

UCSF

UC San Francisco Previously Published Works

Title

Cyclone: an accessible pipeline to analyze, evaluate, and optimize multiparametric cytometry data

Permalink

<https://escholarship.org/uc/item/3b98t3fg>

Authors

Patel, Ravi K

Jaszczak, Rebecca G

Im, Kwok

et al.

Publication Date

2023

DOI

10.3389/fimmu.2023.1167241

Peer reviewed



## OPEN ACCESS

## EDITED BY

Thien-Phong Vu Manh,  
INSERM U1104 Centre d'immunologie de  
Marseille-Luminy (CIML), France

## REVIEWED BY

Cyrille Mionnet,  
INSERM U1104 Centre d'immunologie de  
Marseille-Luminy (CIML), France  
Aurélien Corneau,  
Sorbonne Universités, France

## \*CORRESPONDENCE

Alexis J. Combes

✉ alexis.combes@ucsf.edu

Gabriela K. Fragiadakis

✉ gabriela.fragiadakis@ucsf.edu

†These authors have contributed equally to  
this work

‡Lead authors

RECEIVED 16 February 2023

ACCEPTED 04 August 2023

PUBLISHED 04 September 2023

## CITATION

Patel RK, Jaszczak RG, Im K, Carey ND,  
Courau T, Bunis DG, Samad B,  
Avanesyan L, Chew NW, Stenske S,  
Jespersen JM, Publicover J, Edwards AW,  
Naser M, Rao AA, Lupin-Jimenez L,  
Krummel MF, Cooper S, Baron JL,  
Combes AJ and Fragiadakis GK (2023)  
Cyclone: an accessible pipeline to analyze,  
evaluate, and optimize multiparametric  
cytometry data.  
*Front. Immunol.* 14:1167241.  
doi: 10.3389/fimmu.2023.1167241

## COPYRIGHT

© 2023 Patel, Jaszczak, Im, Carey, Courau,  
Bunis, Samad, Avanesyan, Chew, Stenske,  
Jespersen, Publicover, Edwards, Naser, Rao,  
Lupin-Jimenez, Krummel, Cooper, Baron,  
Combes and Fragiadakis. This is an open-  
access article distributed under the terms of  
the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Cyclone: an accessible pipeline to analyze, evaluate, and optimize multiparametric cytometry data

Ravi K. Patel<sup>1†</sup>, Rebecca G. Jaszczak<sup>1†</sup>, Kwok Im<sup>1,2,3</sup>,  
Nicholas D. Carey<sup>1,4,5</sup>, Tristan Courau<sup>1,2,3,6</sup>, Daniel G. Bunis<sup>1</sup>,  
Bushra Samad<sup>1</sup>, Lia Avanesyan<sup>4,5,7,8</sup>, Nayvin W. Chew<sup>1,2,3</sup>,  
Sarah Stenske<sup>4,5</sup>, Jillian M. Jespersen<sup>4,5</sup>, Jean Publicover<sup>4,5</sup>,  
Austin W. Edwards<sup>1</sup>, Mohammad Naser<sup>1</sup>, Arjun A. Rao<sup>1</sup>,  
Leonard Lupin-Jimenez<sup>1</sup>, Matthew F. Krummel<sup>2,3,6</sup>,  
Stewart Cooper<sup>5,7,8</sup>, Jody L. Baron<sup>4,5,7</sup>, Alexis J. Combes<sup>1,2,3,4,6\*‡</sup>  
and Gabriela K. Fragiadakis<sup>1,9\*‡</sup>

<sup>1</sup>UCSF CoLabs, University of California San Francisco, San Francisco, CA, United States, <sup>2</sup>Department of Pathology, University of California San Francisco, San Francisco, CA, United States, <sup>3</sup>ImmunoX Initiative, University of California San Francisco, San Francisco, CA, United States, <sup>4</sup>Department of Medicine, Division of Gastroenterology, University of California San Francisco, San Francisco, CA, United States, <sup>5</sup>UCSF Liver Center, University of California San Francisco, San Francisco, CA, United States,

<sup>6</sup>UCSF Immunoprofiler Initiative, University of California San Francisco, San Francisco, CA, United States,

<sup>7</sup>The Ibrahim El-Hefni Liver Biorepository at California Pacific Medical Center (IELBC), San Francisco, CA, United States, <sup>8</sup>Division of General and Transplant Hepatology, California Pacific Medical Center & Research Institute, San Francisco, CA, United States, <sup>9</sup>Division of Rheumatology, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

In the past decade, high-dimensional single-cell technologies have revolutionized basic and translational immunology research and are now a key element of the toolbox used by scientists to study the immune system. However, analysis of the data generated by these approaches often requires clustering algorithms and dimensionality reduction representation, which are computationally intense and difficult to evaluate and optimize. Here, we present Cytometry Clustering Optimization and Evaluation (Cyclone), an analysis pipeline integrating dimensionality reduction, clustering, evaluation, and optimization of clustering resolution, and downstream visualization tools facilitating the analysis of a wide range of cytometry data. We benchmarked and validated Cyclone on mass cytometry (CyTOF), full-spectrum fluorescence-based cytometry, and multiplexed immunofluorescence (IF) in a variety of biological contexts, including infectious diseases and cancer. In each instance, Cyclone not only recapitulates gold standard immune cell identification but also enables the unsupervised identification of lymphocytes and mononuclear phagocyte subsets that are associated with distinct biological features. Altogether, the Cyclone pipeline is a versatile and accessible pipeline for performing, optimizing, and evaluating clustering on a variety of cytometry datasets, which will further power immunology research and provide a scaffold for biological discovery.

## KEYWORDS

cyclone, CyTOF, spectral flow cytometry, spatial expression data, multi-parametric analysis, FlowSOM, clustering optimization

## Introduction

The advent of high-dimensional single-cell technologies has transformed our ability to study the complex array of cell types, states, and behaviors comprising the immune system (1, 2). Single-cell proteomics including mass cytometry (CyTOF) and high-dimensional flow cytometry enables the detection of up to 50 extra- and intracellular proteins on hundreds of thousands of cells from a sample (3, 4). These technologies have been applied to a wide variety of patient cohorts and animal models to gain insights into immune set points, responses, and pathology including cancer, infection, hyperinflammatory disorders, and therapeutic intervention (5–8).

The high dimensionality of these data has necessitated the development and application of tools that can parse these data in a semi-automated way. This includes identifying cell populations via clustering algorithms (9–11), dimensionality reduction approaches for stratifying samples (12, 13), and visualization software and statistical packages for downstream analysis (14–16). Due to the diversity and complexity of the immune system, the use of clustering algorithms and dimensionality reduction has become increasingly standard for immune monitoring across tissues and species. However, to date, a consensus process has not been established, rendering the comparison of this type of analysis difficult. While many tools have been both introduced and evaluated, many researchers such as wet-lab immunologists with limited computational experience struggle to navigate the vast landscape of tools for processing and analyzing cytometry datasets. Challenges in algorithm selection, run accessibility and scalability, and chaining the tools for each stage of analysis (preprocessing (17, 18), batch correction, clustering, and downstream analysis) pose significant barriers in the analysis of cytometry data. Moreover, many of these clustering algorithms require selecting a clustering resolution (i.e., selecting the number of clusters), which is largely arbitrary and may reduce the unsupervised nature of these methods. Therefore, an integrated cluster evaluation (19, 20) step is needed to compare different resolutions, guide clustering optimization, and facilitate a tool's usage.

Here, we present the CYtometry CLustering Optimization and Evaluation (Cyclone) ([github.com/UCSF-DSCOLAB/cyclone/](https://github.com/UCSF-DSCOLAB/cyclone/)) pipeline for the analysis of a wide range of cytometry data, including but not limited to CyTOF, fluorescence-based cytometry, and multiplexed immunofluorescence (IF). Cyclone clusters data using FlowSOM (9)—selected based on scalability and fidelity to manual gating—and allows users to optimize and evaluate cluster resolution based on both stability and user exploration. We present Cyclone's performance on CyTOF datasets as well as other single-cell technologies including spectral flow cytometry and imaging. We designed Cyclone to be interoperable with outputs of the leading batch correction algorithms (21, 22) and to feed into accessible downstream analysis tools (15, 16). We additionally accommodate both large datasets as well as showcase Cyclone's performance on downsampled data, increasing its accessibility to those with either extensive or limited computational resources. Leveraging Cyclone

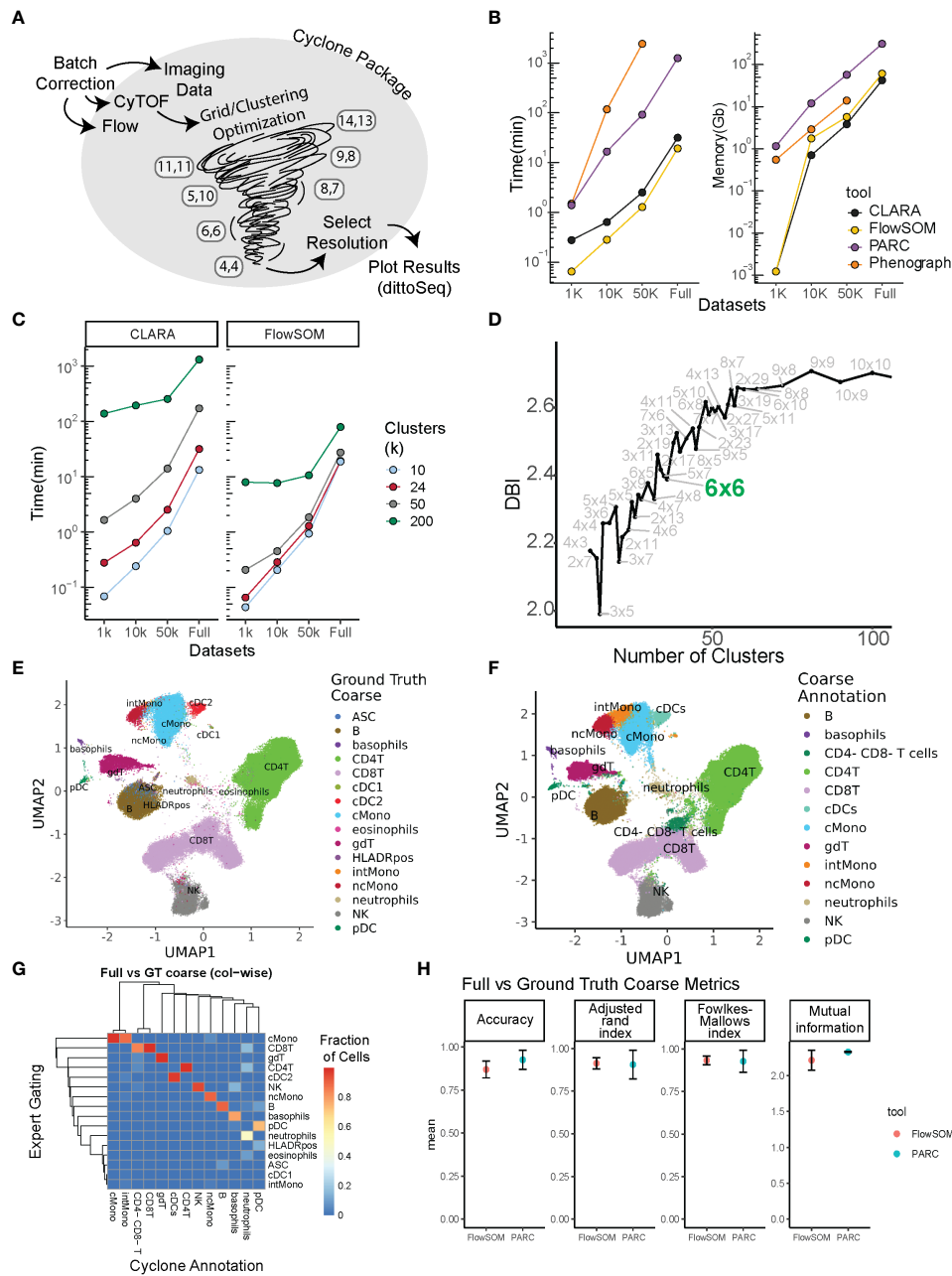
on a seven-color multiplexed immunofluorescence dataset obtained from human colorectal and kidney tumors, we identified a distinct tumor-associated classical dendritic cell subset. We share the R-based pipeline publicly along with extensive documentation for its use in conjunction with upstream and downstream tools. With this largely “plug and play” pipeline, we hope to lower the barrier to entry into high-dimensional cytometry analysis and provide a consensus tool to the research community.

## Results

### Building a scalable pipeline for analysis of high-dimensional cytometry data

To create a functional and accessible workflow for cytometry datasets, we were interested in building a pipeline optimized to meet criteria that address challenges in analyses of this nature (Figure 1A). This included 1) a framework optimized for interoperability, versatility, and ease of use, 2) designed to receive standard inputs from pre-processing steps including batch correction and 3) provide a set of outputs that can easily serve as inputs to downstream analysis and visualization tools. The pipeline required both the selection of a scalable and accurate clustering algorithm robust to downsampling and the ability to tune and evaluate clustering resolution to best meet the specifics of the biological system and scientific inquiry. Due to the wealth of existing single-cell technologies and the need for multiple measurement types to fully understand complex biological systems, we were additionally interested in developing a pipeline that would accommodate a range of single-cell cytometric data modalities including imaging.

A critical component of the pipeline was to select a clustering algorithm that could 1) scale to a large number of cells and parameters while maintaining reasonable runtimes and memory usage and 2) recapitulate populations identified via expert manual gating. We therefore evaluated a subset of available clustering algorithms, focusing on popular algorithms shown to perform well in the literature (23, 24). These included PhenoGraph (11), a graph-based community detection algorithm; CLARA (25), an extension of the partitioning around medoids algorithm; FlowSOM (9), which uses self-organizing map clustering; and PARC (10), a recent combinatorial graph-based algorithm optimized for scalability. We applied these algorithms to a CyTOF dataset measuring the expression of 42 proteins at the single-cell level on peripheral blood mononuclear cells (PBMCs) from 17 individuals (Methods, Supplemental Table 1). We also subset the dataset to various sizes ranging from 1,000 to 10,000 to 50,000 cells per individual to evaluate both speed and memory usage on a range of dataset sizes (Figure 1B). While memory usage was similar across algorithms (except for PhenoGraph performance on the full dataset, which did not conclude in a reasonable amount of time), FlowSOM and CLARA utilized the least runtime across dataset sizes. We therefore proceeded to evaluate those two clustering algorithms on the full dataset. Notably, increasing the requested cluster number identified by the algorithm across the



Patel RK, Jaszczak RG, et al

FIGURE 1

Building a scalable pipeline and optimizing clustering for analysis of high-dimensional cytometry data. (A) Graphical depiction of Cyclone pipeline. Cyclone can intake multiple types of cytometry data, including Flow, CyTOF, and imaging data, and works for both raw and batch-corrected data. FlowSOM clustering includes a grid/clustering optimization step, where a set list of grids is calculated. After the user selects a desired resolution, Cyclone generates useful plots as well as processed data frames that can be easily handed off to other analysis tools and plotting packages, like dittoSeq or ggplot2. (B) Clustering tool performance on data through the lens of scalability (runtime and memory required) of tested tools. CLARA and FlowSOM were similar in their time and memory requirements, while PARC and PhenoGraph have less feasible runtime or memory requirements. (C) Clustering time increases as cluster number (k) increases. FlowSOM still performs better than CLARA when measuring time to cluster, regardless of cluster number requested. (D) Optimization of clustering via evaluation of the different resolutions, leveraging Davis–Bouldin index as indicator (subset of full number of grids assessed; full amount of grids evaluated in the supplement). (E) Ground Truth expert “coarse”-level annotation identifying broad cell types based on manual gating. (F) FlowSOM cluster annotation at “coarse”-level based on CyTOF panel expression. (G) Heatmap comparing “coarse”-level annotations assignment based on the “ground truth” manual gating (rows) of full dataset vs. FlowSOM clustering annotated by expert immunologist based on Cyclone cluster heatmap outputs (columns). (H) Comparison metrics based on “coarse”-level annotations from two expert immunologists. Various performance metrics were used to assess the accuracy of clusters called in the FlowSOM clustering compared to ground truth.

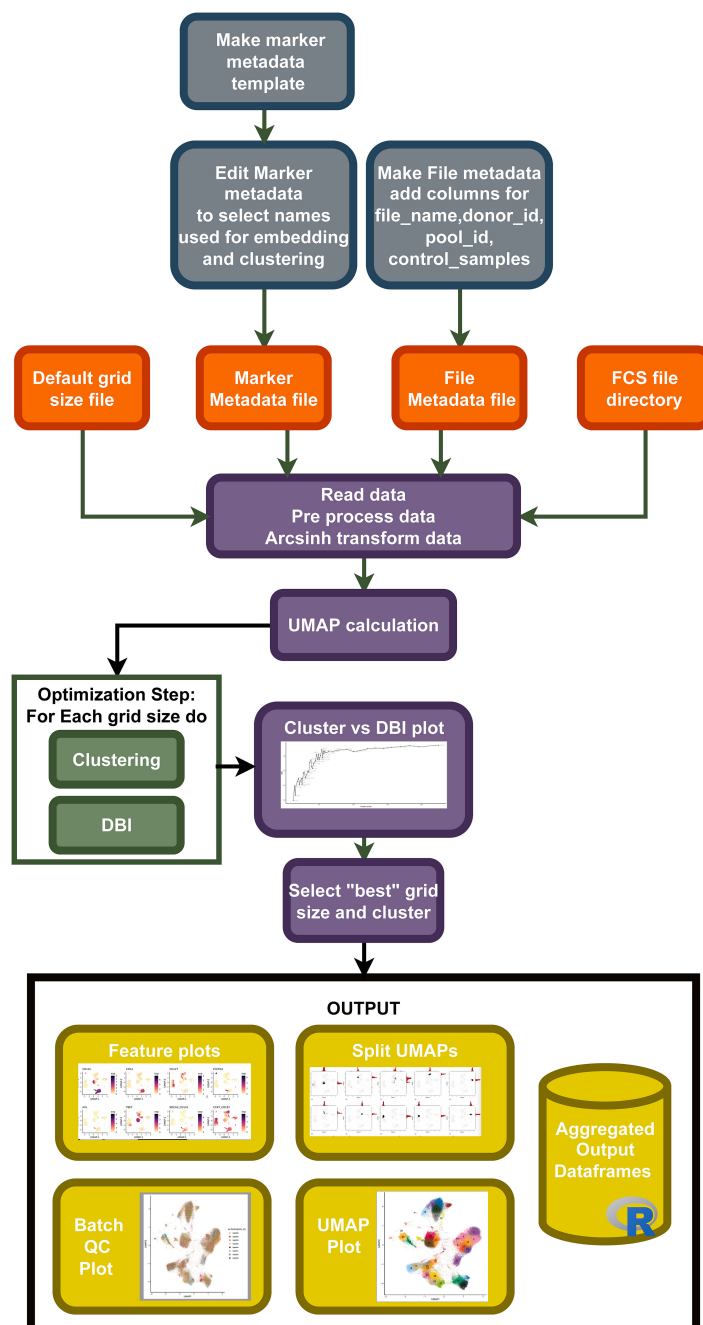
variously sized datasets strongly increased runtime for CLARA as compared to FlowSOM (Figure 1C), which appeared more time efficient even with larger cluster counts regardless of cell numbers. This indicated that CLARA takes more time than FlowSOM to perform a high-resolution clustering to identify cell subsets.

While our final pipeline does focus on R-based implementation for clustering (FlowSOM and CLARA) and permits maximum interoperability with other R-based normalization, debarcoding, and batch correction packages, early investigations explored annotations from both R and python packages; thus, we benchmarked FlowSOM and PARC cluster annotations against manual gating (Supplemental Figure 1) to evaluate their performance in identifying cell populations. To compare the accuracy of annotating clusters based on CyTOF panel expression to gating ground truth, we developed a custom cell barcode scheme (see Methods—FCS modifications) to identify cells in the FCS and facilitate comparison of a cell's ground truth gating-based assignment versus how the cell was annotated by two expert immunologists. To select a resolution/grid for annotation and evaluation, we calculated the Davies–Bouldin index (DBI) (19), a within-dataset similarity metric used to evaluate cluster resolution, across a variety of cluster grid sizes (Figure 1D, Supplemental Figure 2A). Thirty-six clusters ( $6 \times 6$  grid, FlowSOM) or 37 clusters (Resolution 1.3, PARC) were selected for independent annotation by two expert immunologists to compare to “ground truth” manual gates from a third expert immunologist. Both ground truth gating and cluster annotations were performed for major immune cell populations (“coarse-level”, e.g., all  $CD4^+$  T-cell subsets including memory or naïve are annotated as  $CD4^+$  T) (Figures 1E, F) and for more fine-grained sub-populations (“fine-level”, Supplemental Figures 2B–E); these annotations were then visualized in UMAP space. Both FlowSOM and PARC coarse annotations performed well in the recapitulation of manual ground truth gating as captured by four evaluation metrics, including accuracy, adjusted Rand index (26), Fowlkes–Mallows index (27), and mutual information (28) (Figures 1G, H). As expected, while performance was not as strong on fine annotations, metrics showed reasonably accurate clustering for both clustering algorithms (Supplemental Figure 2C). Some sources of error included the fact that global clustering did not isolate small subsets including cDC1s or antibody-secreting cells (ASCs) as their own cluster at the selected resolution due to its low abundance. In addition, the accuracy of some subsets of  $CD4^+$  and  $CD8^+$  T cells varied due to their definition based on markers that have a continuum of expression rather than clear positive and negative expression (e.g., CD45RA). However, the unsupervised nature of the clustering facilitated the identification of cell subsets not included in our manual gating, such as  $CD4^-CD8^-$  T cells and intermediate monocytes (Figure 1G), and B-cell subsets based on the expression of CXCR5 expression or CD38 (Supplemental 2B). Notably, while the majority of clusters were easily recognizable and annotated, a small number of clusters were observed to be dispersed across the UMAP visualization (Supplemental Figure 2F), rather than having a more homogenous phenotype (Supplemental

Figure 2G), which illuminated the utility of visualizing each cluster's dispersion as an output for a developed pipeline. Taken together, these data demonstrate that while PARC and FlowSOM perform similarly compared to expert gating and have similar performance on low numbers of cells, FlowSOM outperforms when considering scalability to higher numbers of cells.

## The Cyclone pipeline is a method for data clustering and evaluation across resolutions for use in the analysis of cytometry data

Based on our observations regarding runtime, memory usage, cluster evaluation methods, and performance on manual gating recapitulation, we developed the Cyclone pipeline in R (Figure 2). We designed Cyclone to produce a series of outputs that users can access to select cluster resolution, evaluate cluster quality, annotate clusters, and utilize resulting cluster statistics (phenotypes and abundance) for downstream analyses. To start Cyclone, the user provides FCS or matrix data files as well as files to specify markers and file metadata. Cyclone expects any normalization and batch correction steps to be performed prior to the use of the pipeline. For CyTOF data, this can be performed by established R packages, including *premassa* ([github.com/ParkerICI/premassa](https://github.com/ParkerICI/premassa)) for bead-based normalization and *CytoNorm* (21) or *cyCombine* (22) for batch correction. After reading in and arcsinh transforming the data, Cyclone calculates UMAP dimensionality reduction. Cyclone works with either FlowSOM or CLARA for clustering; selecting FlowSOM enables DBI-based cluster resolution optimization prior to user grid selection, while with CLARA, the user selects a single resolution for clustering. If FlowSOM is selected, Cyclone then performs iterative optimization of clustering across a variety of cluster grid sizes, which can then be compared using DBI. After the user selection of desired grid, Cyclone performs clustering, generates summary matrices, proceeds through an optional SCAFFoLD step, and then generates output files and visualizations. While not providing batch correction, Cyclone does provide a means for assessing batch or any other input file metadata via UMAP (split by batch) visualization (Figure 3A) as well as clustered heatmaps of cluster frequency with batch or other metadata information overlaid as rugs (Figure 3B). In addition to a DBI plot for cluster resolution selection, Cyclone outputs UMAPs colored by cluster (Figure 3C) as well as heatmaps of marker expression per cluster for ease of cluster annotation (Figure 3D). Additional UMAPs and histogram plots are provided per cluster, showing the cluster's distribution across the UMAP for an evaluation of cluster dispersion as an indication of cluster quality (Figure 3E). Feature UMAP plots of marker expression in all cells are also exported by default (Figure 3F). Taken together, the Cyclone pipeline provides a means of clustering, as well as evaluating and annotating these clusters, to then be readily used in downstream analyses.



Patel RK, Jaszczak RG, et al.

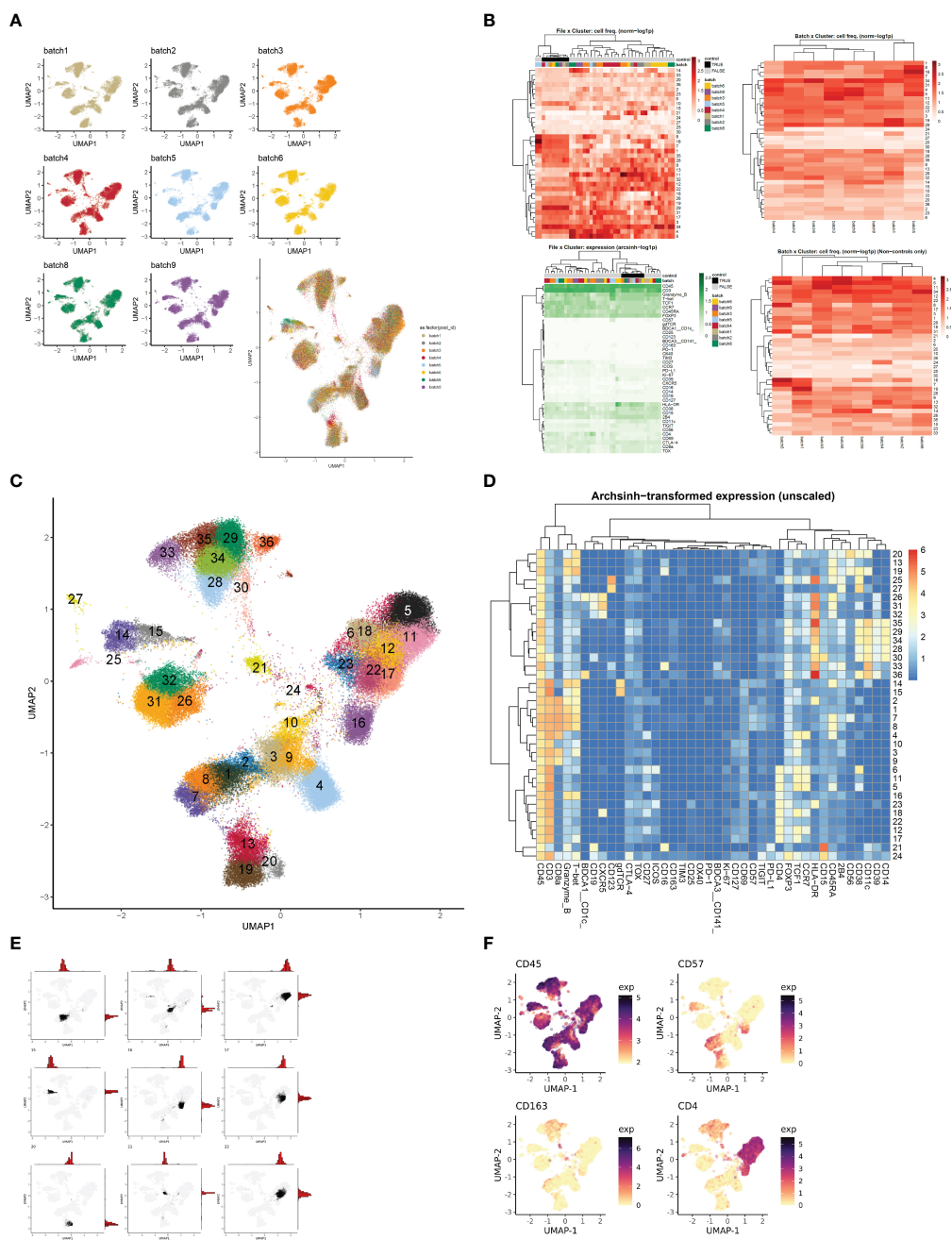
FIGURE 2

Integration of different pipeline pieces and packages into a single method: “Cyclone”. Using information from initial metadata files (see [Supplemental Tables](#)), FCS files are read into the pipeline. After arcsinh transformation, UMAP is calculated and clustering is performed on a default range of grid sizes, resulting in a cluster VS DBI plot for clustering grid selection. After user input to select a specific grid, Cyclone generates feature plots of each antibody included in the panel, split UMAPs to assess cluster dispersion, various QC heatmaps and UMAPs, and a final UMAP annotated by cluster number. Additionally, expression matrices and cell metadata matrices are saved.

### Cyclone provides accessibility and interoperability with upstream processing and downstream analyses

We developed Cyclone to be easily leveraged by researchers with various backgrounds, especially those with limited coding knowledge and minimal computational resources. We also

prioritized interoperability with both upstream and downstream tools for cytometry data processing and analysis ([Figure 4A](#)). In addition to significant documentation (including vignettes to get users started with the pipeline), we evaluated Cyclone on downsampled datasets to determine whether downsampling to fewer cells could provide an alternative for users to run Cyclone locally rather than needing additional compute resources for



**FIGURE 3** Evaluation and interpretation of default pipeline outputs and results. **(A)** UMAPs of batch information. If batches were a part of the CyTOF run, Cyclone exports UMAP plots split by batch information to assess batch correction or batch effect. **(B)** Heatmaps of file x cluster and batch x cluster depicting cell frequency per cluster, and file x feature depicting arcsinh transformed data. **(C)** UMAP annotated by cluster number based on user-selected grid. **(D)** Heatmap of median archsinh-transformed expression (unscaled) per cluster used to annotate clusters. **(E)** Example plots of UMAP + histogram showcasing cluster density/dispersion. **(F)** Example plots of protein feature expression.

analyses. Thus, we downsampled our evaluation dataset to 50,000 cells per sample and coarsely annotated the resulting optimized clusters (Figure 4B, Supplemental Figures 3A, B). We then compared the accuracy of these coarse downsampled annotations to the coarse annotations on the full dataset and encouragingly found strong concordance between the full and downsampled datasets (Figure 4C). When doing fine annotations (Supplemental Figure 3C), metrics again showed reasonably accurate annotations

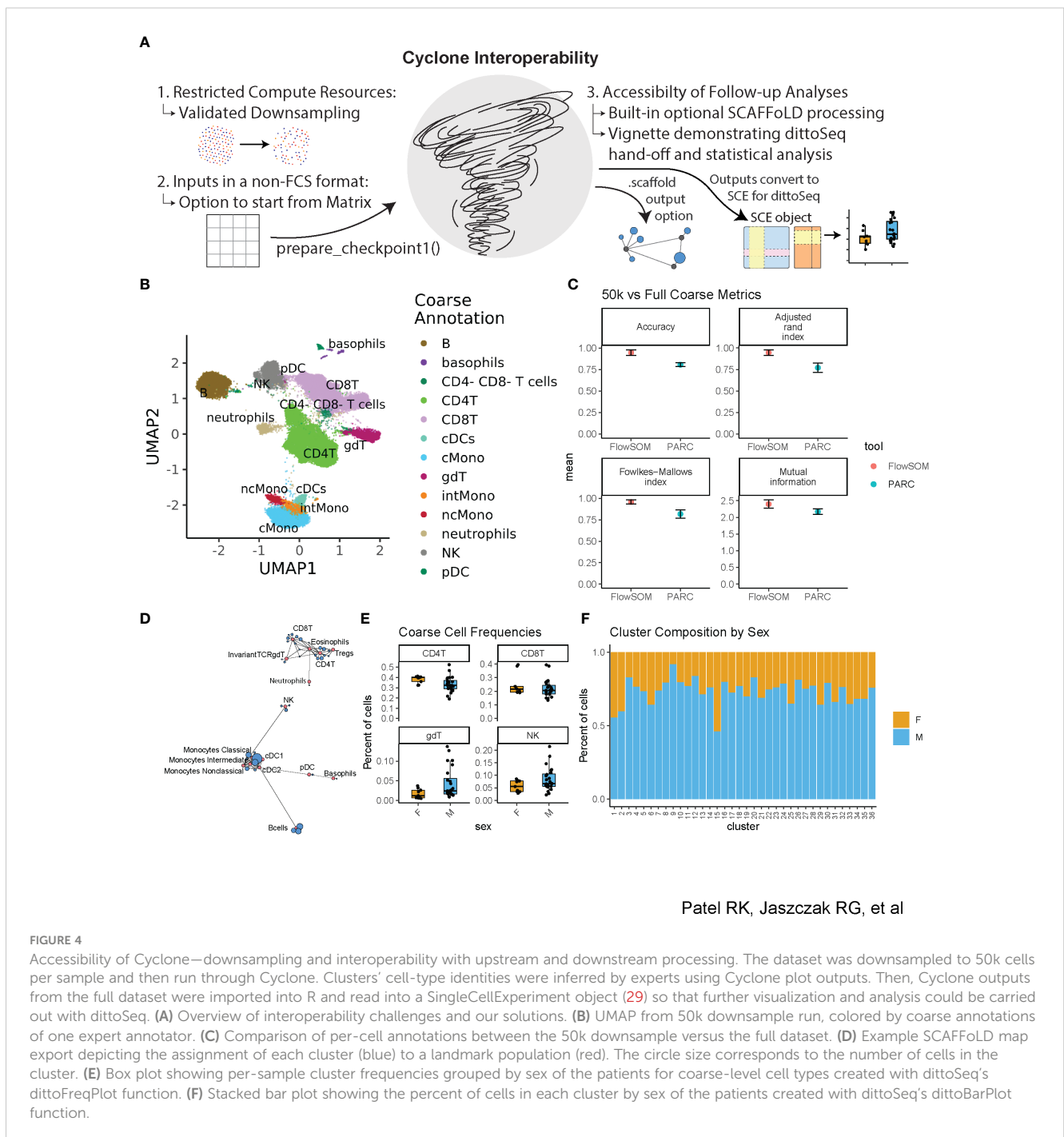
when compared to the full dataset fine annotations (Supplemental Figure 3D). Thus, data downsampling presents a practical option for dataset clustering with Cyclone should computing resources be a challenge for some users.

To optimize Cyclone interoperability with upstream and downstream processing, we considered batch correction as a primary upstream target and SCAFFoLD map analysis and accessible visualization as primary downstream targets. For batch

correction, we ensured a pathway exists to enter the Cyclone workflow with cytometry data in either a matrix format [as output by cyCombine (22)] or adjusted-FCS files [as output by CytoNorm (21)] to accommodate interoperability with outputs from both commonly used batch correction methods. CytoNorm’s adjusted-FCS files can enter into the pipeline as normal, but Cyclone’s prepare\_checkpoint1() provides an entry point for cyCombine’s matrix format. The function takes the same primary inputs as a run without a batch correction step, except for accepting a matrix (with cells in rows and markers in columns) instead of the FCS directory. The function then performs all the same optional steps (arcsinh transformation, control sample

removal, and subsampling) before outputting a Checkpoint1.Rdata file. Afterward, users can run Cyclone as normal to continue from the second step of UMAP calculation.

To aid in cluster annotation, SCAFFoLD connects clusters to unique landmark populations based on the cosine similarity in the feature expression space (15). For powering SCAFFoLD map downstream analysis, we added an optional step directly into the Cyclone pipeline; if landmark population FCS files are provided, an output \*.scaffold file can be used to generate a SCAFFoLD map (15) via the “scaffold::scaffold.run()” command (Figure 4D). For powering accessible visualization and other downstream follow-ups, we ensured that the colors of Cyclone plot outputs use color



Patel RK, Jaszczak RG, et al

FIGURE 4

Accessibility of Cyclone—downsampling and interoperability with upstream and downstream processing. The dataset was downsampled to 50k cells per sample and then run through Cyclone. Clusters’ cell-type identities were inferred by experts using Cyclone plot outputs. Then, Cyclone outputs from the full dataset were imported into R and read into a SingleCellExperiment object (29) so that further visualization and analysis could be carried out with dittoSeq. (A) Overview of interoperability challenges and our solutions. (B) UMAP from 50k downsample run, colored by coarse annotations of one expert annotator. (C) Comparison of per-cell annotations between the 50k downsample versus the full dataset. (D) Example SCAFFoLD map export depicting the assignment of each cluster (blue) to a landmark population (red). The circle size corresponds to the number of cells in the cluster. (E) Box plot showing per-sample cluster frequencies grouped by sex of the patients for coarse-level cell types created with dittoSeq’s dittoFreqPlot function. (F) Stacked bar plot showing the percent of cells in each cluster by sex of the patients created with dittoSeq’s dittoBarPlot function.



blindness-accessible color palettes and that data outputs can be used with dittoSeq—a color blindness-friendly visualization tool (16). Although dittoSeq was designed for single-cell RNAseq data, it proves generalizable to other data modalities including high-dimensional cytometry. To directly power dittoSeq integration and downstream statistical analyses, we provide an additional vignette showing how to 1) transform Cyclone data objects into a SingleCellExperiment [(29), Chapter four] object compatible with dittoSeq, 2) generate useful visualizations, 3) add cell-type annotations for each cluster as well as further annotations such as spatial information in the case of imaging data, and 4) run statistics on differences in cluster or cell-type frequencies between samples. Shown as examples are boxplots of how cluster frequencies compare between male and female subjects (Figure 4E) and the percent composition of each cluster in terms of subject sex (Figure 4F); these functions make probing metadata categories of interest easy to code and assist in producing publication-ready visualization. These accessible (in both the color palate and easily leveraged graphing functions) visualizations are readily created with dittoSeq from Cyclone outputs.

## The Cyclone pipeline generalizes to flow cytometry datasets

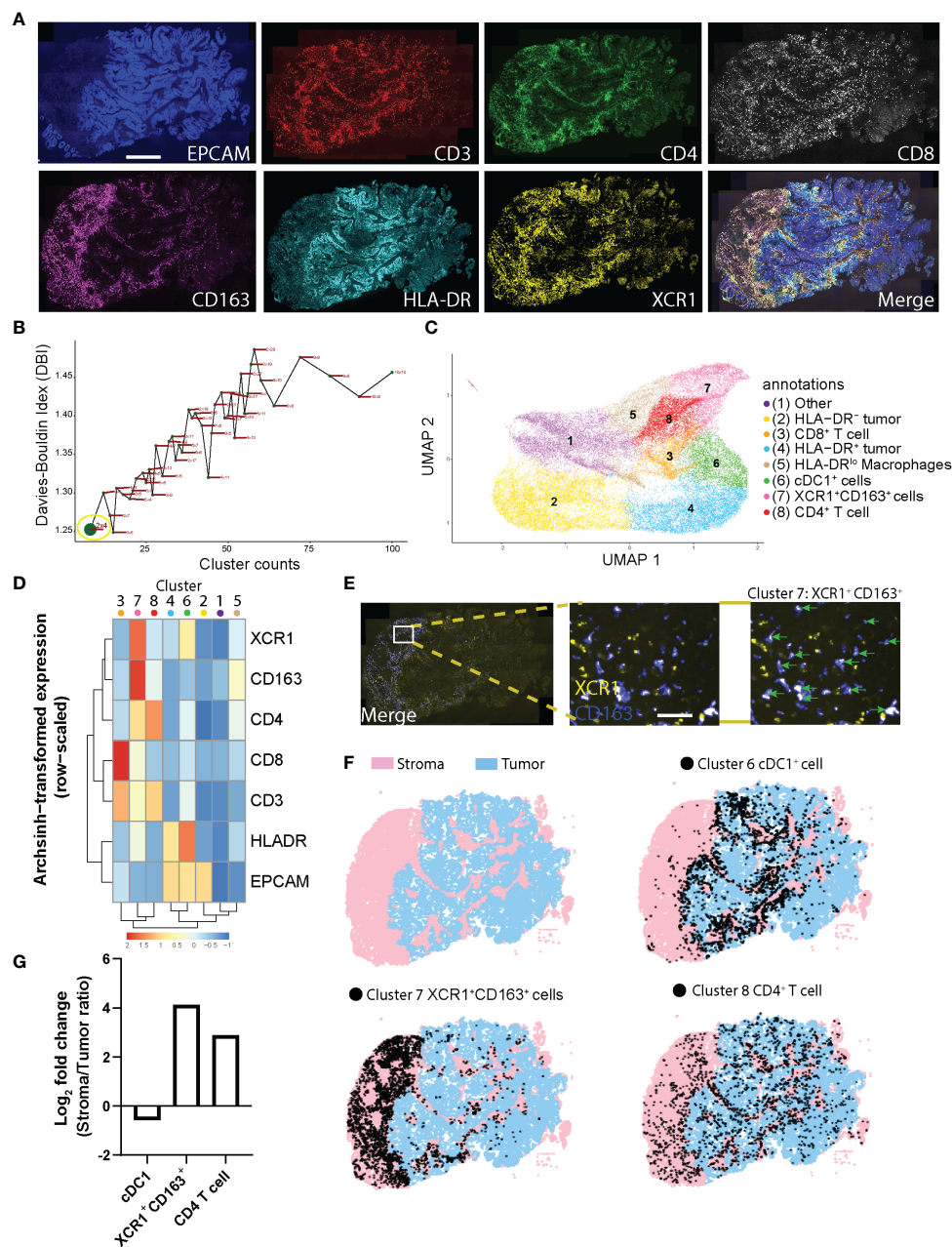
While optimization and development of Cyclone were initially for CyTOF datasets, we realized the need for such analysis techniques on a broader set of similar data modalities, including other cytometry platforms such as spectral flow cytometry. To evaluate Cyclone's utility for this analysis, we applied the Cyclone pipeline to a spectral flow cytometry dataset of mouse liver samples from a transgenic mouse model for hepatitis B virus (HBV) response (30). Liver leukocytes collected from eight mice on day 8, the peak of the immune response in this model, were analyzed using a myeloid-focused spectral flow cytometry panel (Supplemental Table 2). In a repeat experiment, day 8 liver leukocytes from nine mice were analyzed using a lymphoid-focused panel (Supplemental Table 3). The Cyclone pipeline could be used as built, but we used config files to specify the co-factor for the arcsinh transformation to a value more typical of flow cytometry datasets (see Spectral flow data generation and analysis—Cyclone analysis in Methods). For both experiments, CD45<sup>+</sup> live cells were provided to the Cyclone pipeline. After selecting the local minimum DBI, we identified and annotated 21 clusters in the T Cell panel (Figure 5A) and 17 clusters from the Myeloid panel (Figure 5B). Both panels were able to distinguish several T-cell and NK-cell subsets (Supplemental Figure 4A) as well as liver resident macrophages (Kupffer cells) and monocyte-derived macrophages (Supplemental Figure 4B). As was previously observed for CyTOF data, unsupervised clustering using Cyclone largely recapitulated the populations identified by manual expert gating at a coarse-level (Figures 5C, D). Cyclone further enabled the identification of cell subsets across immune cell lineages (Figures 5C, D), which could have been missed by manual gating, such as NK-cell subsets with various expression of KLRG1 or CD62L (Figure 5A, Supplemental Figure 4A). To evaluate the

biological information contained within the unsupervised clustering, we took advantage of the presence of tetramer staining to identify the properties of the HBV-specific T cells. While the tetramer staining was not used as a clustering parameter, we found that the vast majority of CD8<sup>+</sup> (Figure 5E) and CD4<sup>+</sup> (Figure 5F) antigen-specific T cells were enriched in one or two effector T-cell clusters with high expression of activation markers identified by Cyclone for each T-cell subset (Figure 5G). Taken together, we observed that Cyclone performed well on spectral flow data and enabled the unsupervised identification of cell phenotypes that are associated with distinct biological features.

## Cyclone enables unsupervised discovery of cell-type compartmentalization within the tumor microenvironment

Understanding how the phenotype of individual cells relates to the function of multicellular compartments within tissues requires the ability to identify cellular phenotypes with multiple proteins while simultaneously quantifying the spatial distribution and interactions of these cells across large regions of tissue. Along with its applications demonstrated already, Cyclone provides a unique opportunity to analyze the spatial distribution of individual cell phenotypes as well as cellular neighborhoods in an unsupervised manner. To that end, we tested how the Cyclone pipeline compared to a prototypical imaging data analysis pipeline containing image visualization. First, we created a 7-plex immunofluorescent staining panel to be used on a colorectal tumor tissue (CRC1) consisting of a tumor marker (EPCAM), T-cell markers (CD3, CD4, and CD8), and myeloid markers (CD163, HLA-DR, and XCR1) (Figure 6A). Next, we utilized DeepCell (31) segmentation software to demarcate individual cells by inputting nuclei (DAPI<sup>+</sup>) and membrane markers (CD3<sup>+</sup>, CD4<sup>+</sup>, CD8<sup>+</sup>, and CD163<sup>+</sup>) and labeled each cell-type annotation based on rational gating parameters (e.g., CD4<sup>+</sup> T cell = CD3<sup>+</sup>CD4<sup>+</sup>CD8<sup>-</sup>XCR1<sup>-</sup>) (Supplemental Figure 5A, table). We found that determining the manual thresholding on certain markers such as CD163<sup>+</sup> expression was not visually clear in designating a suitable cutoff as compared to CD3<sup>+</sup> expression and that this could ultimately lead to variability in the frequency of cell types annotated (Supplemental Figure 5A, histogram plots). For example, the threshold cutoff at 11 observed CD163<sup>+</sup> expression on low-background, segmented cells, while the threshold cutoff at 13 missed CD163<sup>+</sup> cells. Thus, we opted for a middle ground by choosing threshold 12 for the downstream comparison with our Cyclone pipeline. To run the Cyclone pipeline on this multiplexed immunofluorescence data, we followed the following procedure. Once raw expression values of each marker for every cell were curated and assigned to each cell identified by DeepCell, we obtained a cell per protein expression matrix used to run the Cyclone pipeline (see Methods). After DBI evaluation, we chose a grid size of 2 × 4, which had the lowest DBI value (Figure 6B). This resolution generated eight unique clusters spanning immune and non-immune cells (Figure 6C). These clusters comprised tumor cells expressing different levels of HLA-DR (Clusters 2 and 4), CD4<sup>+</sup> and CD8<sup>+</sup> T cells (Clusters 3 and 8),





Patel RK, Jaszczak RG et al

FIGURE 6

Utilization of Cyclone pipeline for imaging data. (A) Representative immunofluorescence imaging of colorectal tumor biopsy using tumor marker [EPCAM (blue)], T cell markers [CD3 (red), CD4 (green), and CD8 (white)], and myeloid markers [CD163 (purple), HLA-DR (cyan), and XCR1 (yellow) staining]. Scale bar denotes 500 μm. (B) Scatterplot of the Davies–Bouldin index and cluster size over multiple iterations of Louvain clustering and varying parameters using seven markers used in immunofluorescence as in panel (A) Grid size 2 × 4 was chosen for downstream analysis (green dot, yellow circle). (C) UMAP visualization of 45,177 cells from the colorectal tumor biopsy with specific populations annotated based on (D) an arcsinh-transformed expression heatmap of all markers row-scaled. (E) Representative immunofluorescence imaging of tumor biopsy (merge; left) with inset (middle) of XCR1<sup>+</sup> (yellow) and CD163<sup>+</sup> (blue) staining. Cluster 7 from Cyclone pipeline was overlaid and annotated (green arrows) representing XCR1<sup>+</sup>CD163<sup>+</sup> cells. Scale bar denotes 20 μm. (F) Full image representation (left) of stromal (pink) and tumor (blue) regions in tumor tissue biopsy (top left) with cDC1<sup>+</sup> cell Cluster 6 (top right), XCR1<sup>+</sup>CD163<sup>+</sup> cell Cluster 7 (bottom left), and CD4<sup>+</sup> T cell Cluster 8 (bottom right) overlays (black dots). (G) Log<sub>2</sub> fold change bar plot on stroma/tumor ratio of each cluster annotation in CRC1 tumor sample.

confirm that they had both XCR1<sup>+</sup> and CD163<sup>+</sup> expression (Figure 6E). Additionally, these cells had an enriched spatial distribution within the stromal compartment as opposed to cDC1s and other clusters enriched in the tumor compartment

(Figures 6F, G, Supplemental Figure 5B). We next validated that XCR1<sup>+</sup>CD163<sup>+</sup> cell types could be found in other colorectal (CRC2) and kidney (KID1) tumor samples subjected to Cyclone (Supplemental Figure 5C). However, these samples observed

different stromal and tumor enrichments for CD4<sup>+</sup> T cells, cDC1s, and XCR1<sup>+</sup>CD163<sup>+</sup> cells compared to CRC1.

Notably, while conventional manual gating for XCR1<sup>+</sup>CD163<sup>+</sup> cells is possible, the thresholding strategy of comparing independent markers with strict cutoffs likely misses out on what Cyclone identifies when leveraging the entire set of markers to identify cell types (Supplemental Figure 5D). Taken together, we highlight the successful application of the Cyclone pipeline on multiplexed imaging data in tumor tissue to discriminate two different XCR1<sup>+</sup> subsets and their spatial features within the tumor microenvironment.

## Discussion

In this work, we present the Cyclone pipeline—a versatile and accessible pipeline for performing, optimizing, and evaluating clustering on cytometry datasets. The pipeline takes in single-cell measurements, performs high-dimensional clustering, allows the user to select a clustering resolution with guided metrics, and provides outputs for facile cluster annotation and downstream analysis. We confirmed the fidelity of FlowSOM clustering to expert manual gating, as well as its performance on high-dimensional datasets; to date, we have successfully applied the Cyclone pipeline with FlowSOM clustering to a 42-parameter CyTOF dataset of 50 million cells. We have released the pipeline code and documentation with the aim of making it accessible to the greater community, where it has already begun to be applied.

The selection of FlowSOM was based on identifying a clustering algorithm 1) that provided accurate identification of cell populations in clustering as compared to manual gating and 2) that was reasonably scalable to large datasets. In our evaluation, we found that FlowSOM could handle datasets of up to 50 million cells, while also identifying both our coarse and fine manually gated cell populations at a high level of accuracy (Figure 1). Our assessment agrees with the benchmarking literature that also found FlowSOM to capture both abundant and rare cell populations with reasonable runtime (23). Central to this work was to build a tool that would be available to and easy to use by the research community, including wet-lab scientists, such that researchers are better empowered to engage with their cytometry datasets. In our collaborative research model, we brought together the computational rigor from data scientists with the perspectives, challenges, and biological intuition of wet-lab biologists for the design and development of the Cyclone pipeline. This included the consideration of high-performance computational resource access, which is not always available or accessible to wet-lab scientists. We therefore validated Cyclone's performance on downsampled data to offer an option for running the pipeline locally on a personal laptop. This also included addressing the challenges of varied inputs (i.e., starting with varied file formats, either FCS files or a more general matrix format), as well as a computational hand-off to downstream analysis tools, either of which could erect a barrier to researchers less proficient with coding. While commercially available tools offer clustering pipelines such as OMIQ, CellEngine, and Cytobank, it was important to us to offer a freely available pipeline that can be flexible to the computational

resources available to the researcher as well as accommodate a large number of cells. In addition, we pair this clustering with the ability to evaluate clustering resolutions and perform dimensionality reduction to better enable downstream analysis.

Cyclone is readily generalized to a variety of cytometry datasets beyond CyTOF data, demonstrating its versatility as a pipeline for high-dimensional datasets that can be extensible and adaptable as technologies continue to evolve. We demonstrate the application of Cyclone to spectral flow cytometry data, in this case in the setting of a mouse model of viral infection, as well as immunofluorescence imaging data of the tumor microenvironment. Excitingly, Cyclone has been further applied to a CO-Detection by indEXing (CODEX) (32, 33) dataset as well as a series of other CyTOF datasets (32) and spectral flow datasets (data not shown), further confirming its versatility.

These additional applications of Cyclone not only validated that the pipeline was functional in these settings but also demonstrated its strengths in identifying elements of the biological systems that may be overlooked with manual gating. Clustering of the tumor microenvironment imaging data revealed a CD163<sup>+</sup>XCR1<sup>+</sup> cell subset that was enriched in the stroma region. It is now well accepted that XCR1 expression defines the classical dendritic cells of type 1 (cDC1) across a wide range of organisms (34). It is therefore tempting to suggest that these cells may represent a population of CD163<sup>+</sup> DC1, which has been previously described in human breast and lung cancer patients (35). However, in this previous study, these CD163<sup>+</sup> DCs were defined as a discrete subset of DCs distinct from both cDC1s and cDC2s and had the ability to efficiently trigger CD103 expression in CD8<sup>+</sup> T cells *in vitro*, but the expression of XCR1 was not measured. The origin of these XCR1<sup>+</sup>CD163<sup>+</sup> cells remains unclear, and more work is warranted. Nevertheless, we anticipate that the flexibility and operability of the Cyclone pipeline will help address this question as well as aid in the investigation of these cells' spatial relationship within the tumor microenvironment and better define their potential role in the tumor immune response. In the spectral flow dataset, HBV-tetramer staining could be integrated with the clustering of T-cell subsets to better phenotype these antigen-specific cells. We found that the majority of tetramer+ CD4<sup>+</sup> or CD8<sup>+</sup> T cells clustered together in their respective compartments, indicating a shared phenotype. Unsurprisingly for this day 8 timepoint in the immune response, MHC Class I Tetramer+ CD8<sup>+</sup> T cells and MHC Class II Tetramer+ CD4<sup>+</sup> T cells fell primarily into clusters identified as effector T cells with high expression of markers of activation, including CD44 and PD-1, and markers of high proliferative capacity, including Ly108 and TCF-1. Notably, those two markers have been previously identified in antigen-specific CD8<sup>+</sup> T cells during Lymphocytic choriomeningitis virus (LCMV) chronic infection and cancer and have been associated with a cycling T-cell stage, which precedes exhaustion program upon chronic stimulation (36).

This work has several limitations, which can be explored or developed in future work. The set of clustering algorithms we compared for selection was restricted to four popular algorithms that are prevalent in cytometry analysis and have been elsewhere benchmarked (24), rather than us doing a more exhaustive search *de*

*novus*. In addition, though we selected FlowSOM as our default and fully optimized algorithm, CLARA and FlowSOM performed similarly in our comparisons; while the user has the option to select CLARA clustering, the DBI optimization only works with FlowSOM in our implementation. We additionally noted in our evaluation of the clustering that the selected resolutions sometimes failed to capture low-abundance populations as their clusters, such as cDC1s, eosinophils, or ASCs. This is unsurprising and a common challenge in global clustering approaches, which can be remedied either by 1) “over-clustering” the data (i.e., selecting a higher resolution such that you may better capture lower abundance populations but larger populations are further partitioned and need to be subsequently merged back together into a single population) or 2) “subclustering” the data (e.g., taking only the myeloid subsets and clustering them with relevant markers such that only those cells and markers are partitioning the space). We also noted decreased accuracy clusters defined by markers with a continuum of expression such as CD45RA (e.g., naïve v. memory T cells); however, because the placement of this manual gate on a continuum is somewhat arbitrary, modest discrepancy between the manually defined abundance and cluster abundance seems of low importance as long as each is applied consistently across samples of interest. In addition, while we found these algorithms to be robust to downsampling, we have found that FlowSOM and CLARA could not accommodate a larger CyTOF data set of ~90 million cells, regardless of resource dedication. Further optimization of those algorithms is needed to be able to accommodate increasingly large cytometry datasets. Finally, while we have invested in interoperability and extensive documentation and vignettes for ease of use, Cyclone could be even more accessible to a non-coding user base with the development of a graphical user interface (GUI) such as an R-shiny application (<https://shiny.rstudio.com>) or as a workflow in web-based tools such as the University of California San Francisco (UCSF) Data Library (<https://datalibrary.ucsf.edu/>), the Chan Zuckerberg Initiative’s CELLxGENE (<https://github.com/chanzuckerberg/cellxgene>), or CellEngine (<https://github.com/primitybio/cellengine-python-toolkit>), which could be the subject of future efforts. In sum, Cyclone takes the next step forward in the optimization and democratization of cytometry-based analysis tools to further power biological discovery.

## Methods

### Mass cytometry data generation and preprocessing

#### Sample collection

Blood samples from patients were obtained under institutional review board (IRB) #11-07994 protocol approved by the UCSF IRB, and #2012.059-2 (SCoo) protocol approved by Sutter Health IRB the IRB of record for the study. Written informed consent was obtained from all patients. De-identified healthy donor sample was obtained from Vitalant Research Institute (San Francisco, CA, USA). Blood was collected into sterile EDTA vacutainer tubes

(VWR International, Road Radnor, PA, USA) and processed within 24 hours of collection.

#### Sample preparation

PBMCs were isolated using Ficoll-Paque Plus (GE Healthcare, Chicago, IL, USA) density gradient centrifugation; after isolation, cells were aliquoted in  $0.5 \times 10^7$  cells per vial in cell freezing media (10% dimethyl sulfoxide (DMSO) in fetal bovine serum (FBS)) and cryopreserved.

#### CyTOF panel and staining

Mass cytometry was performed as described (37) with modifications. Briefly, primary conjugates of mass cytometry antibodies were prepared using the MaxPAR antibody conjugation kit (Fluidigm, South San Francisco, CA, USA) according to the manufacturer’s recommended protocol. Following labeling, antibodies were diluted in Candor PBS Antibody Stabilization solution (Candor Bioscience GmbH, Wangen, Germany) supplemented with 0.02% NaN<sub>3</sub> to between 0.1 mg/mL and 0.3 mg/mL, and stored long-term at 4°C. Each antibody clone was titrated to optimal staining concentrations using unstimulated or anti-CD3/CD28 stimulated PBMC samples. All mass cytometry antibodies and concentrations used for analysis can be found in **Supplemental Table 1**. Mass cytometry experiments were performed over the course of nine separate experiments. Each PBMC sample was thawed at 37°C and washed in pre-warmed RPMI-1640 media (Sigma-Aldrich Life Sciences, Burlington, MA, USA) supplemented with 10% FBS (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) in the presence of 250U Pierce Universal Nuclease for Cell Lysis (Thermo Fisher Scientific, Rockford, IL, USA); cells were counted using the Beckman Vi-Cell XR Cell Counter. Only samples with viability >75% were used (85% viability on average);  $2.5 \times 10^6$  cells/sample were stained for 1 min with 25 mM of cisplatin (Sigma-Aldrich) in phosphate-buffered saline (PBS) plus EDTA, before undergoing quenching 1:1 with PBS/EDTA/bovine serum albumin (BSA) to determine viability. Staining was performed on a shaker (90 rpm). For staining, cells were first resuspended in cell staining media (CSM) (Fluidigm, South San Francisco, CA, USA) with 5 µL of Human TruStain FcX™ block (BioLegend, San Diego, CA, USA) for 5 min at room temperature to block Fc receptors, followed by staining with CXCR5 antibody in CSM (3 µg/mL) for 30 min at 4°C. Cells were washed, fixed with Fix I Buffer from The Cell-ID™ 20-Plex Pd Barcoding Kit following the manufacturer’s instructions (Fluidigm, South San Francisco, CA, USA), and barcoded by mass-tag labeling with distinct combinations of stable Pd isotopes diluted in Maxpar Barcode Perm Buffer (Fluidigm, South San Francisco, CA, USA) as described previously (5). Twenty barcoded samples were pooled into a single fluorescence-activated cell sorting (FACS) tube (BD Biosciences, San Jose, CA, USA) and stained with a cocktail containing surface marker antibodies (**Supplemental Table 1**) in a final volume of 1,000 µL of CSM for 30 min at room temperature. Samples drawn at different timepoints per patient were barcoded together. Cells were then permeabilized with perm wash buffer (eBioscience, Thermo Fisher Scientific) following the

manufacturer's instructions and then incubated with a cocktail containing intracellular marker (Supplemental Table 1) antibodies diluted in perm wash buffer (eBioscience, Thermo Fisher Scientific) for 1 hour at 4°C. Cells were finally stained with 191/193Ir DNA intercalator (Fluidigm, South San Francisco, CA, USA) diluted in PBS with 1.6% paraformaldehyde (PFA) (Electron Microscopy Sciences, Hatfield, PA, USA) 24 hours prior to data acquisition.

### Data acquisition

For acquisition, cells were washed and resuspended at  $1 \times 10^6$ /mL in deionized water + 10%EQ four-element calibration beads (Fluidigm) and run on a Fluidigm CyTOF2 Helios mass Cytometer within 1 week of staining.

### Data preprocessing (premess) -> bead normalization)

After data collection, we used the premess pipeline (<https://github.com/ParkerICI/premess>) to normalize data and deconvolute individual samples. From the individual sample files, normalization beads were excluded based on Ce140 and Eu153 signals. Single-cell events were identified based on Ir191 DNA signal measured against event length, and CD45<sup>-</sup> Pt195<sup>+</sup> dead cells were excluded (Supplemental Figure 1). Potential batch effects were minimized by including a control sample from the same individual in each experimental run.

### FCS modifications (addition of unique cell IDs for tracking)

The FCS files do not contain cell identifiers. To accurately compare cell identity from either the manually gated annotations or the FlowSOM or PARC clusters annotated by two immunology experts, we added unique identifiers to all CD45<sup>+</sup> live gated cells across all FCS files and used these files for all downstream analyses. We created the unique identifiers by combining the sample identifier and the index of the single cells in each FCS file to generate unique identifiers, such as "<sample\_id>\_<cell\_id>". In this way, each cell gained a unique barcode id used for future comparative analyses.

### Batch correction

The samples were processed across nine batches. While the bead normalization of CyTOF data controls for the batch effects introduced due to instrument change, it does not address all factors affecting batch-to-batch variations (38). To evaluate the batch effects, we first clustered the single cells using CLARA (25) and compared the batch compositions across clusters. We observed uneven distribution of batches across clusters (data not shown). To account for this residual batch-to-batch variation, we corrected the signal for batches using CytoNorm (21). We used the control samples (the same sample that was replicated across batches) to train the model and adjusted the batch effects in non-control samples using nCells = 4k, nClus = 10, and the grid size of  $5 \times 5$

(xdim = 5; ydim = 5). We determined that CytoNorm adjustment removed the majority of batch effects from the data (Figure 3A).

### Manual gating

The main mass cytometry gating scheme can be found in Supplemental Figure 1 and shows the exclusion of beads, debris, dead cells, CD45<sup>-</sup> cells. Following this removal, we show the main gating strategy for identifying major immune cell populations from the mass cytometry dataset.

### Generation of subsets

Different numbers of cells were selected from batch-corrected FCS files to produce specific subsets of the data: 1k subset = 1,000 cells per FCS file, 10k subset = 10,000 cells per FCS file, 50k subset = 50,000 cells per FCS file, and full subset = all cells from FCS files.

### Evaluation of runtime and memory usage for different clustering tools

To scope clustering algorithms for our pipeline, we tested four Python- or R-based widely used clustering tools for cytometry data: CLARA, FlowSOM, PhenoGraph, and PARC. CyTOF analysis allows users to capture hundreds of thousands of cells, and clustering such large datasets requires runtime- and memory-efficient tools that do not compromise clustering performance. One aim of our work was to design a scalable pipeline for large datasets (containing many samples, each with hundreds of thousands of cells). Therefore, we decided to compare the runtime and memory usage of the selected clustering tools. Different parameters affect the runtime and memory usage in different tools. We observed the parameters affecting the number of clusters were the ones that controlled algorithm runtime. To perform a fair comparison between the tools, we identified these parameters affecting cluster counts and used values that produced a similar number of clusters across all four tools. We used  $k = 24$  in CLARA,  $xdim = 6$  and  $ydim = 6$  in FlowSOM,  $k = 25$  in PhenoGraph, and  $resolution = 1.3$  in PARC. These parameters resulted in 30–37 clusters called by each tool. We measured the time for clustering and the memory usage on CentOS nodes of a high-performance computer cluster.

## Pipeline

### Inputs

In order to begin a Cyclone run, metadata files "file\_metadata.csv", "marker\_metadata.csv", and "config.yml" must be generated. These files are unique to your dataset, but Cyclone requires certain metadata to locate the FCS files and associate metadata with them. In the file metadata csv, column "file\_name" records the name of each FCS file, "donor\_id" denotes sample origin, "pool\_id" denotes

batch identification (if any), and “control\_sample” is a Boolean indicating whether the sample is a control or not. The file metadata csv is created via any scripting language based on the FCS files present. To create the “marker\_metadata.csv”, we provide “cyclone/make\_marker\_metadata\_csv.R”, which will read in an FCS file, and create a file with the following columns: channel\_name, marker\_name, used\_for\_UMAP, used\_for\_clustering, and used\_for\_scaffold. It may be advantageous to use a marker for clustering, but not use that marker for UMAP calculation. Thus, the Boolean values for each marker in the “used\_for\_” columns provided granular controls for using markers for UMAP, clustering, and SCAFFoLD analysis. The final file “config.yml” controls how Cyclone will be run. In brief, it contains the absolute path of FCS files and save location for pipeline outputs, location of metadata files, and parameters associated with data processing (arcsinh\_cofactor, default 5), UMAP (n\_neighbors, default 15; min\_dist, default 0.1; spread, default 0.1; learning\_rate, default 0.5; “random” to assign initial embedding positions), and clustering (k, default 3) to tune node connectivity. Cyclone is able to cluster with both FlowSOM and CLARA, and the choice is specified in “config.yml”. CLARA editable parameters are k (default 20), metric (default euclidean), and samples (default 50). The FlowSOM editable parameters are xdim (default 6) and ydim (default 6), which define the grid (6 × 6). If meta\_clustering with FlowSOM is TRUE, (default FALSE), the “nClus” of FlowSOM (i.e. the number of meta-clusters) can be defined using “k” (default 3).

## Processing steps

At the start of the Cyclone run, the pipeline references the metadata.csv files to read in FCS data files and creates a raw expression matrix. After the arcsinh transformation of count values, a transformed expression matrix is created using values specified in config.yml. Cyclone also creates a “cell\_metadata” object, which associates each cell/event with file and marker metadata. Next, Cyclone uses the uwot package with default parameters to calculate UMAP on the transformed matrix with the “used\_for\_UMAP” column of the markers’ metadata file. Each cell/event UMAP dimensions are assigned to the cell\_metadata object. Next, clustering is performed using either FlowSOM or CLARA. In FlowSOM, a default series of grids (cyclone/grid\_sizes.csv) are specified, and different resolutions and clustering parameters are calculated and evaluated with the DBI. After cluster optimization, the Cyclone pipeline exits to await user evaluation of the Cluster VS DBI plot (Figure 1D, Supplemental Figure 2A). If CLARA is used for clustering, the user must specify clustering parameters, and no DBI-based optimization is performed. After specifying a specific grid (FlowSOM) or specifying “k” in config.yml to calculate clusters (CLARA), clustering is calculated on the transformed matrix with the “used\_for\_clustering” column from the markers’ metadata file. Cluster assignment is stored in the cell\_metadata object. After clustering, Cyclone calculates cluster frequency matrices (raw and normalized) and calculates cluster median expression matrix. If SCAFFoLD analysis is selected, a gated directory of landmark FCS files is required. For each cluster, Cyclone obtains the closest landmark population and stores this assignment in

cluster\_metadata. After calculating statistics, SCAFFoLD analysis is saved in a cluster\_metadata object. With analysis completed, Cyclone outputs several helpful plots.

## CytoTOF cluster annotation and benchmarking using manually gated populations

After clustering the data using FlowSOM or PARC, we performed manual annotation of the resulting clusters from both tools based on the median expression of markers. Since each individual may annotate clusters differently, we attempted to account for human-to-human variations in the manual annotation of clusters by having two immunology experts independently annotate the clusters from FlowSOM and PARC. We established “coarse” and “fine”-levels of annotations. Coarse annotations describe cells of different compartment groups (e.g., CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and B cells). Fine annotations further parse cell types into subtypes according to their phenotype (e.g., CD8<sup>+</sup> T cells are split further into naïve, central memory, effector memory, and effector memory re-expressing CD45RA). We then calculated similarities between cluster annotations of single cells and the single-cell annotation based on a third immunology expert gating of the CyTOF data (ground truth annotation). We performed this comparison for subsets of data multiple times and calculated averages. Specifically, we subsampled cluster annotation and manual gating annotation of randomly selected 10,000 cells, calculated accuracy, and adjusted the Rand index, Fowlkes–Mallows index, and mutual information. We repeated this process 10 times with different random seeds and calculated the mean across the iterations for each similarity metric. We performed this analysis for FlowSOM (Figure 1H, Supplemental Figure 2C) and PARC annotations (data not shown) from both immunology experts.

## dittoSeq visualizations

Cyclone outputs (checkpoint1.Rdata and checkpoint8.Rdata) were used to create a SingleCellExperiment object (29) in R containing the arcsinh transformed expression matrix, UMAP embeddings, clustering, and cell and sample-level metadata. dittoSeq (16) functions dittoFreqPlot and dittoBarPlot were then used to create boxplots of cluster frequencies per sample and stacked bar plots of batch composition per cluster, respectively.

## Spectral flow data generation and analysis

### Mice

Wild-type (WT) C57BL/6 mice were purchased from Jackson Laboratory (Bar Harbor, ME, USA). HBVEnvRag<sup>-/-</sup> mice were previously described (30). Briefly, HBVEnvRag<sup>-/-</sup> mice were generated using HBVEnv<sup>+</sup> mice [lineage 1075D; gift from F. Chisari, Scripps Research Institute (39)] backcrossed to Rag1<sup>-/-</sup>

C57BL/6 mice for 15 generations. HBVEnvRag<sup>-/-</sup> mice contain the entire envelope (subtype ayw) protein coding region under the constitutive transcriptional control of the mouse albumin promoter. Young (3 weeks old, before weaning) or adult (8 to 12 weeks old) HBVEnvRag<sup>-/-</sup> mice were given 10<sup>8</sup> syngeneic splenocytes pooled from adult (8 to 12 weeks) WT mice in 0.5 mL of phosphate-buffered saline via tail vein injection. Mice were maintained at the Laboratory Animal Resource Center (LARC) facility at UCSF where health was monitored daily by the LARC staff. Experimental procedures were performed in accordance with Institutional Animal Care and Use Committee (IACUC)-approved protocols, and all efforts were made to minimize animal suffering.

### Sample preparation

Mice were anesthetized in chambers with 1.5% oxygen and 3% isoflurane. Samples for the T cell-focused panel were isolated from the liver after perfusion and digestion. Briefly, mice were perfused via the inferior vena cava using digestion media [Hanks' Balanced Salt Solution (HBSS), crude collagenase (0.2 mg/mL; Crescent Chemical, Islandia, NY, USA), and DNase I (0.02 mg/mL; Roche Diagnostics, Basel, Switzerland)]. Livers were forced through a 70µm filter using a syringe plunger, and debris was removed by centrifugation (30 g for 3 min). Supernatants were collected and centrifuged for 10 min at 650 g. Cells were isolated from the Percoll interface using a 60%:40% Percoll gradient. Samples for the myeloid-focused panel were isolated from the liver after 6 min of perfusion via the inferior vena cava using digestion media as above. Livers were chopped and further digested with liberase and DNase I (Roche Diagnostics) [1 Wunsch Units (WU) and 0.8 mg, respectively, in 10 mL of RPMI-1640 containing 5% FBS] for 30 min at 37°C in a shaking water bath. Livers were forced through a 70µm filter, and debris was removed by centrifugation (30 g for 3 min). Supernatants were collected and centrifuged for 10 min at 650 g. Cells were isolated from the interface of a 60%:40% Percoll gradient.

### Staining and acquisition

Cells were prepared as above. Samples stained with the T cell-focused panel were first stained with custom HBV-specific tetramers developed by the National Institutes of Health (NIH) Tetramer Core Facility at Emory University. Cells were stained first with MHC Class II Tetramer for 1 hour at 37°C protected from light. Next, these cells were then stained with MHC Class I Tetramer for 1 hour at 4°C protected from light. For both panels, cells were then stained with Live/Dead Fixable Blue (Thermo Fisher Scientific) according to the manufacturer's instructions. Next, surface markers on cells were stained according to standard protocols with anti-mouse antibodies detailed in [Supplemental Table 2](#) (Myeloid-focused Panel) or [Supplemental Table 3](#) (T cell-focused Panel). Finally, cells were fixed and permeabilized using FoxP3/Transcription Factor Staining Buffer Set (Thermo Fisher Scientific, cat. 00-5523-00) and stained with anti-mouse antibodies targeting intracellular markers according to standard protocols.

### Data acquisition and preprocessing

Single-color reference controls were collected for live unmixing with calculated autofluorescence immediately before fully stained

sample acquisition. For sample acquisition, cells were run the same day as preparation and staining on an Aurora flow cytometer (Cytek, Fremont, CA, USA) with a 5-laser setup at the UCSF Flow CoLab. Following sample collection, spectral signatures were checked to ensure reliable unmixing, and channel spillover was adjusted in SpectroFlo (Cytek). Prior to Cyclone and manual gating analyses, FCS files were gated on leukocyte size/granularity, singlets, live, and CD45<sup>+</sup> events in FlowJo ([Supplemental Figure 4C](#)).

### Cyclone analysis

Cyclone pipeline was run as described above, specifying an arcsinh transformation cofactor of 6000, with all other values kept as default ([Supplementary Table 4](#)). For both panels, forward scatter, side scatter, CD45, Live/Dead Blue, and autofluorescence were excluded for UMAP generation and clustering. In addition, for the T cell-focused panel, channels for tetramers were excluded for UMAP generation and clustering. For the T-cell panel, 21 clusters were identified, all of which were identified as specific lymphocyte populations (B cells, NK cells, T cells, or ILCs) or were assumed to belong to the myeloid compartment ([Supplemental Figure 4A](#)). For the myeloid panel, 20 clusters were identified. Among these clusters, three were unidentifiable by markers in the panel ([Supplemental Figure 4B](#)) and excluded from subsequent analysis ([Figures 5B, D](#)).

### Manual gating

Expert manual gating was performed in FlowJo to assign unique cell-type identities to events. Example plots for T cell-focused gating strategy ([Supplemental Figure 4D](#)) and Myeloid-focused panel ([Supplemental Figure 4E](#)) are provided.

### Annotation comparison

FCS files for manual gates were exported from FlowJo and then read into a flowFrame object in R using the FlowCore package. Raw data were used to uniquely match events and assign annotations from manual gating and Cyclone clustering. Cells that received both manual gating and Cyclone clustering annotations were used to generate annotation concordance heatmaps ([Figures 5C, D](#)).

### Tetramer+ cell visualization

Tetramer+ events were identified as having raw fluorescent intensities 3 or more standard deviations above the mean intensity.

## Imaging data generation and analysis

### Sample preparation

All patients consented by the UCSF IPI clinical coordinator group for tissue collection under a UCSF IRB-approved protocol (UCSF IRB #20-31740). Samples were obtained after surgical excision with biopsies taken by pathology assistants to confirm the presence of tumor cells. Freshly resected samples were placed in ice-cold PBS or Leibovitz's L-15 medium in a 50-mL conical tube, immediately transported to the laboratory for sample labeling, and formalin fixed for imaging analysis. Clinical data on three samples were denoted as follows: CRC1 = IPIICRC072, CRC2 = IPIICRC057, and KID1 = IPIKID090.



## Staining and imaging immunofluorescent 7-plex panel

A 7-plex immunofluorescent panel was created using the Ventana BenchMark Ultra (Roche Diagnostics) automated staining platform. All reagents were from Discovery (Ventana Medical Systems, Tucson, AZ, USA) and used according to the manufacturer's instructions, except as noted. Heat-Induced Epitope Retrieval (HIER) was performed with the Cell Conditioning 1 (CC1) solution (cat. 950-124) for 64 min at 97°C. The primary antibodies used were CD3 (1:100, clone: D7A6E from Cell Signaling Technology, Danvers, MA, USA), CD4 (RTU, clone: SP35 from Ventana), CD8 (1:100, clone: D8A8Y from Cell Signaling Technology), CD163 (1:250, clone: EPR19518 from Abcam, Cambridge, UK), HLA-DR (1:500, clone: EPR3692 from Abcam), XCR1 (1:40, clone: D2F8T from Cell Signaling Technology), and EpCAM (1:50, clone: D9S3P, from Cell Signaling Technology). The tissue was counterstained with DAPI (Akoya cat. FP1490) for nucleus localization. The staining was conducted in two cycles: the first cycle had CD3, CD4, CD8, CD163, HLA-DR, and XCR1; the second cycle had EpCAM. Both cycles had DAPI. The slide was scanned using a whole slide scanner after each staining cycle. Finally, the images from both cycles were registered to achieve the 7-plex image shown in [Figure 6](#).

## Data preprocessing

Cell segmentation was performed by utilizing ark-analysis (v0.2.9) DeepCell (31) software with nuclei (DAPI) and membrane (CD3, CD4, CD8, and CD163) as modalities for segmentation. Mean fluorescent intensity was measured from each cell region of interest (ROI) and then arcsinh transformed to input into Cyclone pipeline as "trans\_exp".csv file. Manual gating was performed in a custom Napari application version 0.4.14 to classify cells as positive or negative for each marker. For stroma versus tumor region separation, Qupath (version 0.3) software was used to annotate each region using EpCAM for colorectal or PanCK for kidney as a marker reference. Finally, data cluster labels generated from the Cyclone pipeline or manual gating were generated as a csv file corresponding to each cell ROI and integrated with a segmented imaging file for cluster overlay in Napari.

## Cyclone analysis

After determining an optimal grid size, the Cyclone output heatmap showing arcsinh transformed and scaled marker expression data were used to annotate clusters and scaled by row. Log2 fold change of each cluster in stroma versus tumor was determined by calculating the frequency of stromal over frequency of tumor in each cluster and transformed by a log of 2. Contour plots comparing Cyclone versus manual gating were generated in python (version 3.8.12) by plotting arcsinh transformed cell ROIs annotated as XCR1 and CD163.

## Data availability statement

The FCS files for the CyTOF experiments used in this study are deposited and publicly available at FlowRepository ([http://](http://flowrepository.org/id/FR-FCM-Z6LS)

[flowrepository.org/id/FR-FCM-Z6LS](http://flowrepository.org/id/FR-FCM-Z6LS)). The cyclone pipeline and package are available at [github.com/UCSF-DSCOLAB/cyclone/](https://github.com/UCSF-DSCOLAB/cyclone). The documentation on running the cyclone is available at <https://github.com/UCSF-DSCOLAB/cyclone/blob/main/vignettes/Running.md> and the documentation on follow up analyses is available at <https://github.com/UCSF-DSCOLAB/cyclone/blob/main/vignettes/FollowUp.html>.

## Ethics statement

The studies involving humans were approved by UCSF and Sutter Health Institutional Review Boards. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. All animal experiments done in this study were approved by the UCSF Institutional Animal Care and Use Committee. The study was conducted in accordance with the local legislation and institutional requirements.

## Author contributions

RP and RJ contributed equally to this manuscript. RP, RJ, KI, NDC, TC, NWC, BS, DB, LL-J, JJ, JP, and LA, generated and analyzed data and contributed to the manuscript by providing figures, tables, and important intellectual contributions. RP, RJ, AC, and GF had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the analyses. RP, RJ, KI, DB, NDC, AC, and GF wrote and edited the manuscript. RP, RJ, BS, DB, AR, and LL-J performed computational analyses of the data. MN and NWC were part of the UCSF Immunoprofiler team who performed tissue preparation and staining for the immunofluorescence performed on the different human tumor specimens. RP, RJ, BS, DB, and AR developed the Cyclone package and wrote the tutorial. AE and KI performed computational analysis of the multiplexed immunofluorescence data on human tumors. MK, SC, JB, AC, and GF were actively involved in the direction of projects and provided financial resources for the obtainment of the data. AC and GF co-led this project. All authors edited and critically revised the manuscript for important intellectual content and gave final approval for the version to be published.

## Funding

Acquisition and analysis of certain human samples described in this study were partially funded by contributions from AbbVie, Amgen, Bristol-Myers Squibb, and Pfizer as part of the UCSF Immunoprofiler Initiative. Data acquisition was further supported by R01DK103735 and P30DK026743. Further support came from the Bakar ImmunoX Initiative and ImmunoX Computational Biology Initiative at UCSF funding for AC, GF, DB, RP, and RJ.

## Acknowledgments

We thank all members of the Combes and Fragiadakis labs, UCSF Bakar ImmunoX Initiative, UCSF CoLabs, and the UCSF Immunoprofiler Consortium for discussion and guidance while developing this study. We would like to thank Suprita Trilok for help with spectral flow cytometry. We would like to thank Drs David Erle, Michael G Kattah, and Elvira Mennillo for scientific discussion and editing the manuscript. We would like to particularly thank Isabelle Tingin, Garry Shumakher, Elizabeth Edmiston, and Meghan Zubradt for their constant support and help during this study. Finally, we thank all patients and their families for placing their trust in us.

## Conflict of interest

AC is a shareholder and member of the scientific advisory board of Foundry Innovations. AC receives funding from Genentech, Corbus, and Ely Lilly. GF receives funding from Eli Lilly. MK is a founder and shareholder of Pionyr Immunotherapeutics and Foundry Innovations.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1167241/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

CytoF manual gating of human PBMC. (A) Pre-Gating: gating out beads, debris, dead cells, RBC and granulocytes. (B) Hierarchical gating was applied to identify 22 "landmark" immune populations: CD14<sup>+</sup> CD16<sup>-</sup> classical monocytes, CD14<sup>-</sup> CD16<sup>+</sup> nonclassical monocytes, CD14<sup>+</sup> CD16<sup>+</sup>

intermediate monocytes, cDC1, cDC2, pDC, basophils, Natural Killer cells, regulatory CD4<sup>+</sup> T cells, CD4<sup>+</sup> T cells (Naive, T<sub>CM</sub>, T<sub>EM</sub>, T<sub>EMRA</sub>), CD8<sup>+</sup> T cells (Naive, T<sub>CM</sub>, T<sub>EM</sub>, T<sub>EMRA</sub>),  $\gamma\delta^+$  T cells, B cells, ASC (antibody producing cells) HLA-DRpos (CXCR5- B cells), plasmablasts. Also shown gating of eosinophils (CD15<sup>+</sup> CD16<sup>+</sup> HLA-DR<sup>-</sup>) and neutrophils (CD15<sup>+</sup> CD16<sup>-</sup> HLA-DR<sup>-</sup>).

### SUPPLEMENTARY FIGURE 2

Assessment of "Fine"-level annotations and metrics for evaluating cyclone outputs. (A) Full Davis-Bouldin index plot showing up to 200 potential clusters identified through FlowSOM. (B) Heatmap of full dataset "fine"-level annotations identifying cell types and cell subtypes, based on ground truth (GT) manual gating (rows) compared to annotated FlowSOM clusters (columns). (C) Comparison metrics based on "fine"-level annotations from two individuals. Various performance metrics were used to assess the accuracy of clusters called in the FlowSOM clustering compared to ground truth. (D) Ground Truth expert cluster "fine"-level annotation identifying broad cell types and specific cell subtypes based on manual gating. (E) FlowSOM clustering "fine"-level annotations based on CyTOF panel expression. (F) Depiction of a cluster dispersed across UMAP space (Cluster 26) with a heterogenous protein expression profile compared to (G) a cluster with uniform protein expression and tight UMAP localization (Cluster 28).

### SUPPLEMENTARY FIGURE 3

"Fine"-level annotations after running Cyclone on the downsampled dataset. The dataset was down-sampled to 50k cells per sample and then run through cyclone. Clusters' cell type identities were inferred by experts using Cyclone plot outputs. (A) UMAP annotated by cluster number. (B) Heatmap of median  $\text{arcsinh}$  transformed expression (unscaled) per cluster, used to annotate clusters. (C) UMAP from 50k down-sample run, colored by fine annotations. (D) Comparison of per-cell annotations between the 50k down-sample versus the full dataset. Various performance metrics were used to assess the accuracy of clusters called in the downsampled dataset compared to the full dataset.

### SUPPLEMENTARY FIGURE 4

Spectral flow cytometry cell type identification. (A) Heatmap of markers used for UMAP generation, clustering, and identification for "fine"-level Cyclone clusters for the spectral flow cytometry dataset with a T cell-focused panel presented in Figure 5A. (B) Heatmap of markers used for UMAP generation, clustering, and identification for "fine"-level Cyclone clusters for the spectral flow cytometry dataset with a Myeloid-focused panel presented in 5B. (C) Representative two-dimensional flow plots demonstrating pre-gating on live CD45<sup>+</sup> cells before analysis with either Cyclone or expert manual gating in FlowJo. (D) Manual gating strategy for samples in the T cell-focused panel. (E) Manual gating strategy for samples in the Myeloid-focused panel.

### SUPPLEMENTARY FIGURE 5

Resolution of manual gating and Cyclone clustering. (A) Table (top, left) depicting cell type annotations based on the 7 markers used and cell frequency histograms with manual gating thresholds (red line) on above (yellow) or below (blue) threshold for CD163 and CD3 marker. Thresholds were set on CD163 and CD3 at 11, 12, and 13 to denote the cutoff and visualization of annotated macrophage<sup>+</sup> cells (pink dots). Circled white regions indicate examples of over-thresholding (threshold 11) and under-thresholding (threshold 13). Scale bar denotes 20 $\mu$ m. (B) Frequency bar plot of each cluster annotation divided into stromal (pink) or tumor (blue) compartments. (C) Log2 fold change bar plot on stromal/tumor ratio for each annotated cluster in CRC1, CRC2, and KID1 tumor samples. (D) Scatterplot of cDC1<sup>+</sup> (pink), XCR1<sup>+</sup>CD163<sup>+</sup> (blue) cells, and other between cyclone pipeline (left) and manual gating (right).

## References

- Ginhoux F, Yalin A, Dutertre CA, Amit I. Single-cell immunology: Past, present, and future. *Immunity* (2022) 55(3):393–404. doi: 10.1016/j.immuni.2022.02.006
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. *Cell* (2019) 177(7):1888–902. doi: 10.1016/j.cell.2019.05.031
- Bendall SC, Simonds EF, Qiu P, Amir ED, Krutzik PO, Finck R, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* (2011) 332(6030):687–96. doi: 10.1126/science.1198704
- Nolan JP, Condello D. Spectral flow cytometry. *Curr Protoc Cytom* (2013) 27(63):1.27.1–1.27.13. doi: 10.1002/0471142956.cy0127s63
- Spitzer MH, Carmi Y, Reticker-Flynn NE, Kwek SS, Madhiredy D, Martins MM, et al. Systemic immunity is required for effective cancer immunotherapy. *Cell* (2017) 168(3):487–502. doi: 10.1016/j.cell.2016.12.022

6. Gaudillière B, Fragiadakis GK, Bruggner RV, Nicolau M, Finck R, Tingle M, et al. Clinical recovery from surgery correlates with single-cell immune signatures. *Sci Transl Med* (2014) 6(255):255ra131. doi: 10.1126/scitranslmed.3009701
7. Rahil Z, Leylek R, Schürch CM, Chen H, Bjornson-Hooper Z, Christensen SR, et al. Landscape of coordinated immune responses to H1N1 challenge in humans. *J Clin Invest* (2020) 130(11):5800–16. doi: 10.1172/JCI137265
8. Vallvé-Juanico J, George AF, Sen S, Thomas R, Shin M-G, Kushnour D, et al. Deep immunophenotyping reveals endometriosis is marked by dysregulation of the mononuclear phagocytic system in endometrium and peripheral blood. *BMC Med* (2022) 20(1):158. doi: 10.1186/s12916-022-02359-4
9. Quintelier K, Couckuyt A, Emmaneel A, Aerts J, Saeys Y, Van Gassen S. Analyzing high-dimensional cytometry data using FlowSOM. *Nat Protoc* (2021) 16(8):3775–801. doi: 10.1038/s41596-021-00550-0
10. Stassen SV, Siu DMD, Lee KCM, Ho JWK, So HKH, Tsia KK. PARC: ultrafast and accurate clustering of phenotypic data of millions of single cells. *Bioinformatics* (2020) 36(9):2778–86. doi: 10.1093/bioinformatics/btaa042
11. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir ED, Tadmor MD, et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* (2015) 162(1):184–97. doi: 10.1016/j.cell.2015.05.047
12. Amir ED, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat Biotechnol* (2013) 31(6):545–52. doi: 10.1038/nbt.2594
13. Anchang B, Hart TDP, Bendall SC, Qiu P, Bjornson Z, Linderman M, et al. Visualization and cellular hierarchy inference of single-cell data using SPADE. *Nat Protoc* (2016) 11(7):1264–79. doi: 10.1038/nprot.2016.066
14. Polikowsky HG, Drake KA. Supervised machine learning with CITRUS for single cell biomarker discovery. *Methods Mol Biol* (2019) 1989:309–32. doi: 10.1007/978-1-4939-9454-0\_20
15. Spitzer MH, Gherardini PF, Fragiadakis GK, Bhattacharya N, Yuan RT, Hotson AN, et al. An interactive reference framework for modeling a dynamic immune system. *Science* (2015) 349(6244):1259425. doi: 10.1126/science.1259425
16. Bunis DG, Andrews J, Fragiadakis GK, Burt TD, Sirota M. dittoSeq: universal user-friendly single-cell and bulk RNA sequencing visualization toolkit. *Bioinformatics* (2020) 36:5535–5536. doi: 10.1093/bioinformatics/btaa1011
17. Finck R, Simonds EF, Jager A, Krishnaswamy S, Sachs K, Fantl W, et al. NorMalization of mass cytometry data with bead standards. *Cytometry A* (2013) 83(5):483–94. doi: 10.1002/cyto.a.22271
18. Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, Bruggner RV, et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat Biotechnol* (2012) 30(9):858–67. doi: 10.1038/nbt.2317
19. Davies DL, Bouldin DW. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell* (1979) 1(2):224–7. doi: 10.1109/TPAMI.1979.4766909
20. Patterson-Cross RB, Levine AJ, Menon V. Selecting single cell clustering parameter values using subsampling-based robustness metrics. *BMC Bioinf* (2021) 22(1):39. doi: 10.1186/s12859-021-03957-4
21. Van Gassen S, Gaudillière B, Angst MS, Saeys Y, Aghaeepour N. Cytonorm: A normalization algorithm for cytometry data. *Cytometry A* (2020) 97(3):268–78. doi: 10.1002/cyto.a.23904
22. Pedersen CB, Dam SH, Barnkob MB, Leipold MD, Purroy N, Rassenti LZ, et al. cyCombine allows for robust integration of single-cell cytometry datasets within and across technologies. *Nat Commun* (2022) 13(1):1698. doi: 10.1038/s41467-022-29383-5
23. Weber LM, Robinson MD. Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry A* (2016) 12:1084–96. doi: 10.1002/cyto.a.23030
24. Liu X, Song W, Wong BY, Zhang T, Yu S, GN L, et al. A comparison framework and guideline of clustering methods for mass cytometry data. *Genome Biol* (2019) 20(1):297. doi: 10.1186/s13059-019-1917-7
25. Kaufman L, Rousseeuw PJ eds. *Finding groups in data: an introduction to cluster analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc (1990).
26. Hubert L, Arabie P. Comparing partitions. *J Classification* (1985) 2(1):193–218. doi: 10.1007/BF01908075
27. Fowlkes EB, Mallows CL. A method for comparing two hierarchical clusterings. *J Am Stat Assoc* (1983) 78(383):553–69. doi: 10.1080/01621459.1983.10478008
28. Cover TM, Thomas JA. *Elements of information theory*. Hoboken, NJ, USA: John Wiley & Sons, Inc (2005).
29. Amezcua RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with Bioconductor. *Nat Methods* (2020) 17(2):137–45. doi: 10.1038/s41592-019-0654-x
30. Baron JL, Gardiner L, Nishimura S, Shinkai K, Locksley R, Ganem D. Activation of a nonclassical NKT cell subset in a transgenic mouse model of hepatitis B virus infection. *Immunity* (2002) 16(4):583–94. doi: 10.1016/S1074-7613(02)00305-9
31. Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat Biotechnol* (2022) 40(4):555–65. doi: 10.1038/s41587-021-01094-0
32. Mennillo E, Kim YJ, Rusu I, Lee G, Dorman LC, Bernard-Vazquez F, et al. Single-cell and spatial multi-omics identify innate and stromal modules targeted by anti-integrin therapy in ulcerative colitis. *BioRxiv* (2023). doi: 10.1101/2023.01.21.525036
33. Goltsev Y, Samusik N, Kennedy-Darling J, Bhate S, Hale M, Vazquez G, et al. Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. *Cell* (2018) 174(4):968–981.e15. doi: 10.1016/j.cell.2018.07.010
34. Crozat K, Guiton R, Contreras V, Feuillet V, Dutertre C-A, Ventre E, et al. The XC chemokine receptor 1 is a conserved selective marker of mammalian cells homologous to mouse CD8alpha+ dendritic cells. *J Exp Med* (2010) 207(6):1283–92. doi: 10.1084/jem.20100223
35. Bourdely P, Anselmi G, Vaivode K, Ramos RN, Missolo-Koussou Y, Hidalgo S, et al. Transcriptional and functional analysis of cd1c+ human dendritic cells identifies a CD163+ subset priming CD8+CD103+ T cells. *Immunity* (2020) 53(2):335–352.e8. doi: 10.1016/j.immuni.2020.06.002
36. Beltra J-C, Manne S, Abdel-Hakeem MS, Kurachi M, Giles JR, Chen Z, et al. Developmental relationships of four exhausted CD8+ T cell subsets reveals underlying transcriptional and epigenetic landscape control mechanisms. *Immunity* (2020) 52(5):825–41. doi: 10.1016/j.immuni.2020.04.014
37. Allen BM, Hiam KJ, Burnett CE, Venida A, DeBarge R, TenVooren I, et al. Systemic dysfunction and plasticity of the immune macroenvironment in cancer models. *Nat Med* (2020) 26(7):1125–34. doi: 10.1038/s41591-020-0892-6
38. Schuyler RP, Jackson C, Garcia-Perez JE, Baxter RM, Ogolla S, Rochford R, et al. Minimizing batch effects in mass cytometry data. *Front Immunol* (2019) 10:2367. doi: 10.3389/fimmu.2019.02367
39. Chisari FV, Filippi P, McLachlan A, Milich DR, Riggs M, Lee S, et al. Expression of hepatitis B virus large envelope polypeptide inhibits hepatitis B surface antigen secretion in transgenic mice. *J Virol* (1986) 60(3):880–7. doi: 10.1128/jvi.60.3.880-887.1986