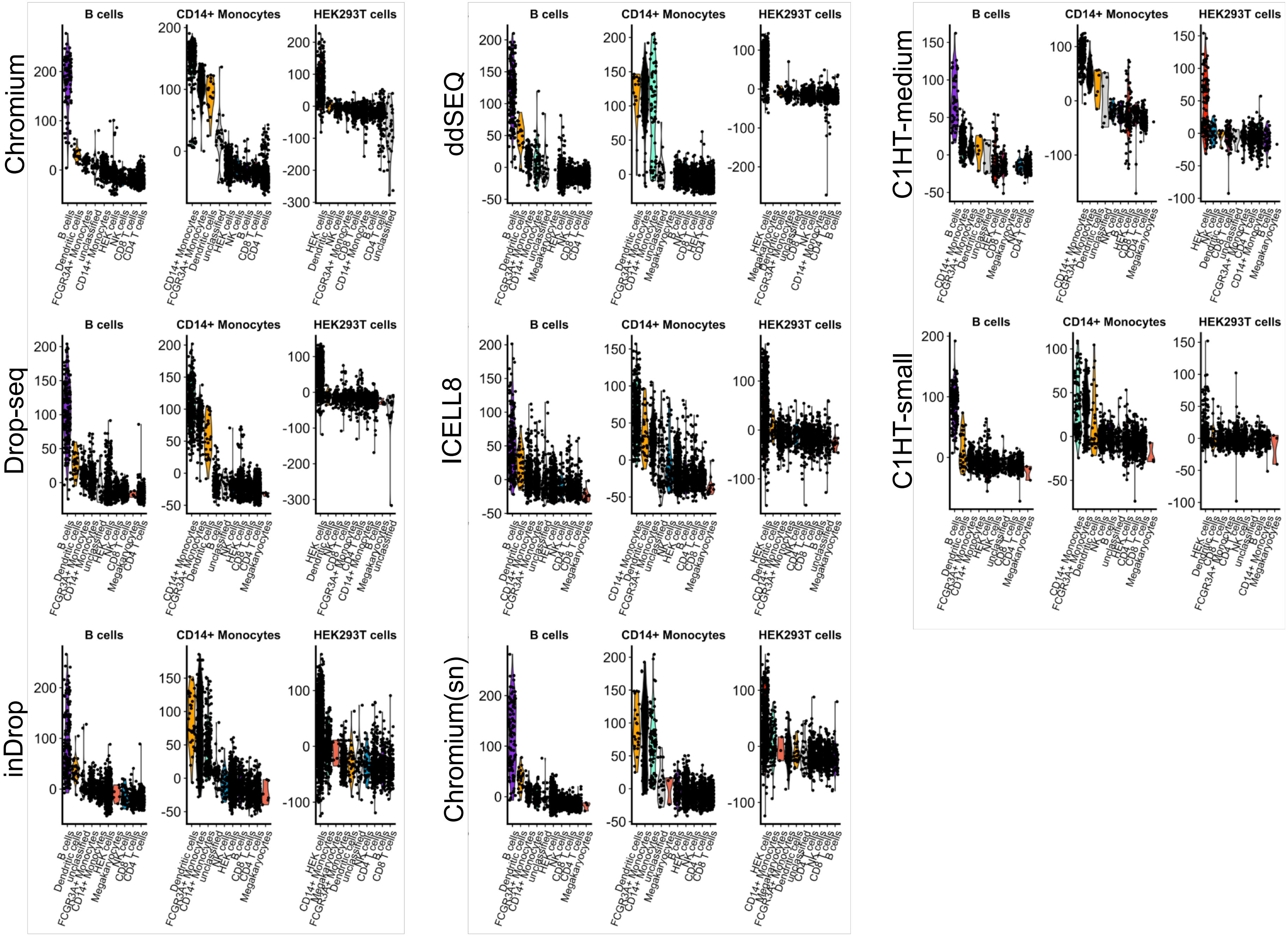
Supplementary Figure 28



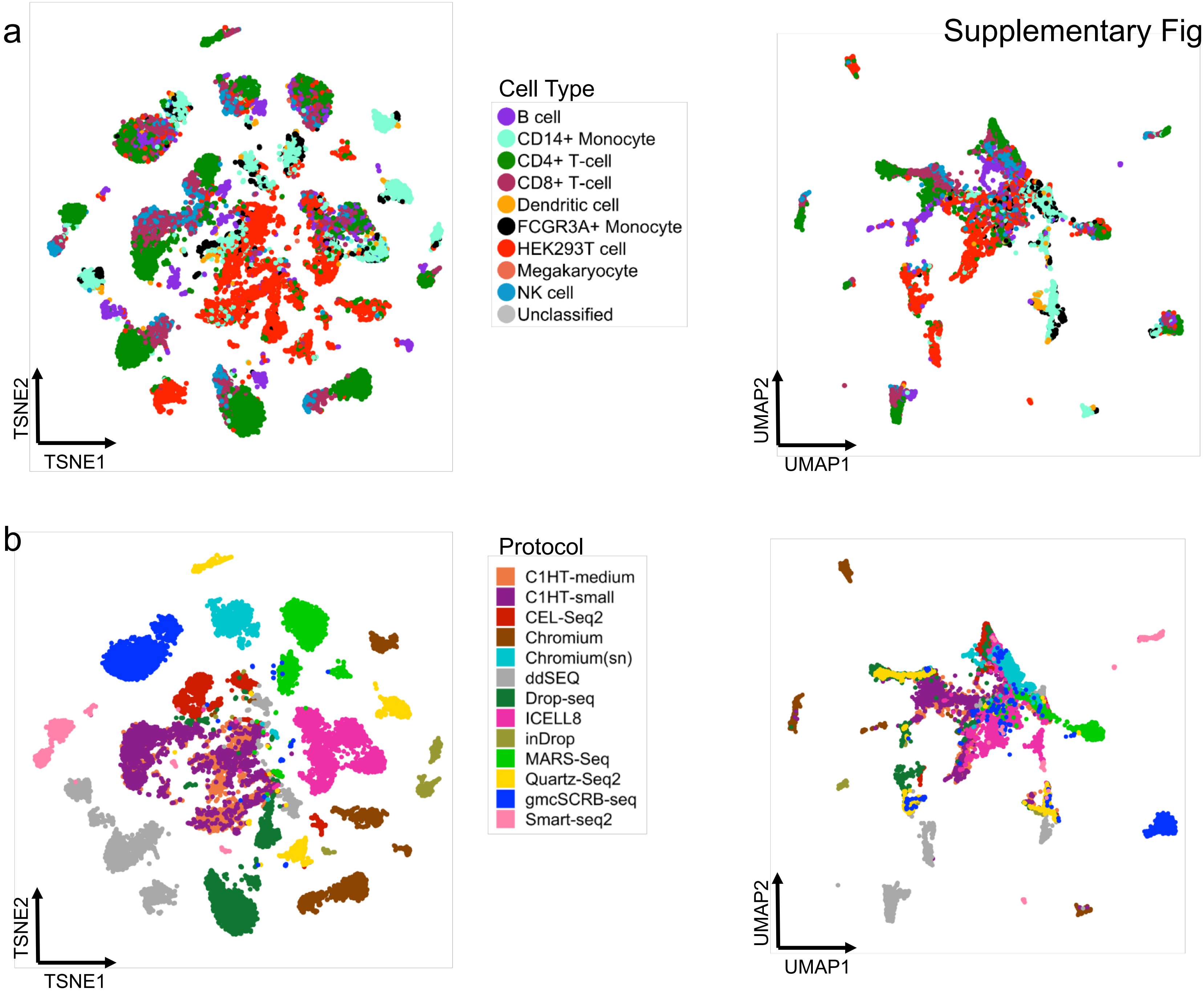
Supplementary Figure 29



Supplementary Figure 30



Supplementary Figure 31



Supplementary Figure 32