Pairing high-parameter spectral flow cytometry with CITE-seq

Using a novel automated artificial intelligence-based analysis platform to characterize immunophenotypes and cell states

Zoey Zhou[1], Andrew McDavid[1], Kurt Van Gunst[1], Ashley L. Wilson[1], Denise Allen[1], Dillon Hammill[1], Evan Greene[1], Cherie Green[1]

ozette

BOOTH #733

ozette.com

BACKGROUND

Spectral flow cytometry and Cellular Indexing of Transcriptomes and Epitopes sequencing (CITE-seq) both enable high parameter single-cell surface protein analysis. Flow cytometry provides a time and resource efficient assessment of millions of cells, while CITE-seq throughput is typically limited to profiling of 10^4 to 10^5 cells per replicate. However, the CITE-seq assay generates paired single-cell RNA-seq data that may provide value to researchers investigating gene expression. Analyzing flow cytometry and CITE-seq data often requires use of disparate analysis tools and platforms, and integrating results can be a complex and resource intensive process. Here we apply Ozette Discovery™, an artificial intelligence-based analysis platform [1], to analyze paired spectral flow cytometry and CITE-seq datasets generated from peripheral blood mononuclear cells (PBMC) from healthy donors with known cytomegalovirus (CMV) status.

FIGURE 1

Experimental design and workflow

Peripheral blood mononuclear cells (PBMC) were isolated from 8 healthy donors with known CMV status and cryopreserved.

PBMC were characterized by spectral flow cytometry using a 48-color pan-immune profiling panel on a 5-laser Cytek Aurora™ full spectrum cytometer. Output files were unmixed using single color reference controls, and preprocessing was performed to select for single, live PBMC.

CITE-seq was performed on the same set of PBMC using the 10X Genomics™ 3' Gene Expression v3.1 kit and a 143-target antibody panel with 6 isotype controls. The CITE-seq panel was composed of the Biolegend TotalSeq-B Universal Cocktail and a titrated 9-plex custom antibody mix, such that the contents of the CITE-seq panel overlap the 48-color flow cytometry panel. Single-cell RNA-seq and antibody-derived tag sequencing libraries were generated and sequenced on an Illumina NovaSeq™. A series of preprocessing steps were performed to select for single, live, PBMC. Background adjustment and normalization of antibody-tag derived counts were performed.

Data from both assays was uploaded to the Ozette Discovery™, which was used to determine positive marker expression thresholds and identify immunophenotypes based on cell surface expression and detection frequency.

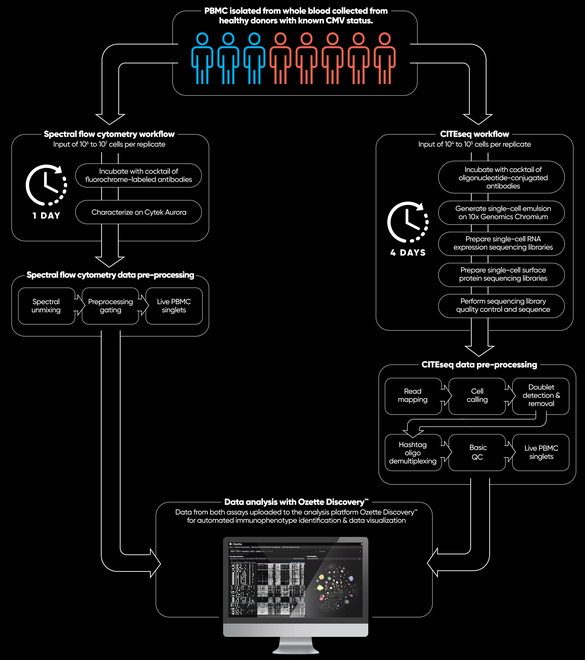


FIGURE 2

Spectral flow cytometry results in significantly higher cell profiling throughput

Using spectral flow cytometry, a total of 10.6M cells were profiled across all samples. CITE-seq profiled approximately 81K cells across all samples.

FIGURE 3

Higher number of robust immunophenotypes and proportion of rare phenotypes are identified from spectral flow cytometry compared to CITE-seq

A greater number of immunophenotypes were identified from (A) spectral flow cytometry data than from (B) CITE-seq data, shown as single-cell embedding visualizations where each uniquely colored island represents a robust cell phenotype. A greater number of rare phenotypes were identified in spectral flow cytometry data, shown as (C) distribution of phenotype abundance. Positive correlation was observed in the abundance of the union of all phenotypes in detected in each assay, when comparing across assays (D).

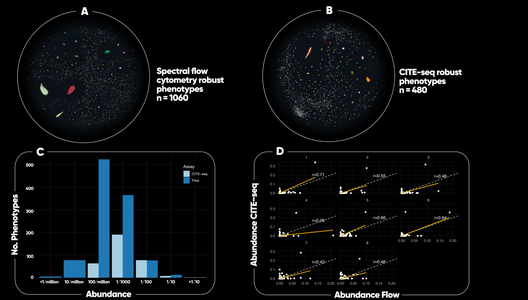


FIGURE 4

Similar frequencies of major immune cell lineages identified from spectral flow cytometry and CITE-seq

Abundance of major immune cell lineage groups, defined by expression of key markers, identified from spectral flow cytometry and CITE-seq data, grouped by donor. CD4+ T cells predominate.

Figure 4b: Correlation of major lineage abundance between spectral flow cytometry and CITE-seq. Flow sorted cells tend to exhibit higher abundance of CD4+ and CD8+ T cells.

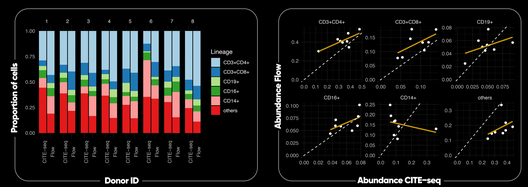


FIGURE 5

Differentially abundant immunophenotypes identified when stratified by donor CMV status

(A) Differential abundance of CD3+ CD4+ CD2+ CD95+ KLRG1+ T cell phenotypes identified in CMV+ status samples in flow cytometry and CITE-seq data. Each heatmap row represents a phenotype and each column represents a replicate, with n = 43 replicates from 8 donors. The plus sign in the marker columns indicates positive marker expression.

(B) The abundance CD3+ CD4+ CD2+ CD95+ KLRG1+ T cells was higher in CMV+ status samples in both flow cytometry and CITE-seq data.

(C) A similar set of differentially abundant CD3+ CD4+ CD2+ CD95+ KLRG1+ T cell phenotypes associated with CMV status was detected in the CITE-seq data at a lower degree of statistical significance, shown as a volcano plot.

FIGURE 6

Backgating visualizations reveal similar protein detection profiles between spectral flow cytometry & CITE-seq

Identification of analogous CD3+ CD4+ CD2+ CD95+ KLRG1+ T cells from spectral flow and CITE-seq by gating with Ozette Cell Discovery. The thresholds for positive and negative expression are shown in each scatter plot, and the population of interest (light blue) is superimposed onto the set of all PBMCs (grey).

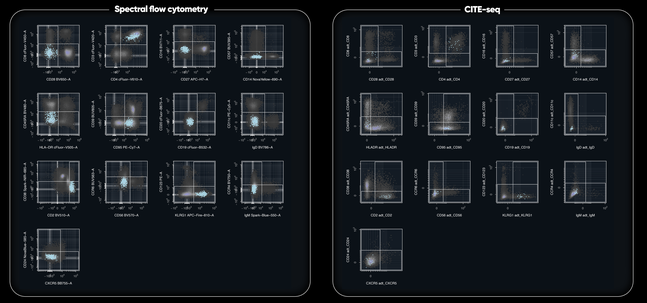


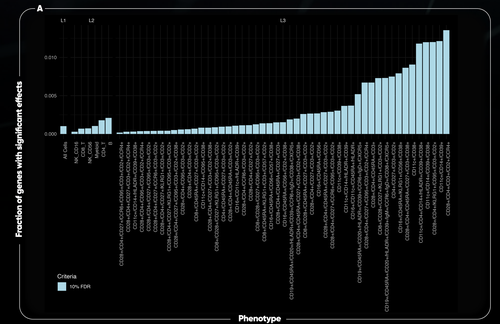
FIGURE 7

Using specific phenotypes discovered from spectral flow cytometry to inform CITE-seq differential expression analysis reveals distinct gene expression profiles

(A) A higher proportion of genes with effects significant at 10% FDR were identified in Ozette Discovery phenotypes (L3) when differential gene expression between CMV status was tested using Limma on pseudobulked RNA-seq counts from single cells. Less granular lineages defined by expression of canonical markers is (L2) and in pseudobulked data consisting of all cells (L1) are also shown.

(B) Distinct gene expression profiles are identified when stratifying by protein expression phenotypes (L3), compared to canonical marker defined lineages (L2). Top differentially expressed genes and their effects (t-statistic) in various phenotypes, in aggregated lineages, and all cells.

(C) RNA expression levels in selected genes were plotted by donor CMV status. Cells were aggregated into All Cells (3888 events; median donor), CD4 T cells (1993 events; median donor) and CD3+ CD4+ CD2+ CD28+ CD95+ KLRG1+ (17 events; median donor).



SUMMARY

1. Spectral flow cytometry workflow is able to profile significantly more cells compared to CITE-seq, and results in identification of a greater number of robust immunophenotypes and improved detection of rare phenotypes.
2. The identification of differentially abundant robust immunophenotypes from flow cytometry data can be used to inform analysis of CITE-seq gene expression data, and identify distinct gene expression profiles associated with specific protein expression phenotypes.
3. Immunophenotypes associated with chronic viral infection were detected in both spectral flow cytometry and CITE-seq data, but flow cytometry data provided additional depth and a higher degree of statistical power.
4. Ozette Discovery streamlines the analysis of both CITE-seq and spectral flow cytometry data on a unified platform, reducing the time needed to uncover biological insights and advance cancer immunotherapy research.

REFERENCES & ETHICS

References:

1. Allen D, Weaver M, Prokopchuk S, Lekschas F, Jiang M, Finak G, Greene E, McDavid A. Protein-based cell population discovery and annotation for CITE-seq data identifies cellular phenotypes associated with critical COVID-19 severity. BioRxiv 584720 [Preprint]. 2024 [cited 2024 June 24]. https://doi.org/10.1101/2024.03.14.584720

Ethics approval:

All donors provided informed consent using Institutional Review Board (IRB) approved consent forms.

Affiliations:

[1] Ozette Technologies, Seattle, WA, USA