

Databases and ontologies

IsoAnalytics: a single-cell proteomics web server

Suzette N. Palmer ^{1,2,3}, Andrew Y. Koh^{1,4,5,*}, Xiaowei Zhan ^{3,5,*}

¹Division of Hematology/Oncology, Department of Pediatrics, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

²Department of Biomedical Engineering, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

³Peter O'Donnell Jr. School of Public Health, Quantitative Biomedical Research Center, Center for the Genetics and Host Defense, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁴Division of Hematology/Oncology, Department of Microbiology, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁵Harold C. Simmons Comprehensive Cancer Center, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

*To whom correspondence should be addressed.

Associate Editor: Guoqiang Yu

Abstract

Motivation: Single-cell proteomics provide unprecedented resolution to examine biological processes. Customized data analysis and facile data visualization are crucial for scientific discovery. Further, user-friendly data analysis and visualization software that is easily accessible for the general scientific community is essential.

Results: We have created a web server, *IsoAnalytics*, that gives users without computational or bioinformatics background the ability to directly analyze and interactively visualize data obtained from the Isoplexis single cell technology platform. We envision this open-sourced web server will increase research productivity and serve as a free, competitive alternative for single-cell proteomics research.

Availability and implementation: *IsoAnalytics* is free and available at: <https://cdc.biohpc.swmed.edu/isoplexis/> and is implemented in Python, with all major browsers supported. Code for *IsoAnalytics* is free and available at: https://github.com/zhanxw/Isoplexis_Data_Analysis.

Contact: andrew.koh@utsouthwestern.edu or xiaowei.zhan@utsouthwestern.edu

Supplementary information: [Supplementary data](#) are available at *Bioinformatics Advances* online.

1 Introduction

Single-cell proteomics is a high-throughput technology that enables quantitative protein profiling of individual cells and for each sample, cytokine secretion profiles for ~1000 live cells are calculated. The Isoplexis platform utilizes flow-cell and multiplex ELISA technology to allow for the detection of up to 32 cytokines from individual immune cells. This fully automated and optimized technology serves as an alternative to flow cytometry and mass cytometry (cytometry by flight of time). The Isoplexis technology has provided key insights for a wide variety of biological processes, including cancer, immunology and oncology (Abbas *et al.*, 2021; Baer *et al.*, 2022; Kwak *et al.*, 2013).

Currently, the only available analysis software for Isoplexis assays is IsoSpeak, which is provided with the machine but offers limited user analyses and options. To offer users' greater flexibility and customizability for Isoplexis data analyses, we have created a web server, *IsoAnalytics*. *IsoAnalytics* provides expanded data analysis, visualization options and contains four categories of analysis and visualization: Clustering, Dimensionality Reduction, Polyfunctionality and Statistics. Our web server is completely interactive and does not require any computational or bioinformatics knowledge or coding ability to analyze and visualize data.

2 Implementation and application

IsoAnalytics was created using Dash (Dash Python, 2022; Plotly, 2015) and can be accessed through <https://cdc.biohpc.swmed.edu/isoplexis/>. All figures generated by *IsoAnalytics* are interactive and customizable. For example, zoom and rotation manipulation are available for each visualization. Selection of conditions or variables to be visualized is included. Additionally, all visualizations can be exported as PNG images.

The *IsoAnalytics* website lands to the 'Overview' tab. It gives descriptions for each data analysis provided. Additionally, the cytokines for each Isoplexis single cell secretome assay and the corresponding dominant functional group classifications are provided for easy access. The next tab is 'Upload' and allows users to directly load data (CSV or Excel format files) onto the website. The user must then select the correct single-cell secretome assay and then select the appropriate experimental condition or conditions to analyze. The order of this selection will determine the order of these conditions on the graphs. After selection of analysis conditions, the user has the option to scale, normalize and/or filter the data. Lastly, the user can select which individual cytokine will be analyzed, which are located on the Clustering and Statistics pages. At any time, the user can return to the 'Upload' tab

and change which treatment conditions and individual cytokines are analyzed.

The data analysis and visualization tabs are Clustering, Dimensionality Reduction, Polyfunctionality and Statistics. Under 'Clustering', hierarchical clustering across all cytokines and cells and user selection of cytokines only, or cells only are performed, and clustered heatmaps are generated for the selected treatment condition(s) (Fig. 1A). Underneath this section, for the selected cytokine from the 'Upload' tab, hierarchical clustering across all cytokines and cells and user selection of cytokines only or cells only are performed for samples that have non-zero values for the selected

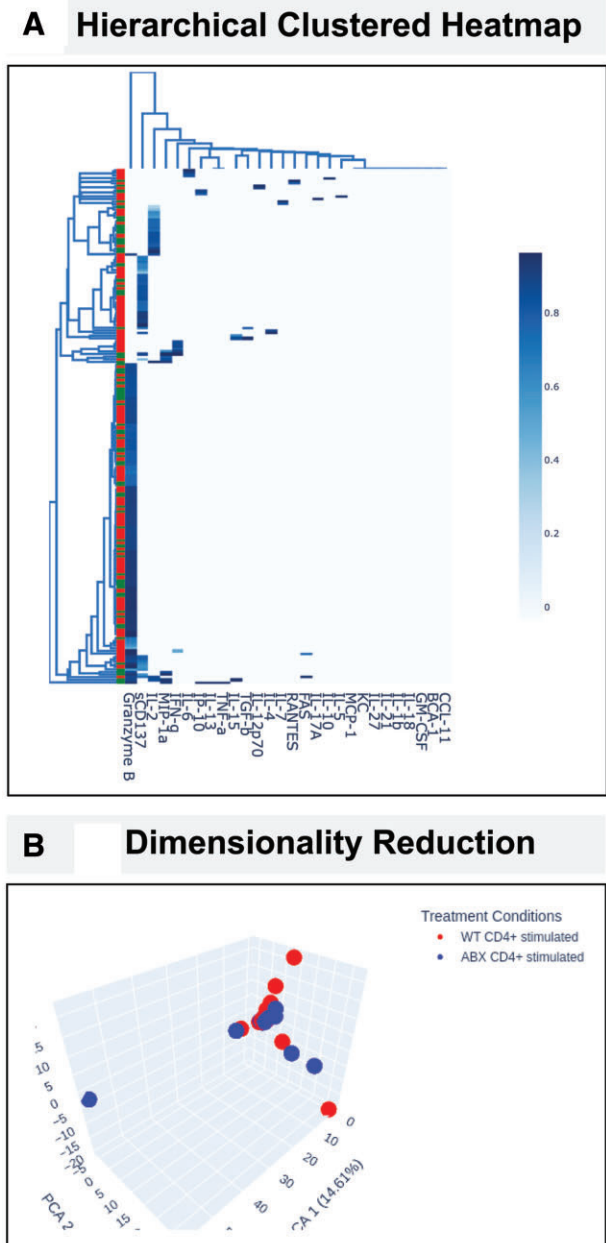


Figure 1. IsoAnalytics webserver provides interactive Hierarchical Clustering and Dimensionality Reduction analyses. **(A)** Hierarchical clustering on Isoplexis data is performed and visualized using a clustered heatmap. The user has the option to visualize clustering across all provided data or can select to view clustering based on samples containing a specific cytokine. **(B)** Dimensionality reduction analysis PCA can be visualized in 3D (top) or 2D

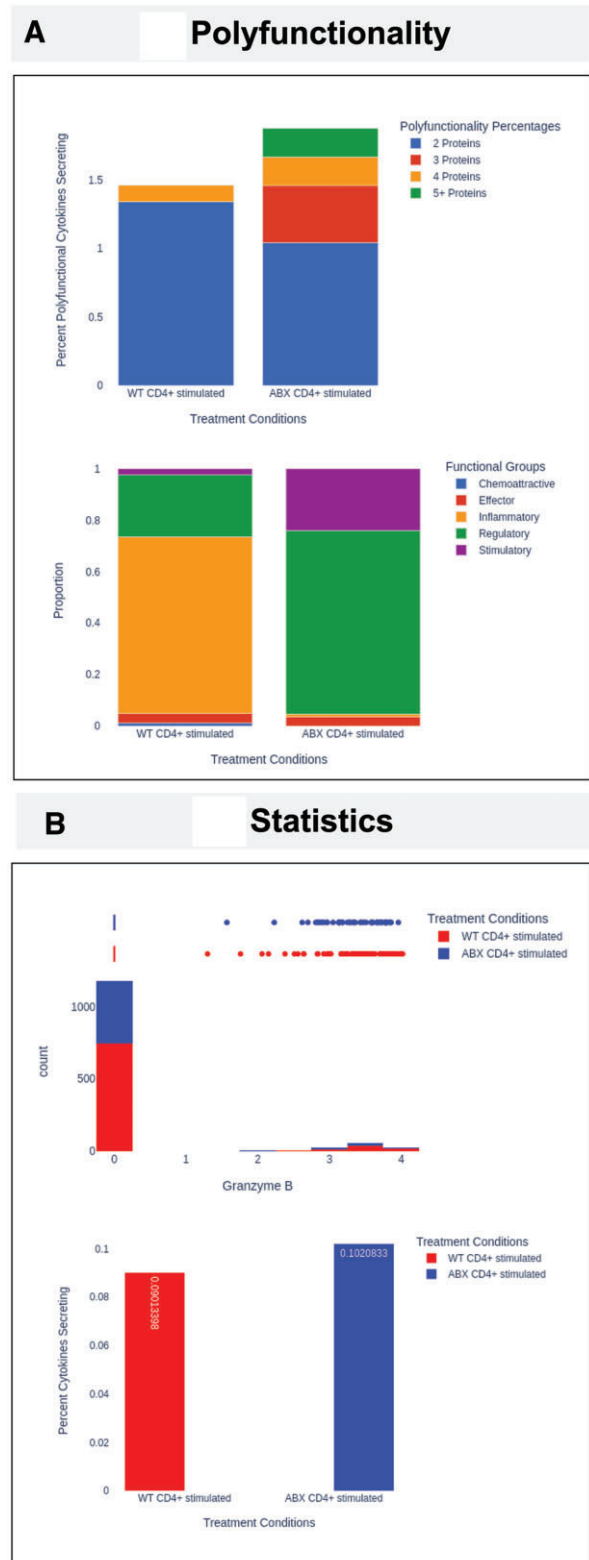


Figure 2. IsoAnalytics webserver provides Polyfunctionality and Statistics analysis. **(A)** The stacked bar graph (top) displays the number of polyfunctional cells (cells expressing more than on cytokine) for each treatment condition. The stacked bar graph (bottom) displays the functional groups for each treatment condition. **(B)** The histogram and box plot (top) displays the distribution of the user-selected cytokine and is plotted based on treatment conditions. The user also has the option to visualize the data as a violin or rug plot, instead of the boxplot. The bar plot (bottom) shows non-zero proportions based on the treatment conditions. More detailed descriptions of the analyses and visualizations are available in the IsoAnalytics [Supplementary Material](#)

cytokine, with accompanying clustered heatmaps generated (Fig. 1A).

The ‘Dimensionality Reduction’ tab displays 2D or 3D visualization of Standard Scalar Normalized PCA (Pearson, 1901; Pedregosa *et al.*, 2011) (Fig. 1B) and t-distributed stochastic neighbor embedding (t-SNE) (Pedregosa *et al.*, 2011; Van der Maaten and Hinton, 2008). For the t-SNE algorithm, the user can modify the algorithm by selecting a different perplexity and number of iterations. The ‘Polyfunctionality’ tab displays the number of polyfunctional cells and a stacked bar graph displaying the percent polyfunctional cells for each treatment condition (Fig. 2A). The Dominant Functional Groups for each treatment condition are displayed using stacked bar graphs (Fig. 2A). The user can select whether to view these data as absolute abundance or as a proportion. Additionally, all data from this tab are available to the user and can be exported as a csv file.

The last tab, ‘Statistics’ displays the Isoplexis data distribution and statistical tests. At the top, non-zero proportions for each treatment condition and cytokine are displayed as an interactive bar graph. The individual cytokine statistics are shown for the selected data or individual treatment conditions. The user can also calculate statistical significance by using the percent cytokines secreting (Non-Zero Proportion) test or the Kolmogorov–Smirnov test (Virtanen *et al.*, 2020) (Fig. 2B). The distribution of the data is displayed using a histogram and density plot, where values for each treatment condition are calculated (Fig. 2B). Above the histogram, the user can also select whether to view the data as a boxplot, violin plot or a rug plot.

3 Results

With the development of novel technology, such as the Isoplexis single-cell functional proteomics platform, new data analysis and visualization tools that are accessible to a wide variety of scientists are essential. We believe IsoAnalytics fills this niche, allowing users without computational or bioinformatics backgrounds to directly and interactively explore data produced by Isoplexis technology. We envision this website will complement the currently available IsoSpeak software, allowing users more options for data analysis and visualization. In future versions, we will include other Isoplexis analyses, such as Codeplex, and will continue to add novel assays that are developed for the Isoplexis platform.

Acknowledgements

The authors would like to thank other members of the Koh lab for crucial feedback that allowed the successful development of the website.

Conflict of Interest

A.Y.K. is a consultant for Prolacta Biosciences. He received research funding from Prolacta Biosciences and Novartis. He is a co-founder of Aumenta Biosciences.

Funding

This work was supported by National Institutes of Health [grant numbers CA231303 to A.Y.K., AI123163 to A.Y.K.]; Crow Family Fund (to A.Y.K.); the University of Texas Southwestern Medical Center and Children’s Health Cellular and ImmunoTherapeutics Program (to A.Y.K.); National Institute of Allergy and Infectious Disease [grant numbers AI169298 to X.Z., 5T32AI005284-43 to S.N.P.]; National Human Genome Research Institute [grant number HG011035 to X.Z.]; and National Institute of General Medical Sciences [grant number GM126479 to X.Z.].

References

- Abbas,H.A. *et al.* (2021) Single-cell polyfunctional proteomics of CD4 cells from patients with AML predicts responses to anti-PD-1-based therapy. *Blood Adv.*, **5**, 4569–4574.
- Baer,C. *et al.* (2022) CCL22 mutations drive natural killer cell lymphoproliferative disease by deregulating microenvironmental crosstalk. *Nat. Genet.*, **54**, 637–648.
- Dash Python (2022). *Dash Python*. <https://dash.plotly.com/> (3 January 2023, date last accessed).
- Kwak,M. *et al.* (2013) Single-cell protein secretomic signatures as potential correlates to tumor cell lineage evolution and cell-cell interaction. *Front. Oncol.*, **3**, 10.
- Pearson,K. (1901) LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.*, **2**, 559–572.
- Pedregosa,F. *et al.* (2011) Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.*, **2**, 2825–2830.
- Plotly (2015). *Collaborative Data Science*. Plotly Technologies Inc., <https://plot.ly> (3 January 2023, date last accessed).
- Van der Maaten,L. and Hinton,G. (2008) Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.*, **9**, 2579–2605.
- Virtanen,P. *et al.*; SciPy 1.0 Contributors (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods.*, **17**, 261–272.