

Preprint

<https://research.libd.org/spatialLIBD>
Advance Access Publication Date: Day Month Year
Application Note



Gene Expression

spatialLIBD: an R/Bioconductor package to visualize spatially-resolved transcriptomics data

Brenda Pardo^{1,2}, Abby Spangler², Lukas M. Weber³, Stephanie C. Hicks³, Andrew E. Jaffe², Keri Martinowich², Kristen R. Maynard², Leonardo Collado-Torres^{2,*}

¹Licenciatura de Ciencias Genómicas, Escuela Nacional de Estudios Superiores Unidad Juriquilla, Universidad Nacional Autónoma de México, Querétaro, 76230, México; ²Lieber Institute for Brain Development, Johns Hopkins Medical Campus, Baltimore, MD 21205, USA ³Department of Biostatistics, Johns Hopkins School of Public Health, Baltimore, MD 21205, USA

*To whom correspondence should be addressed.

Associate Editor: XXXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Spatially-resolved transcriptomics has now enabled the quantification of high-throughput and transcriptome-wide gene expression in intact tissue while also retaining the spatial coordinates. Incorporating the precise spatial mapping of gene activity advances our understanding of intact tissue-specific biological processes. In order to interpret these novel spatial data types, interactive visualization tools are necessary.

Results: We describe *spatialLIBD*, an R/Bioconductor package to interactively explore spatially-resolved transcriptomics data generated with the 10x Genomics Visium platform. The package contains functions to interactively access, visualize, and inspect the observed spatial gene expression data and data-driven clusters identified with supervised or unsupervised analyses, either on the user's computer or through a web application.

Availability: *spatialLIBD* is available at bioconductor.org/packages/spatialLIBD.

Contact: lcolladotor@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The rise of spatially-resolved transcriptomics technologies (Ståhl *et al.*, 2016; Xia *et al.*, 2019) are providing new opportunities to answer questions about the structure and function of complex intact tissues (Maynard *et al.*, 2020, 2021; Kuppe *et al.*, 2020). However, this exciting opportunity requires the development of new computational methods and interactive software. These tools have to accommodate the analysis of transcriptomic data with spatial coordinates and high resolution images into previously developed methods for single cell RNA-seq analysis (scRNA-seq), which lacked these components.

Currently there are limited options for interactive exploration of spatially-resolved transcriptomics data (Dries *et al.*, 2020). The 10x Genomics Loupe Browser (10x Genomics, 2020) utilizes a file created

by the Space Ranger data pre-processing pipeline (10x Genomics, 2020) and performs marker gene analysis, as well as unsupervised graph-based and *k*-means clustering. Loupe visualizes both the spatially-resolved transcriptomics data along with the high resolution imaging data to export gene expression maps. An alternative to Loupe is the Giotto pipeline (Dries *et al.*, 2020), which contains both a data analysis module and a data visualization module. The data analysis module has many functions ranging from pre-processing to more advanced analysis such as characterizing cell-cell interactions. The visualization module contains an interactive workspace for exploring multiple layers of information. However, Loupe and Giotto currently do not support visualizing more than one tissue section at a time, which is useful for comparing replicates and annotating observed spots on the Visium array across samples. Furthermore, while unsupervised clustering methods are widely developed (Zhao *et al.*, 2020; Biancalani *et al.*, 2020), manual annotation of spots using known marker genes is important, as well as cross-sample dimension reduction techniques such as UMAP and t-SNE.

2Brenda Pardo^{1,2}, Abby Spangler², Lukas M. Weber³, Stephanie C. Hicks³, Andrew E. Jaffe², Keri Martinowich², Kristen R. Maynard², Leonardo Collado-Torres^{2,*}

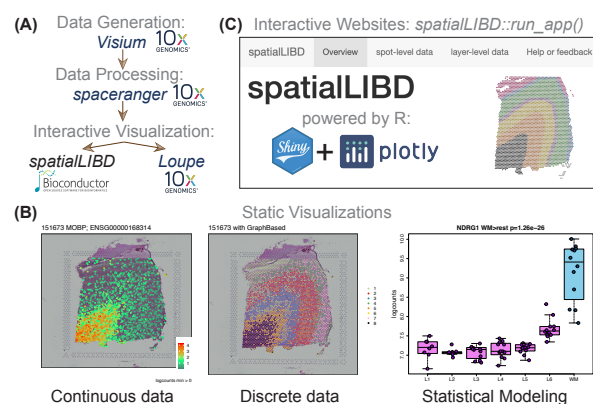


Fig. 1. Overview of spatialLIBD: (A) Using spatially-resolved transcriptomics data generated with the 10x Visium platform and processed with spaceranger, both from 10x Genomics, spatialLIBD leverages the R/Bioconductor ecosystem and presents an alternative for interactive visualization to Loupe browser. (B) spatialLIBD supports static data visualizations of both continuous and categorical measurements along with the histology images, as well as visualizing results from downstream analyses. (C) spatialLIBD supports interactive websites for exploring the data powered by shiny and plotly.

2 Results

spatialLIBD is an R/Bioconductor (Huber *et al.*, 2015) package that provides functions to interactively visualize and explore spatially-resolved transcriptomics data, including functionality to manually annotate spots. It was developed initially for visualizing multiple Visium tissue sections from the human brain (Maynard *et al.*, 2021), yet it is flexible enough to be used with other Visium data from different tissue types (Supplementary File 1, Figure 1A).

The primary functions in *spatialLIBD* allow users (i) to visualize the spot-level spatial gene expression data, or any other continuous variable, and data-driven clusters or any categorical variable, with the histology image in the background (Figure 1B), and (ii) to inspect the data interactively, either on the user’s computer or through web hosting services such as shinyapps.io (Figure 1C). In addition, *spatialLIBD* enables users to visualize multiple samples at the same time, to export all static visualizations as PDF files or all interactive visualizations as PNG files, as well as all result tables as CSV files. Supplementary File 1 shows how to read in data processed with Space Ranger (10x Genomics, 2020) into R, make plots or create an interactive website with *spatialLIBD*.

Finally, *spatialLIBD* also supports the interactive exploration of data from Maynard *et al.* (2021). For example, using the customized website spatial.libd.org/spatialLIBD, researchers can perform spatial registration of their scRNA-seq clusters. This data can also be used for developing new analytical methods (Biancalani *et al.*, 2020; Zhao *et al.*, 2020).

3 Methods

spatialLIBD is designed to work with *SpatialExperiment* R/Bioconductor objects (Righelli *et al.*, 2021). It uses *shiny* (Chang *et al.*, 2021) and *plotly* (Sievert, 2020) for the interactive website, and *ggplot2* (Wickham, 2016) for the static images.

4 Discussion

spatialLIBD benefits from the open-source Bioconductor ecosystem (Huber *et al.*, 2015) and relies on the infrastructure provided by *SpatialExperiment* (Righelli *et al.*, 2021). It provides user-friendly functionality for visualizing continuous and categorical measurements along with the tissue images from spatially-resolved transcriptomics data from projects involving one or more samples measured on the 10x Genomics Visium platform. *spatialLIBD* provides a simple interactive website, which can be used for sharing data on the web, as well as

manually annotating spots. *spatialLIBD* has limitations (Supplementary File 1) that are inherent to the methods used to implement it, such as (i) the memory per user required by a server for hosting the web application, (ii) response speeds for the interactive views due to the number of spots, (iii) the resolution of the images displayed limiting the usefulness to magnify specific spots, and (iv) customization of the web application by the end user. Despite these limitations, software solutions like *spatialLIBD* enable faster data exploration and insights in spatially-resolved transcriptomics research projects (Maynard *et al.*, 2021) and can serve as testing ground for ideas for future work, some of which are commercially available (BioTuring, 2021; 10x Genomics, 2020).

Acknowledgements

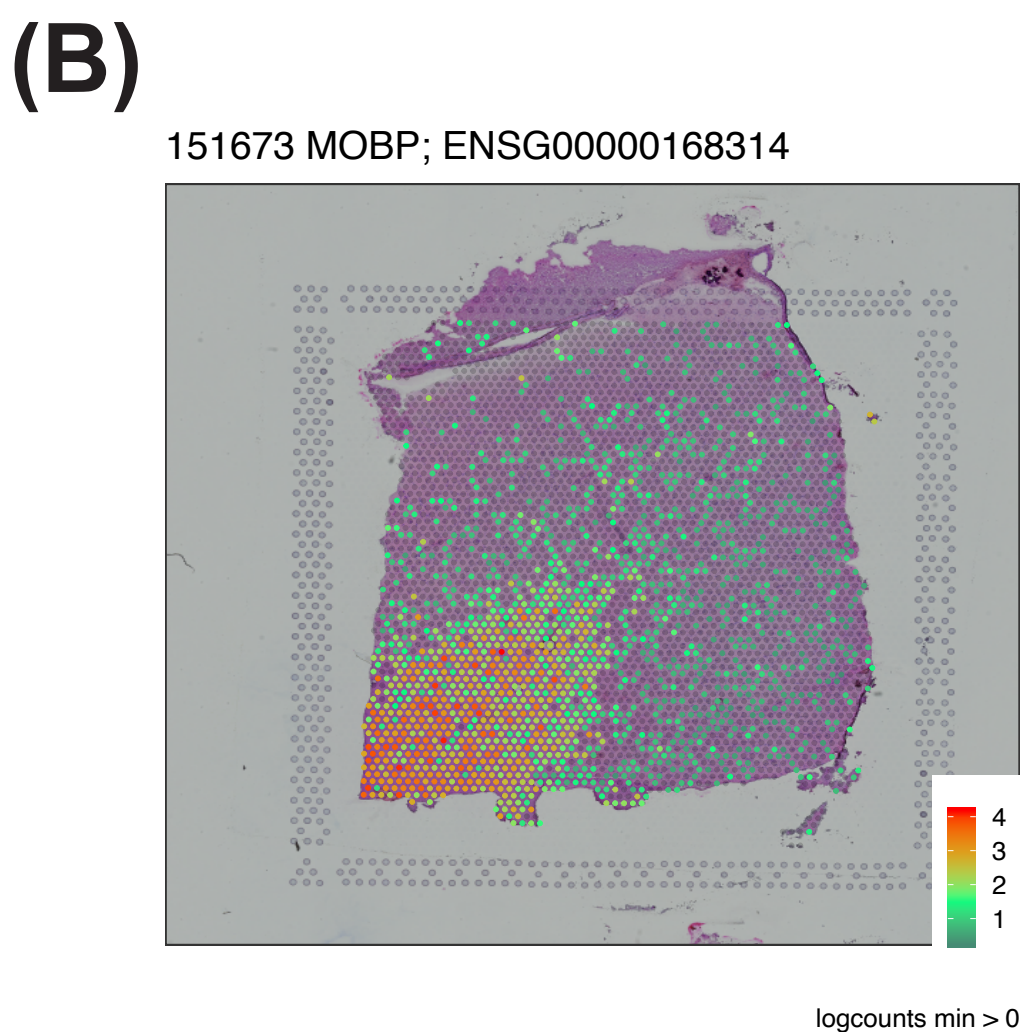
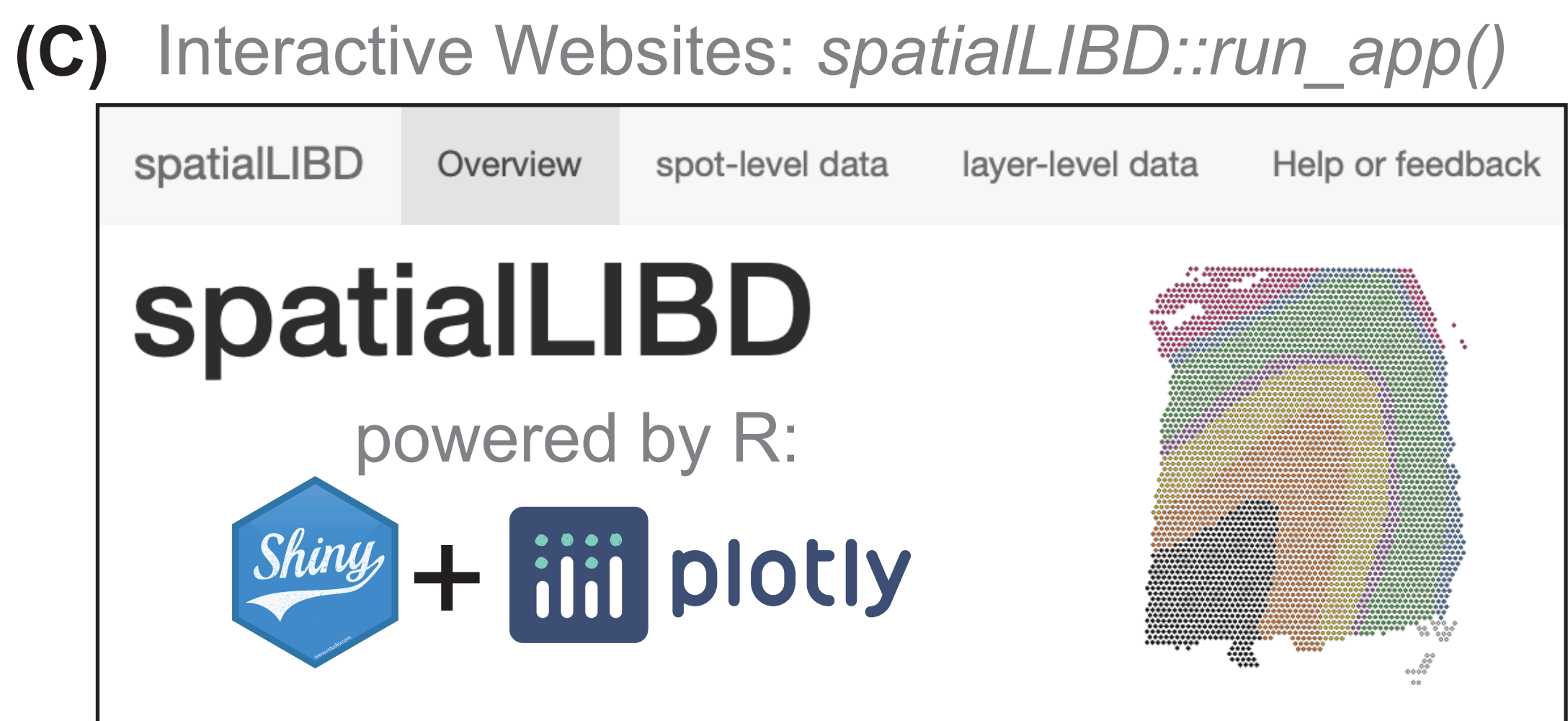
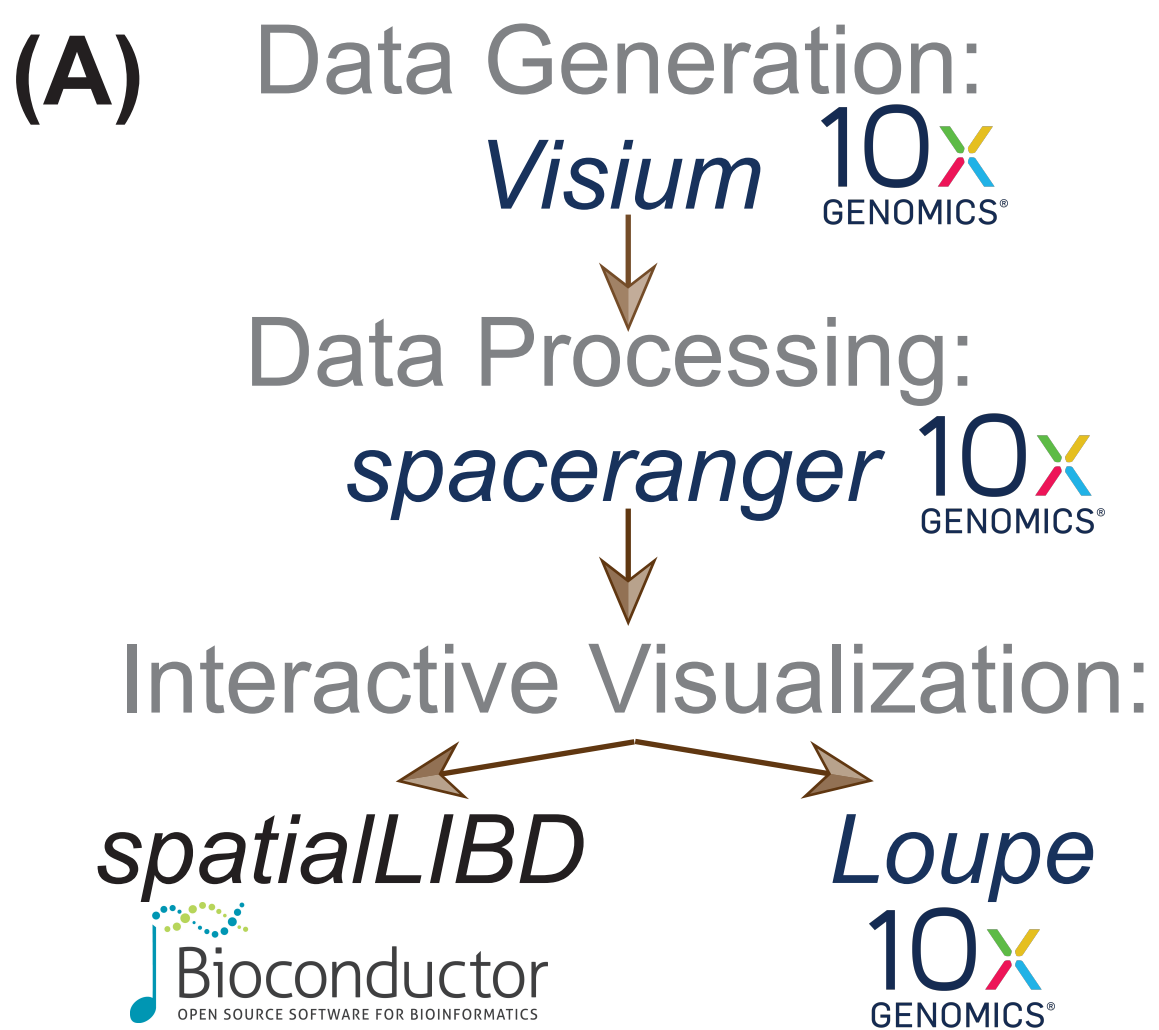
Jesús Vélez Santiago and Ricardo O Ramírez-Flores from Julio Saez-Rodríguez’s group at Heidelberg University tested the applicability of *spatialLIBD* to new datasets and improve the interactive website code.

Funding

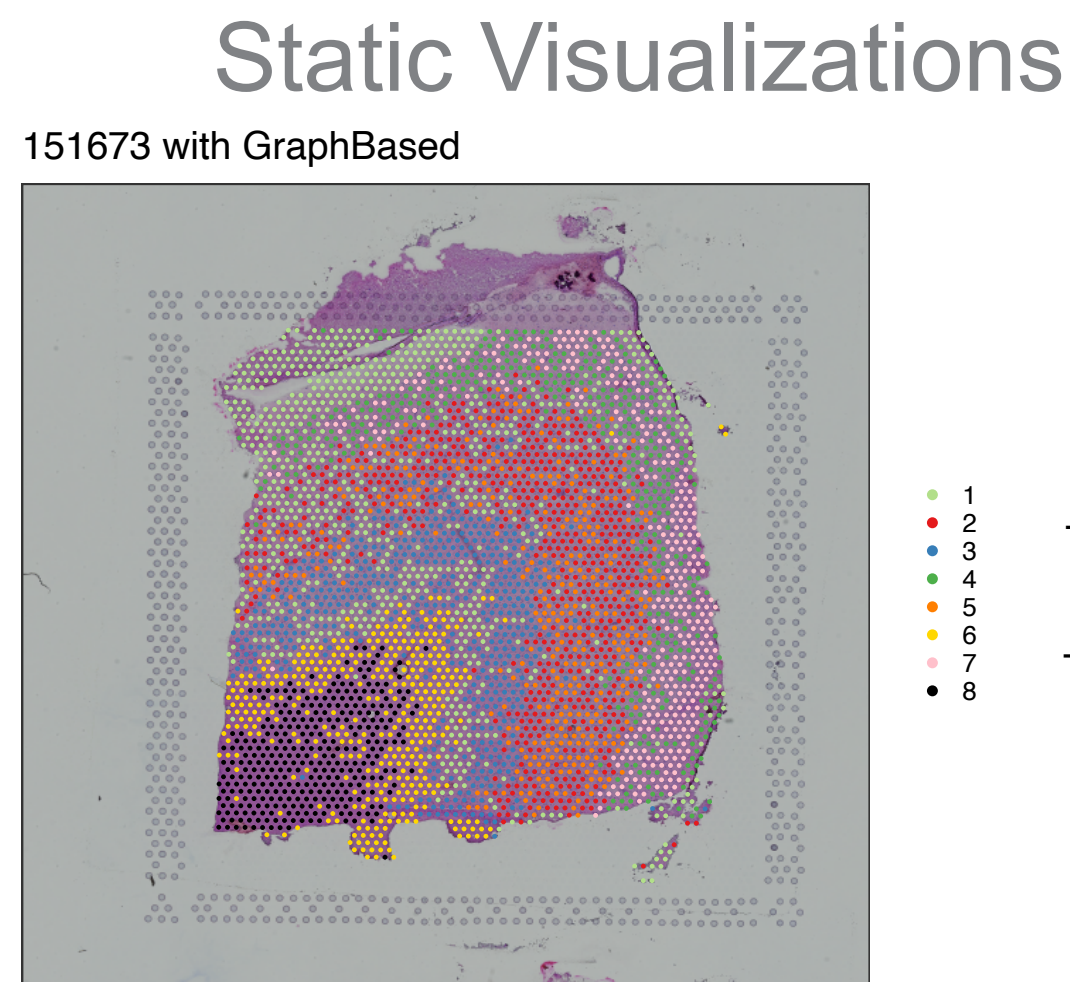
This work was supported by the Lieber Institute for Brain Development, National Institutes of Health grant U01 MH122849 (LMW, SCH, AEJ, KM, KRM, LCT), and CZF2019-002443 (LMW, SCH) from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation.

References

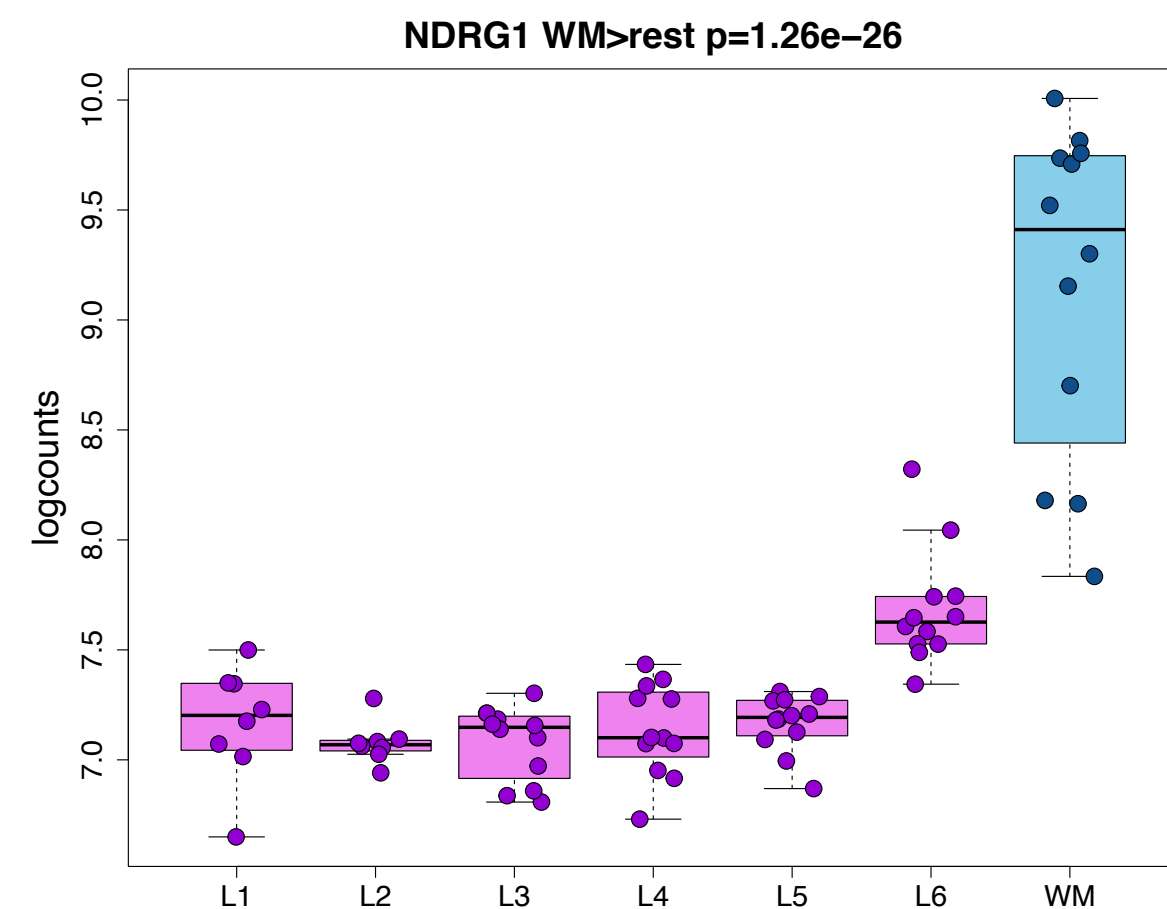
- 10x Genomics (2020). Space ranger and loupe browser. <https://support.10xgenomics.com/spatial-gene-expression/software>.
- Biancalani, T. *et al.* (2020). Deep learning and alignment of spatially-resolved whole transcriptomes of single cells in the mouse brain with tangram. *bioRxiv*.
- BioTuring (2021). Bbrowser. <https://bioturing.com/bbrowser/spatial-transcriptomics>.
- Chang, W. *et al.* (2021). shiny: Web Application Framework for R. R package version 1.6.0.
- Dries, R. *et al.* (2020). Giotto, a toolbox for integrative analysis and visualization of spatial expression data. *bioRxiv*.
- Huber, W. *et al.* (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, **12**(2), 115–121.
- Kuppe, C. *et al.* (2020). Spatial multi-omic map of human myocardial infarction. *bioRxiv*.
- Maynard, K. R. *et al.* (2020). Spatial transcriptomics: putting genome-wide expression on the map. *Neuropsychopharmacology*, **45**(1), 232–233.
- Maynard, K. R. *et al.* (2021). Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature Neuroscience*, **24**(3), 425–436.
- Righelli, D. *et al.* (2021). SpatialExperiment: infrastructure for spatially resolved transcriptomics data in R using Bioconductor. *bioRxiv*.
- Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC.
- Stahl, P. L. *et al.* (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, **353**(6294), 78–82.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Xia, C. *et al.* (2019). Spatial transcriptome profiling by merfish reveals subcellular rna compartmentalization and cell cycle-dependent gene expression. *Proceedings of the National Academy of Sciences*, **116**(39), 19490–19499.
- Zhao, E. *et al.* (2020). Bayesspace enables the robust characterization of spatial gene expression architecture in tissue sections at increased resolution. *bioRxiv*.



Continuous data



Discrete data



Statistical Modeling