

CeSTAAN: An atlas of *C. elegans* adult neurons for fast queries of single-nucleus RNA sequencing data

Jonathan St. Ange¹, Henry Langmack², Coleen T. Murphy^{1§}

¹LSI Genomics, Molecular Biology, Princeton University, Princeton, New Jersey, United States

²Princeton High School, Princeton New Jersey

[§]To whom correspondence should be addressed: ctmurphy@princeton.edu

Abstract

RNA-sequencing provides rich transcriptomic data about cell types, and new methods allow characterization at the single-cell level. For neurons, single-nucleus RNA sequencing (snSeq) is the gold standard, as this technique retains neuron identity. We carried out snSeq on sorted nuclei from *C. elegans* neurons of Day 1 adult wild-type (N2) and *daf-2* mutants, and males and genotypically-matched hermaphrodites. To provide the field with an easily accessible atlas of this information, we created *C. elegans* Single-nucleus Transcriptomic Atlas of Adult Neurons (CeSTAAN) (<https://cestaan.princeton.edu/>), described here with examples. CeSTAAN will allow the *C. elegans* field to access this adult neuron transcriptional information.

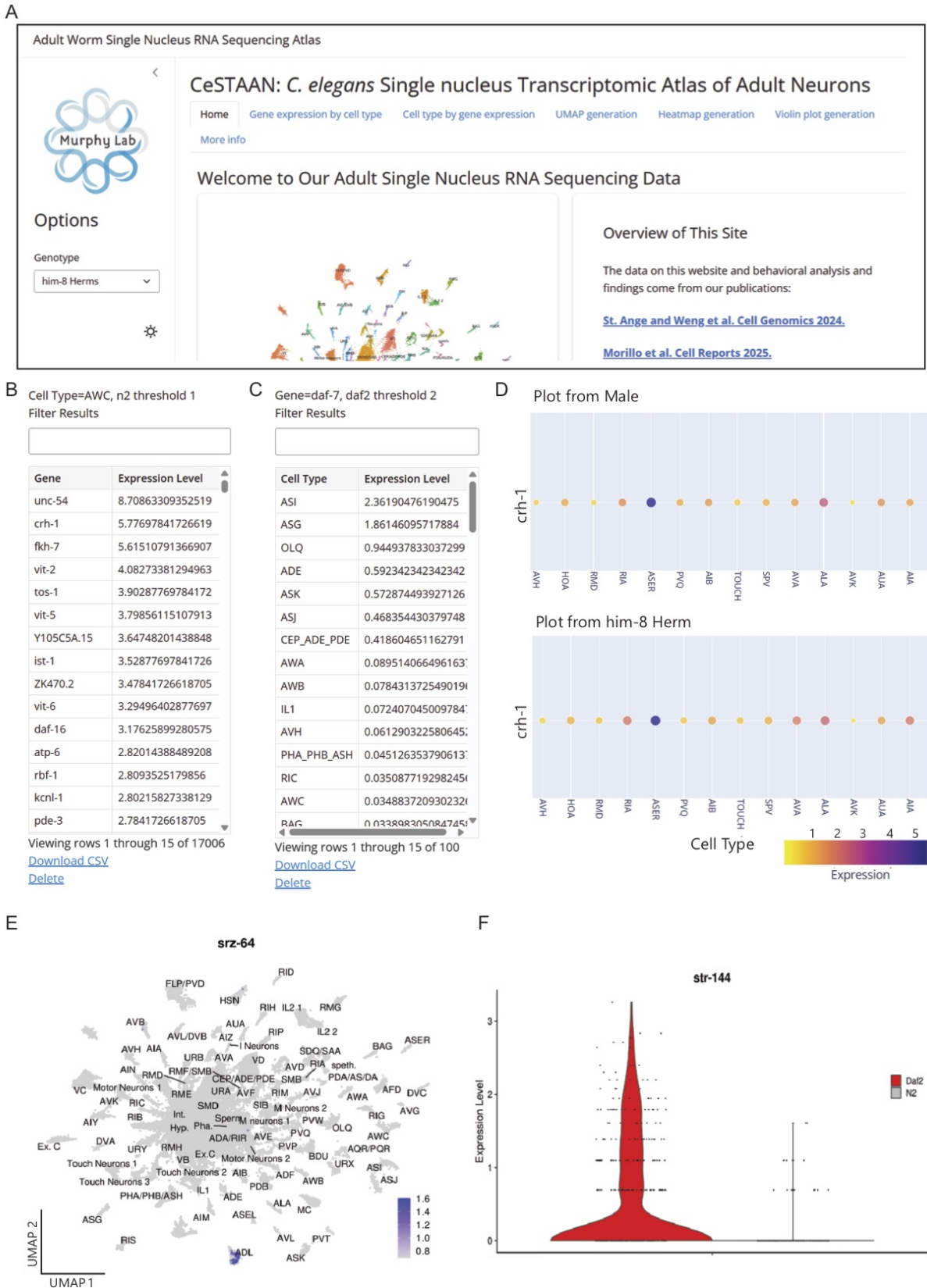


Figure 1. CeSTAAAN Website functionalities:

A) An excerpt from the homepage of the CeSTAAAN website. On the side panel, the user can select the genotype or sex they would like to see transcriptomic information for as well as click between functionality tabs. These tabs allow for various user queries. **B)** ‘Gene Expression by Cell Type’ tab allows the user to input a cell type and receive a downloadable gene expression table. **C)** ‘Cell Type by Gene Expression’ tab allows the user to input a gene and receive a

downloadable table of that gene's expression level across the nervous system. The website is also capable of plot generation, e.g., **D**) Excerpts from the interactive heatmap tab containing information on expression levels and percent expression of genes in all neurons captured in these data sets. The user can also input a gene or multiple genes to generate **E**) UMAP plots or **F**) violin plots from the data. Dot size is correlated with fraction of expressing cells in a given cluster.

Description

The data available on CeSTAAN comes from our recent manuscripts using the single-nucleus RNA sequencing technique, as shown in St. Ange & Weng et al. (2024)¹ and Morillo et al. (2025)².

Shiny python serves as the main structure for the website, while SQL databases quickly serve queries back to the user and R generates visuals for presentations if the user desires. The user can pick from several different tabs: Gene expression by cell type, Cell type by gene expression, UMAP generation, Heatmap generation, Violin plot generation, and more info. The output of the expression tabs are dynamic searchable tables that are available for any genotype or sex at any of four different gene expression and percent cell expression thresholds. UMAPs and Violin Plots generate from the .rds file containing the Seurat objects³ and running R script. Finally, the Heatmap generation tab uses the SQL databases and python to generate an interactive heatmap that the user can zoom in and out of.

All plots and tables generated by the website are downloadable, and the more info tab contains a download button for our Seurat object files that are the fully processed objects from our papers.

Methods

All datasets used for this site as well as scripts used to create the site are available at: <https://github.com/svenmh/murphy-lab-project>. These datasets come from St. Ange & Weng et al. (2024) as well as Morillo et al. (2025). All generation, processing, and quality control metrics for these datasets are available in their respective manuscripts. The raw read sequence data from these manuscripts are available on NCBI as BioProjects [PRJNA1027859](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1027859) and [PRJNA1195922](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1195922) respectively.

The expression values reported on the website are normalized Single-Cell Transform (SCT) expression values. SCT is a standard normalization process for single-cell data that addresses the large amount of zeros in the data and the variation between cells⁴. SCT values are very small where an average expression value in the [N2/daf-2](#) dataset is ~0.05 and an average expression value in the male/hermaphrodite dataset is ~0.03. Importantly, however, these datasets were not all normalized, together so adult [N2](#) animals can only be compared to the genotypically- and age-matched [daf-2](#) animals, while adult male data can only be compared to the genotypically- and aged-matched hermaphrodites.

The data for the website are stored as SQL databases containing average normalized expression values and percent expression values for each dataset. These databases are generated from excel sheets using a python script, which are all uploaded on github. For plot generation, the main Seurat objects are used. RDS files containing Seurat objects for the Adult Day 1 [N2/daf-2](#) dataset and the Adult genotypically-matched male/hermaphrodite datasets are also on github. These files are opened using R, and PNGs are generated and handed back to the user. The main website uses a shiny python script as the interface for user input and reactive tables and plots. It connects to python and R files to respond to user inputs and quickly serve query results back to the user.

Reagents

Software/Tool	Website	Version	Publication reference
Python	https://www.python.org/	v3.9.21	-
R	https://www.r-project.org/	v4.5.0	-
SQLite	https://www.sqlite.org/	v3.34.1	-
Seurat	https://satijalab.org/seurat/	V4.4.0	(Hao et al. 2021)
Shiny for Python	https://shiny.posit.co/py/	v1.4.0	-

Acknowledgements: We thank Sven Heinicke and Mark Schroeder of the Princeton University Genomics Computational Core for their help in setting up the web server for the application.

References

Hafemeister C, Satija R. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 20(1): 296. PubMed ID: [31870423](#)

Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al., Satija R. 2021. Integrated analysis of multimodal single-cell data. *Cell* 184(13): 3573-3587.e29. PubMed ID: [34062119](#)

Morillo KS, St Ange J, Weng Y, Kaletsky R, Murphy CT. 2025. Single-nucleus neuronal transcriptional profiling of male *C. elegans* uncovers regulators of sex-specific and sex-shared behaviors. *Cell Rep* 44(8): 116016. PubMed ID: [40682776](#)

St Ange J, Weng Y, Kaletsky R, Stevenson ME, Moore RS, Zhou S, Murphy CT. 2024. Adult single-nucleus neuronal transcriptomes of insulin signaling mutants reveal regulators of behavior and learning. *Cell Genom* 4(12): 100720. PubMed ID: [39637862](#)

Funding: Supported in part by National Science Foundation Graduate Research Fellowship Program (United States) DGE-2039656 to JMS, the Princeton Learning Lab Program to HL & CTM and by the Simons Collaboration on Plasticity in the Aging Brain (SCPAB) to CTM.

Supported by NSF GRFP DGE-2039656 to Jonathan St. Ange.

Author Contributions: Jonathan St. Ange: conceptualization, data curation, methodology, visualization, writing - original draft, writing - review editing, software, investigation. Henry Langmack: software, methodology. Coleen T. Murphy: conceptualization, funding acquisition, resources, supervision, writing - review editing.

Reviewed By: Anonymous

Nomenclature Validated By: Anonymous

WormBase Paper ID: WBPaper00068848

History: Received July 31, 2025 **Revision Received** November 3, 2025 **Accepted** November 24, 2025 **Published Online** November 25, 2025 **Indexed** December 9, 2025

Copyright: © 2025 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: St. Ange J, Langmack H, Murphy CT. 2025. CeSTAAN: An atlas of *C. elegans* adult neurons for fast queries of single-nucleus RNA sequencing data. *microPublication Biology*. [10.17912/micropub.biology.001785](#)