LIVEtools: an R Package for Visualizing *C. elegans* Embryo Cell Tracking Data

Yuntian Gan[1], Sreekar Kutagulla[2], John Isaac Murray[1§], Priya Sivaramakrishnan[3§]

[1]Department of Genetics, Perelman School of Medicine, University of Pennsylvania

[2]Children's Hospital of Philadelphia Research Institute

[3]Center for Computational and Genomic Medicine, Children's Hospital of Philadelphia

[§]To whom correspondence should be addressed: jmurr@pennmedicine.upenn.edu; sivaramakp@chop.edu

Abstract

Automated cell tracking in the *C. elegans* embryo using the StarryNite software pipeline has been a powerful tool to quantify the embryonic lineage, to define the spatiotemporal dynamics of gene expression, and to characterize developmental defects in mutants. The raw lineage traces are typically curated and visualized using the Java-based tool AceTree. However, these legacy visualization tools are often hard to deploy, are unstable on modern computer systems, and lack flexibility, especially for batch processing. *LIVEtools* is an R package that aims to simplify and automate *C. elegans* cell-tracking data visualization.
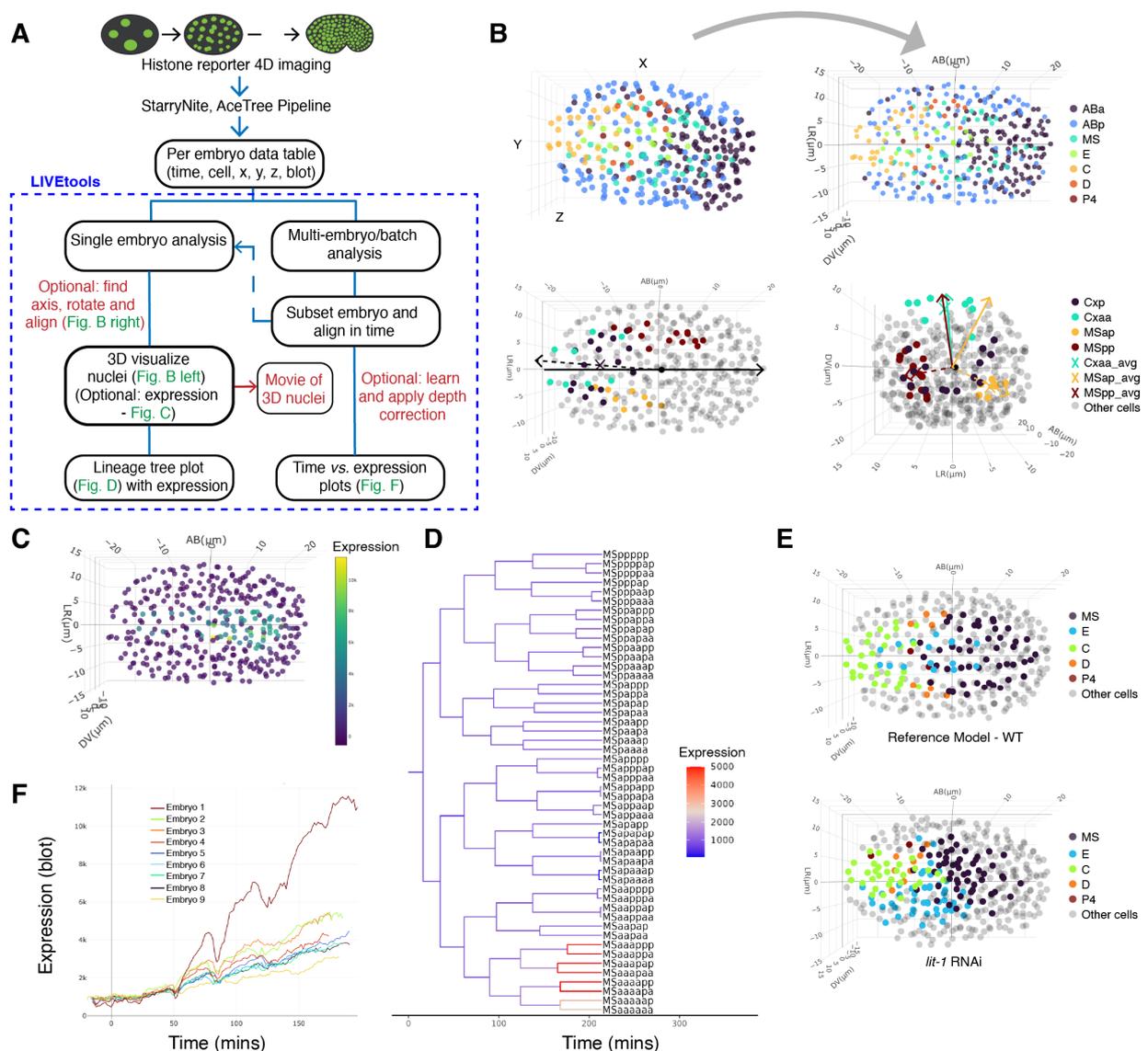
Figure 1. Illustration and outputs of LIVEtools functionalities:

(A) Flowchart describing the main and optional steps to utilize *LIVEtools* code. 4D confocal images are processed through StarryNite and AceTree to obtain a data table with information on "cell", "time", "x", "y" and "z" coordinates, and optionally gene expression, that serves as the starting point for *LIVEtools*-based analysis. Workflows exist to process a single embryo at a time or a group of embryos in batch as indicated. (B) Automated axis rotation. Top left: 3D visualization of a "raw" embryo dataset before rotation. Top right: the same embryo after rotation to align the AP axis with the x-axis and DV with the z-axis, using the lineages Cxp (posterior), Cxaa (dorsal), MSap (right) and MSpp (left) as axis landmarks. Bottom left (dorsal aspect) and right (anterior aspect) show 3D visualizations of post-rotated embryos with cell groups used for rotation colored as shown in the legend (cell colors are customizable). For DV rotation, left, right, and dorsal groups were specified, resulting in three candidate dorsal orientations based on each group (solid lines) which were averaged to guide the final rotation. (C) A rotated embryo 3D plot colored by mean reporter expression level. Selected lineages/cells can be highlighted and given custom marker shapes, with adjustments for color range (min/max) as needed. (D) Expression of a fluorescent reporter in the MS lineage (with user-defined root and end times) generated using the *plot_trees* function. (E) 3D positions of cells from the specified lineages in a wild-type (top) and *lit-1* RNAi-treated (bottom) embryo. (F) Expression level versus time for a *pha-4* expression reporter across many replicate embryos/conditions after batch processing and time alignment. Expression values are averaged across cells from arbitrary user-specified lineages (in this case ABalp, ABara, ABplpappp, ABprpappp and their descendants).

Description

Embryos of the nematode *Caenorhabditis elegans* develop through a fully mapped invariant lineage (Sulston et al., 1983). This lineage has proven to be a powerful experimental tool, and methods to automatically track the lineage from live imaging 4D microscopy data have allowed researchers to define the spatiotemporal dynamics of gene expression across the lineage (Araya et al., 2014; Ma et al., 2021; Mace et al., 2013; Murray et al., 2008, 2012; Sarov et al., 2012). Applying these approaches to mutants or other perturbed embryos can identify defects in the lineage and in cellular gene expression or position (Du et al., 2015; Ho et al., 2015; Moore et al., 2013; Murray et al., 2022; Rumley et al., 2022; Walton et al., 2015).

A common analysis pathway involves imaging embryos where all nuclei express a fluorescently tagged histone, and using the StarryNite analysis program for cell tracking (Bao et al., 2006). The resulting cell lineage tracks can be further curated and visualized in the Java program AceTree (Boyle et al., 2006). AceTree allows viewing expression, lineage, and cell positions, and fixing of any tracking errors made by StarryNite. AceTree also includes visualization tools, including exporting lineage trees as image files, displaying the 3D positions of cells as model projections, and plotting expression versus time along a lineage trajectory. However, these tools have limited flexibility and scalability due to their hard-coded nature and scale poorly due to the need to interact with a graphical user interface separately for each embryo. Furthermore, the original Java3D-based 3D viewer is incompatible with modern computers due to requirements for obsolete Java versions (Java 6) and libraries (Java3D). More recent updates to AceTree (Katzman et al., 2018) increased stability and replaced the obsolete 3D viewer with a new version based on WormGuides (Santella et al., 2015), but reproducible figure design in AceTree is still inefficient.

Here, we present a codebase, *LIVEtools*, that allows efficient and reproducible manipulation and visualization of *C. elegans* lineage tracing data in the widely used R statistical programming language. This toolkit works from a simple tabular data format for each embryo, as typically produced by the StarryNite pipeline, and can also directly parse the raw StarryNite format (.zip) files. The user can incorporate additional metadata such as voxel size to convert between pixel and micron coordinates, and timestamp data to convert ordinal time points to elapsed time. A Sulston-compatible cell search tool allows visualization of highly customizable cell selections. Since images can be acquired in any orientation, *LIVEtools* can automate rotation of 3D views to a common orientation (e.g. dorsal aspect) and correct reduced expression intensity due to imaging depth.

The main features (illustrated in the Figure) are:

1. Display interactive 3D models of nuclear coordinates at a given time point using *plotly* (with options to color by lineage or any numerical attribute such as gene expression). These models can be manipulated in a web browser rendered by *plotly* for discovery, or output to image files for figure generation.

2. Display a lineage tree with the option to color by reporter expression level.

3. Subset the embryo (for example, keeping only specific time points, or lineages).

4. Rotate nucleus coordinates to a common aspect by aligning a customizable set of lineages to represent anterior-posterior and dorsal-ventral axes, and aligning these with the x- and z- axes.

For batch processing, there are built-in functions to do the following (and custom batch processing is made possible using custom R scripts):

1. Simultaneously subset embryos.

2. Align the time between multiple embryos by developmental events (such as to compare expression after the birth of a particular blastomere).

3. Learn and correct for reporter signal loss over depth (z) due to imaging artifacts.

4. Plot reporter expression over time in the same trajectory between embryos to allow comparative expression analysis.

5. Save modified embryo data (rotated, depth-corrected).

LIVEtools enables early data inspection and efficient batch processing of arbitrary collections of embryo data. Use of common data processing will also simplify generation of reproducible analyses and figures for journal articles.

We illustrate these functions in (Fig 1. B-F). Fig. 1B illustrates 3D visualization of nuclei and the rotation function, using known or user-supplied cells to define relative positions to center the embryo. Defined orientations can be applied to all time points. Cells in the 3D plot can be colored by reporter expression (Fig. 1C), which can also be visualized through plotting of lineage trees (Fig. 1D). Average expression in specific lineages can be compared across conditions over time, after aligning embryos (Fig. 1F). 3D visualization can help interact with and identify position defects in mutant or perturbed embryos (Fig. 1E). The accompanying GitHub repository includes the code and examples to create these figures, as well as vignettes describing how to perform additional visualization.

Methods

The methods are included and documented in detail in the GitHub repository and conceptually summarized here.

File formats and loading

The user can load data from a StarryNite-format file or a file in "CD" format. CD files are comma-separated text files, where the first line has column headers and additional lines have cell-time points. Required columns are "cell", "time", "x", "y" and "z". By default, expression-plotting functions will use a column called "blot," but other quantitative columns with user-specified names can be selected as well. If a "TIME" file is provided, with columns including the ordinal time point and elapsed time, the true time values can be used for downstream plotting.

3D rotation

Embryo rotation is rigid-body, based on default or user-specified markers of the AP and DV axes such that the AP axis is aligned with the x-axis and DV with the z-axis. Briefly, at a specified time point, a transform matrix is defined that will align the AP axis (defined using PCA to identify the long axis of the embryo, with the specified cell groups ensuring anterior=positive x), and will rotate the embryo around this AP axis such that the specified dorsal and ventral cell group centroids are aligned with the z-axis (or optionally, left and right cell groups are aligned with the y-axis). Note that since the DV and LR axes are linked, providing markers for both axes will result in a solution that minimizes the total angular deviation of the provided cell groups from the y- and z-axis, likely meaning that neither will correspond perfectly to the corresponding axis. Finally, this transform matrix is applied either to the specified time point, or all time points to give a full-embryo rotation.

Depth correction

Depth correction of expression values relies on using replicate embryos (prior to 3D rotation to a common axis). *LIVEtools* learns an exponential decay model for the relationship of expression with z position (depth) across the replicate embryos provided, then uses this to correct expression of the specified embryos to a common reference depth, as previously described (Ma et al., 2021; Murray et al., 2008). Once a depth correction model is learned, it can be applied to other embryos collected using equivalent imaging methods where the decay of signal with depth is similar to that learned on the training set.

Plotting

LIVEtools uses *ggtree* to output lineage trees as vector graphics, and *plotly* to output 3D visualizations. These 3D visualizations can be manually viewed and manipulated or automatically exported to images using provided functions based on the *reticulate* package and *kaleido*.

LIVEtools is available on GitHub at https://github.com/johnmurraylab/LIVE_tools as an R package. The code and example data required to generate the images in the figure are available on GitHub at: https://github.com/johnmurraylab/LIVEtools-paper.

Extended Data

Description: v1.0 release of LiveTools. Resource Type: Software. File: LIVE_tools-micropub.zip. DOI: 10.22002/zay8w-3dn76

References

Araya CL, Kawli T, Kundaje A, Jiang L, Wu B, Vafeados D, et al., Snyder M. 2014. Regulatory analysis of the C. elegans genome with spatiotemporal. Nature. 512: 400-405. DOI: 10.1038/nature13497

Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. 2006. Automated cell lineage tracing in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 103: 2707-2712. DOI: 10.1073/pnas.0511111103

Boyle TJ, Bao Z, Murray JI, Araya CL, Waterston RH. 2006. AceTree: a tool for visual analysis of Caenorhabditis elegans. BMC Bioinformatics. 7: 275. DOI: 10.1186/1471-2105-7-275

Du Z, Santella A, He F, Shah PK, Kamikawa Y, Bao Z. 2015. The Regulatory Landscape of Lineage Differentiation in a Metazoan Embryo. Dev. Cell. 34: 592-607. DOI: 10.1016/j.devcel.2015.07.014

Ho VW, Wong MK, An X, Guan D, Shao J, Ng HC, et al., Zhao Z. 2015. Systems-level quantification of division timing reveals a common genetic. Mol. Syst. Biol. 11: 814. DOI: 10.15252/msb.20145857

Insley P, Shaham S. 2018. Automated C. elegans embryo alignments reveal brain neuropil position. PLoS One. 13: e0194861. DOI: 10.1371/journal.pone.0194861

Katzman B, Tang D, Santella A, Bao Z. 2018. AceTree: a major update and case study in the long term maintenance of. BMC Bioinformatics. 19 DOI: 10.1186/s12859-018-2127-0

Ma X, Zhao Z, Xiao L, Xu W, Kou Y, Zhang Y, et al., Du Z. 2021. A 4D single-cell protein atlas of transcription factors delineates. Nat. Methods. 18: 893-902. DOI: 10.1038/s41592-021-01216-1

Mace DL, Weisdepp P, Gevirtzman L, Boyle T, Waterston RH. 2013. A High-Fidelity Cell Lineage Tracing Method for Obtaining Systematic. G3. 3: 851-863. DOI: 10.1534/g3.113.005918

Moore JL, Du Z, Bao Z. 2013. Systematic quantification of developmental phenotypes at single-cell. Development. 140: 3266-3274. DOI: 10.1242/dev.096040

Murray JI, Bao Z, Boyle TJ, Boeck ME, Mericle BL, Nicholas TJ, et al., Waterston RH. 2008. Automated analysis of embryonic gene expression with cellular resolution. Nat. Methods. 5: 703-709. DOI: 10.1038/nmeth.1228

Murray JI, Boyle TJ, Preston E, Vafeados D, Mericle B, Weisdepp P, et al., Waterston RH. 2012. Multidimensional regulation of gene expression in the C. elegans embryo. Genome Res. 22: 1282-1294. DOI: 10.1101/gr.131920.111

Murray JI, Preston E, Crawford JP, Rumley JD, Amom P, Anderson BD, et al., Zacharias AL. 2022. The anterior Hox gene ceh-13 and elt-1/GATA activate the posterior Hox. PLoS Genet. 18: e1010187. DOI: 10.1371/journal.pgen.1010187

Rumley JD, Preston EA, Cook D, Peng FL, Zacharias AL, Wu L, Jileaeva I, Murray JI. 2022. pop-1/TCF, ref-2/ZIC and T-box factors regulate the development of. Dev. Biol. 489: 34-46. DOI: 10.1016/j.ydbio.2022.05.019

Santella A, Catena R, Kovacevic I, Shah P, Yu Z, Marquina Solis J, et al., Bao Z. 2015. WormGUIDES: an interactive single cell developmental atlas and tool for. BMC Bioinformatics. 16: 189. DOI: 10.1186/s12859-015-0627-8

Sarov M, Murray JI, Schanze K, Pozniakovski A, Niu W, Angermann K, et al., Hyman AA. 2012. A genome-scale resource for in vivo tag-based protein function exploration. Cell. 150: 855-866. DOI: 10.1016/j.cell.2012.08.001

Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev. Biol. 100: 64-119. DOI: 10.1016/0012-1606(83)90201-4

Walton T, Preston E, Nair G, Zacharias AL, Raj A, Murray JI. 2015. The Bicoid class homeodomain factors ceh-36/OTX and unc-30/PITX cooperate. PLoS Genet. 11: e1005003. DOI: 10.1371/journal.pgen.1005003

Conflicts of Interest: The authors declare that there are no conflicts of interest present.

Author Contributions: Yuntian Gan: conceptualization, formal analysis, investigation, methodology, writing - original draft, visualization, validation, software. Sreekar Kutagulla: validation. John Isaac Murray: writing - review editing, validation, supervision, conceptualization. Priya Sivaramakrishnan: conceptualization, writing - review editing, validation, visualization, supervision, project administration, methodology.

Reviewed By: Amanda Zacharias, Anonymous

WormBase Paper ID: WBPaper00069256

Citation: Gan Y, Kutagulla S, Murray JI, Sivaramakrishnan P. 2026. *LIVEtools*: an R Package for Visualizing *C. elegans* Embryo Cell Tracking Data. microPublication Biology. 10.17912/micropub.biology.001933