

## Abstract

Phagocytosis is a critical process for maintaining homeostasis and immune response. The process is carried out most efficiently by a set of immune cells collectively referred to as professional phagocytes that include macrophages, neutrophils, monocytes, and dendritic cells. There is great diversity in phagocyte activity, targets, and interaction with the microenvironment. An automated in-depth characterization of phagocyte subpopulations is of interest to many applications dependent on the plasticity of the system. Herein, we describe an effort to characterize phagocyte subpopulations using a multimodal assay of high-parameter flow and imaging, using pattern recognition and statistical methods for population identification and validation, and a machine learning system for forward propagating results to follow-up data.

Phenotypically, phagocytes can be characterized by CD45 and co-expression of various lineage dependent markers. In this work we used a 20-color assay targeting cell surface, signaling and activation markers. Additionally, the physical process of phagocytosis entails membrane extension, engulfing, and digesting, all of which are detectable by imaging and image derived parameters. Phagocytosis was initiated by introducing *E. coli* to PBMCs, with the *E. coli* labeled with pHrodo dye which will only fluoresce when the environment acidifies, indicating bacterial digestion. The complete input data set thus includes light scatter, 20 cytometric parameters, cell images, and eight image derived morphological parameters for each of light loss, light scatter, and pHrodo signal. We identified cellular populations via unsupervised clustering and tSNE for visualization first using only the cytometric parameters and then including the image derived parameters as well. Comparing the two outcomes, we concluded that adding image-based information improved population identification using both statistical and visual methods. Statistical methods included a dip test and modality score for comparing homogeneity within a cluster and heterogeneity between clusters with average values across all clusters comprising the overall measure of goodness. Visual methods involved inspection of the images per cluster using an overlay of pHrodo expression with morphology to verify the consistency of whole cells with internal fluorescence. Visual inspection revealed consistent but not entirely homogeneous results, including a preponderance of phagocytosis in varying stages germane to particular cluster. We proceeded to refine our clusters by hand-selecting subsets of twenty-five representative cells per cluster, and using nearest-neighbor estimation to expand that set to the one thousand most similar events based on cytometric and image derived measurements. These training sets were then used with Hyperfinder to produce sortable gating strategies to use these cells in future experiments.

## Workflow

### 1A Fluorochrome panel and Image Measurements

Fluorochrome	Target	Measure	Target
BUV563	CD3_CD19 (Dump)	Size	pHrodo
BUV395	CD11b	Max Intensity	pHrodo
BV605	CD11c	Total Intensity	pHrodo
APC-H7	CD14	Long Axis Moment	pHrodo
BV650	CD15	Short Axis Moment	pHrodo
V450	CD16	Eccentricity	pHrodo
BUV661	CD32	Radial Moment	pHrodo
BV786	CD33	Diffusivity	pHrodo
BV711	CD45RA	Size	FSC
BV510	CD56	Long Axis Moment	FSC
BUV737	CD64	Short Axis Moment	FSC
AF 647	CD66b	Eccentricity	FSC
R718	CD86	Radial Moment	FSC
BV421	CD123	Size	SSC
BV480	CD172a	Long Axis Moment	SSC
BV750	CD193	Short Axis Moment	SSC
BUV615	CXCR-1	Eccentricity	SSC
RY586	FCER1	Radial Moment	SSC
BUV805	HLA-DR		SSC
	pHrodo		

### 1B Analysis

**Quality Control**  
PeacoQC<sup>1</sup>

- Input Isolation Tree: 0.6 / Input MAD: 6.00
- Remove Margins: Off
- Run on all measured parameters

**Manual gates for cleanup**

Minimal gating to identify whole non-T or B cells

**Clustering and dimensionality reduction**

Phenograph<sup>2</sup> and optSNE<sup>3</sup>  
Applied twice for comparison:

- Fluorescent parameters only
- Fluorescent + imaging parameters

**Clustering identification**  
Heat mapping – used to understand the big picture of parameter expression  
Marker Enrichment Modeling<sup>4</sup> (MEM) – used to quickly identify differentiating markers on pHrodo+ populations

**Cluster goodness evaluation**  
Euclid

- Dip test and Taylor index applied
- All parameters used in clustering included

### 1C Image Exploration

BD CellView Lens plugin

- Images + statistics were used to explore the difference between pHrodo expressing clusters

Figure 1  
pHrodo+ clusters with representative images highlighted.

Images sets  
Image sets were created by manually selecting 25-50 exemplar cells per pHrodo expressing cluster, then using a nearest neighbors calculated on fluorescent and image derived parameters to find the 1,000 most similar cells for use as training data for HyperFinder gates.

## Methods

### Data Acquisition

Data were acquired on the BD FACSDiscover™ S8. Fluorescent panel and image measurements displayed in Figure 1a were used for all processing. PBMCs were exposed to *e-coli* particles that only fluoresce once internalized due to the presence of the pHrodo dye.

### Preprocessing

Data were cleaned using the PeacoQC algorithm and minimal manual gates designed to remove debris, doublets, cells thought not to use phagocytosis, and one additional gate to highlight cells with some pHrodo signal.

### Analysis

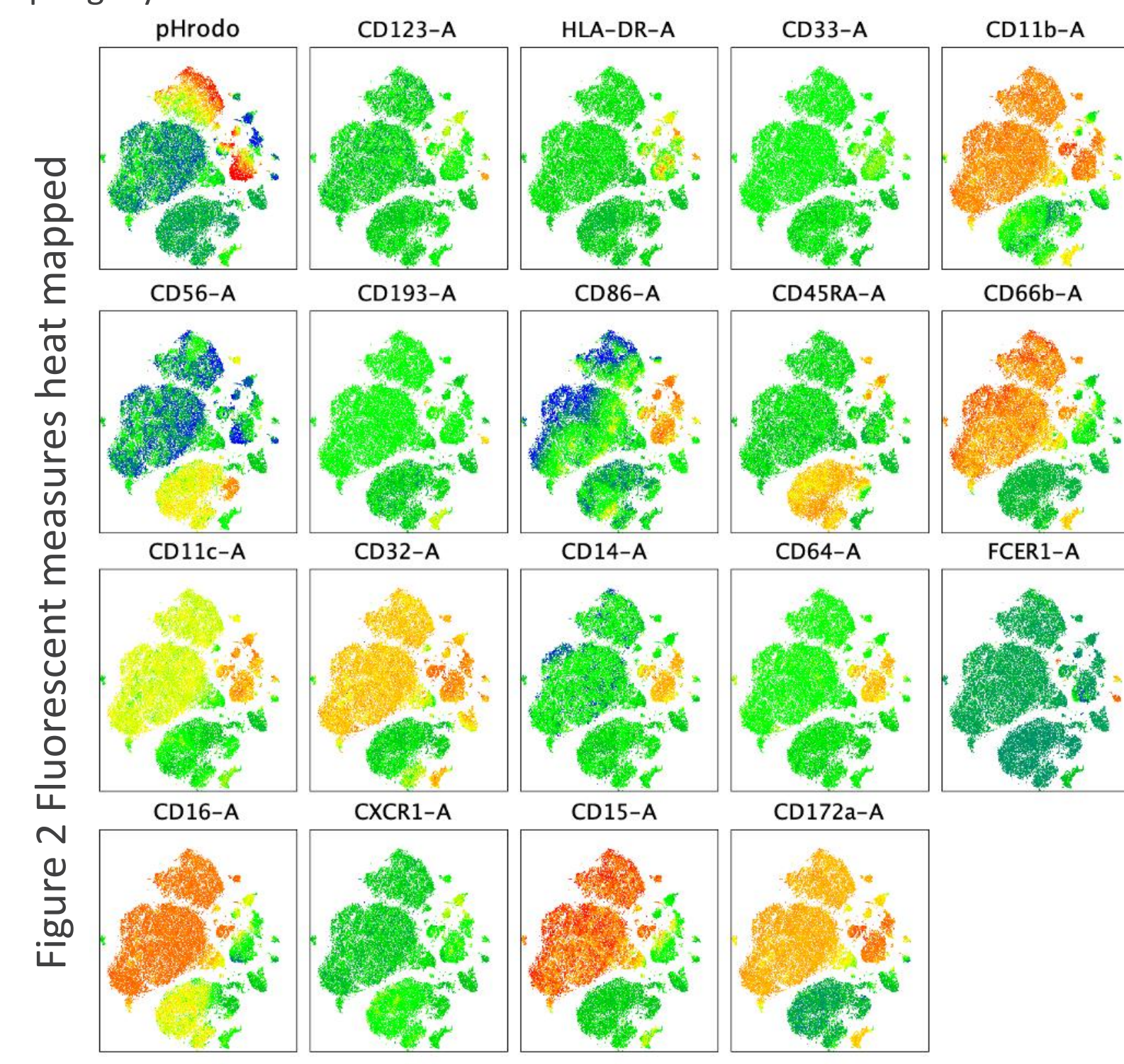
Data were clustered twice using the Phenograph algorithm, once with only the fluorescent panel (minus the dump channel) listed in Figure 1a, and once with the fluorescent panel + the image measurements. Dimensionality reduction was performed using optSNE and the same set of parameters as clustering. Heat mapping and Marker Enrichment Modeling (MEM) were used to understand the factors that led to unique cluster identification. Evaluation of cluster goodness was performed via a dip test and Taylor Index, both contained in the Euclid plugin to FlowJo.

Exemplar images were collected from each cluster and expanded using a nearest neighbor algorithm within CellView Lens. Finally, to make this workflow extensible we used the HyperFinder algorithm on the extended image sets to create a series of sortable gates so that one may recapitulate this population on the BD FACSDiscover™ sorter.

## Results

### 2A Clustering results

Heatmapping of the optSNE results revealed that there were 5 clusters that contained cells expressing pHrodo and that these clusters had heterogeneous expression of several markers thought related to phagocytosis



MEM was applied to the 5 clusters (Pictured in Figure 1) that were pHrodo+ to identify the differentiating markers. MEM scores reveal that CD11b, CD15, CD16, CD56, CD 66b, CD172 defined the differences between these clusters.

- Phenograph.Lin~4 ( CD66b-A+10 CD11b-A+2 CD16-A+2 CD15-A+2 )
- Phenograph.Lin~5 ( CD66b-A+9 CD16-A+3 CD11b-A+2 CD15-A+2 )
- Phenograph.Lin~10 ( CD66b-A-10 CD16-A-5 eGFP-A+2 CD11b-A+2 CD172a-A+2 )
- Phenograph.Lin~17 ( CD11b-A+5 CD66b-A+5 CD32-A+2 CD16-A+2 CD172a-A+2 )
- Phenograph.Lin~26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 )

### 2A Population evaluation

Statistical analysis using the identified fluorescent markers and image derived parameters revealed identifiable phenotypes with varied morphological factors that may contribute to identification as differing clusters.

	Median Area (µm²)	Median Volume (µm³)	Median Radial Moment (µm²)	Median Short Axis Moment (µm²)	Median Long Axis Moment (µm²)	Median Eccentricity	Median Diffusivity (µm²/s)	Median Total Intensity (AU)	Median Max Intensity (AU)	Median Radial Moment (µm²)	Median Short Axis Moment (µm²)	Median Long Axis Moment (µm²)	Median Eccentricity
Cluster 4	1446	5.517	4.282	2.711	3.071	0.11	1.782	20.5	0.12	75254	1446	5.517	4.282
Cluster 5	1446	5.517	4.282	2.711	3.071	0.11	1.782	20.5	0.12	75254	1446	5.517	4.282
Cluster 10	1446	5.517	4.282	2.711	3.071	0.11	1.782	20.5	0.12	75254	1446	5.517	4.282
Cluster 17	1446	5.517	4.282	2.711	3.071	0.11	1.782	20.5	0.12	75254	1446	5.517	4.282
Cluster 26	1446	5.517	4.282	2.711	3.071	0.11	1.782	20.5	0.12	75254	1446	5.517	4.282

- The stats imply that the clusters are:
- Cluster 4 CD15+, CD16+, CD66b+ Neutrophils >Tot eGFP <radial eGFP
  - Cluster 5 CD15+, CD16+, CD66b+ Neutrophils <Tot eGFP >radial eGFP
  - Cluster 10 CD11b<sup>DM</sup>, CD172+ Monocytes, < SSC size, tot SSC, etc.
  - Cluster 17 CD11b+, CD16 Macrophages, > morphological features
  - Cluster 26 CD56+, CD172- NK Cells, smallest sizes, signals

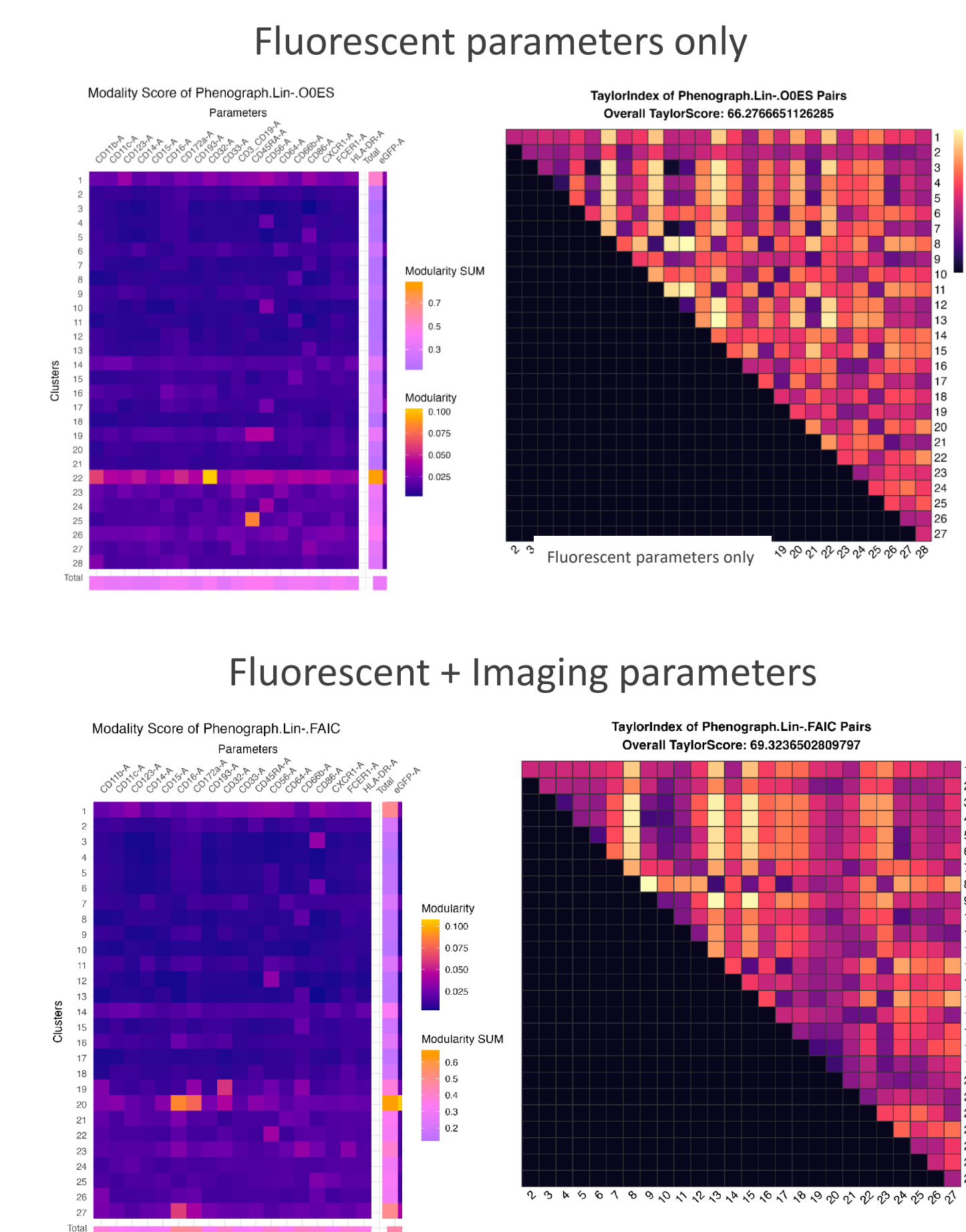
- Using the images shown in Figure 1 helps us understand what some of these differences in measurement mean.
- Cluster 4 and 5 are two halves of a similar population and the images reveal that they are similar neutrophils in different stages of digesting the *E. coli*.
  - Cluster 10 includes primarily larger cells attached to smaller cells. CD11b is an indicator of macrophage adhesion, and it appears we can see it here.
  - Cluster 17 includes primarily large cells connected to each other. It appears we have images of cell communication captured.
  - Cluster 26 includes the smallest cells both by numbers and visibly, adding confirmation that we have identified NK cells.

Hyperfinder, an algorithm that recapitulates a selected population using traditional gates for sorting purposes, was run on all clusters with an average F-measure of 0.84

- HyFl\_BUV615-A vs eGFP target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_BUV395-A vs Radial Moment (eGFP) target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_BUV661-A vs Alexa Fluor 647-A target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_BV510-A vs Size (eGFP) target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_Short Axis Moment (eGFP) vs Eccentricity (eGFP) target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_R718-A vs BV750-A target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_LightLoss (Violet)-A vs BV480-A target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A
- HyFl\_RY586-A vs eGFP target=1\_Phenograph.Lin~00ES\_26 ( CD66b-A-3 CD16-A-3 CD11b-A-2 ) runID=FF0E8A

### 2A Cluster assessment

Clustering results with and without imaging parameters were calculated and Euclid was run on both outcomes to compare the modularity (whether a cluster has a single peak per parameter) and Taylor index (the ratio of the distance between cells within a cluster and the other clusters). We observed a small improvement in both modularity and Taylor index using the image parameters.



## Conclusions

We have demonstrated that the complexity of an active phagocytotic system can be characterized more effectively by including imaging as an additional modality, and then automatically classifying clusters using an algorithm-driven workflow.

- Clustering is an effective way to identify populations in high-dimensional data, and image parameters improved both the quality of clusters and allowed to visualize morphological differences that may impact cluster outcomes.
- Morphology can reveal nuance of cell type difficult to capture with fluorescence only
- MEM and Hyperfinder were effective tools to expedite an analysis and make it extensible
- Euclid can be a useful way to evaluate a clustering outcome.

### References

- PeacoQC: Peak-based selection of high-quality cytometry data. *Cytometry A*. 2022 Apr; 101(4):325-338. doi: 10.1002/cyto.a.24501
- Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. doi: 10.1038/s41467-019-13055-y
- Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*. Volume 162, Issue 1, 184 – 197
- Characterizing cell subsets in heterogeneous tissues using marker enrichment modeling. *Nat Methods*. 2017 Jan 30;14(3):275-278. doi: 10.1038/nmeth.4149

BD Flow cytometers are Class 1 Laser Products. For Research Use Only. Not for use in diagnostic or therapeutic procedures. Trademarks are the property of their respective owners. © 2025 BD, BD, the BD Logo and all other trademarks are property of Becton, Dickinson and Company. BD-148272 (v1.0) 0425