Protocol

Automated Lineage and Expression Profiling in Live *Caenorhabditis elegans* Embryos

John Isaac Murray and Zhirong Bao

Describing gene expression during animal development requires a way to quantitatively measure expression levels with cellular resolution and to describe how expression changes with time. Fluorescent protein reporters make it possible to measure expression dynamics in live cells by time-lapse microscopy, but it can be challenging to identify expressing cells in complex tissues and to compare expression across organisms. This protocol describes how to use automated lineage analysis to identify cells in *Caenorhabditis elegans* embryos expressing fluorescent reporters and how to quantify that expression with cellular resolution. Because *C. elegans* develops through an invariant pattern of cell divisions, every cell's identity and future fate can be predicted from its pattern of previous cell divisions. Automated analysis of images collected from embryos expressing a fluorescent histone transgene in all cells allows lineage tracing and cell identification. This provides a scaffold with which to describe expression of a second color reporter such as a fusion of a second fluorescent protein to a gene of interest or its regulatory sequences. These methods can also be used for analysis of reporter expression, cell division timing, and cell position in genetically perturbed embryos. The protocol describes how to prepare *C. elegans* strains containing nuclear-expressed fluorescent reporters, collect images of appropriate quality from embryos, perform automated lineage analysis, manually edit and curate the lineage, and, finally, extract and display reporter signals.

MATERIALS

It is essential that you consult the appropriate Material Safety Data Sheets and your institution's Environmental Health and Safety Office for proper handling of equipment and hazardous materials used in this protocol.

Reagents

C. elegans lineaging strains

Strains are available from the *Caenorhabditis* Genetics Center (Stiernagle 2006) and include RW10026 and RW10029 (*histone::GFP*) and RW10226 (*histone::mCherry*). RW10029 contains two green fluorescent protein (GFP) loci (*his-72::GFP* and *pie-1_{promoter}::H2B-GFP*). Although only the *his-72::GFP* locus is required for lineage analysis, strains without the *pie-1*-driven marker might require more manual editing of the divisions before the 12-cell stage. RW10026 is an alternative GFP strain for use when the reporter to be studied is linked to the GFP locus on chromosome V in RW10029. RW10226 also includes two loci (*pie-1_{promoter}::H2B-mCherry* and *his-72_{promoter}::H1.1-mCherry*). Both loci are necessary for full lineaging because the *his-72*-driven marker is not expressed at all until the 50-cell stage.

Adapted from *Imaging in Developmental Biology* (ed. Sharpe and Wong). CSHL Press, Cold Spring Harbor, NY, USA, 2011.

© 2012 Cold Spring Harbor Laboratory Press

Cite this article as *Cold Spring Harb Protoc*; 2012; doi:10.1101/pdb.prot070615

Reagents for mounting *C. elegans* embryos as described in **Mounting *Caenorhabditis elegans* Embryos for Live Imaging of Embryogenesis** (Bao and Murray 2011)

Reagents for propagation of *C. elegans* (e.g., NGM plates seeded with OP50 bacteria) (Hope 1999)

Equipment

Computer running Linux, UNIX, Macintosh OS X, or Microsoft Windows

Equipment for mounting *C. elegans* embryos as described in **Mounting *Caenorhabditis elegans* Embryos for Live Imaging of Embryogenesis** (Bao and Murray 2011)

Imaging setup

The major equipment required for this method is a microscope capable of acquiring four-dimensional (4D; i.e., three-dimensional + time) multicolor fluorescence images. It is challenging to acquire sufficiently high-quality images for lineage tracing throughout embryogenesis without killing the embryo or perturbing development. A suitable microscope system must be able to quantitatively and continuously control exposure, to increase exposure with depth in the embryo, and to decrease exposure with time (as the fluorescence signal becomes brighter). We have successfully used a Zeiss LSM 510 confocal laser-scanning microscope equipped with a 63× 1.4-numerical aperture (NA) Plan-Apochromat objective, a 30-mW multiline argon laser running at 35% power, and a 5-mW HeNe 543-nm laser, running Zeiss AIM software with the MultiTime macro for multiple time-series acquisition. It is also possible to generate suitable images with a Yokogawa spinning-disk microscope with a high signal-to-noise electron-multiplying charge-coupled device (EMCCD) camera and MetaMorph software (W. Mohler, pers. comm.). Many commercially available microscopes do not have software controls that allow enough flexibility, so it is important to check for this before purchasing a microscope that will be used for long-term live imaging of fluorescently labeled *C. elegans* embryos.

We use a temperature-controlled stage (Brook Industries, Lake Villa, IL) calibrated to ensure the embryo stays at 20°C during the experiment; this increases consistency between embryos because the overall pace of developmental rate is temperature dependent. Temperature >25°C causes abnormal *C. elegans* development.

Software and accessory programs

The software described in the procedure below—SNLauncher, StarryNite (Bao et al. 2006), AceTree (Boyle et al. 2006), and AceBatch—plus accessory programs are available from <http://waterston.gs.washington.edu/tools.html> for Linux, Macintosh OS X, and Microsoft Windows. These programs should all be extracted to a single directory, and require Java 1.5.06 or later and Java3D. In general, any modern (post-2005) desktop computer should be capable of running all these programs. The major difficulty is file storage because each embryo generates ~10 Gb of data. A centralized redundant file system with a backup is highly recommended if you will be analyzing more than one embryo.

METHOD

The procedure consists of several independent parts that can be performed at different times: strain generation, imaging, automated lineage analysis, manual lineage curation, and reporter signal analysis, each described separately. The starting point is a stable strain expressing a nuclear localized fluorescent reporter gene, and the result of the procedure is a quantitative measurement of the nuclear signal in each cell of the embryo over time.

Strain Generation

1. Using standard *C. elegans* genetics methods (Hope 1999; Stiernagle 2006), cross the reporter into a lineaging strain background.
 - i. To lineage red fluorescent protein (RFP) reporters (e.g., DsRed, mCherry, tdTomato, etc.), cross the reporter into the RW10029 background.
 - ii. To lineage GFP, yellow fluorescent protein, or cyan fluorescent protein reporters, cross the reporter into the RW10226 background.

Imaging

2. Mount embryos under a Vaseline-sealed coverslip by using the bead slurry method.

For details on this procedure, see **Mounting *Caenorhabditis elegans* Embryos for Live Imaging of Embryogenesis** (Bao and Murray 2011). Use Boyd buffer supplemented with 1% methyl cellulose, antibiotics, and 20-μm beads as the mounting solution.

It is important to mount only a small number (five or less) of young embryos. The embryos should be at the two-cell stage or earlier to allow time to begin imaging before the end of the four-cell stage; this is required for AceTree to recognize the identities of the founder cells.

3. Collect 4D images using the microscope.

The precise settings used for image collection will depend on the microscope used. In particular, the absolute level of excitation light needed is highly microscope dependent and can vary substantially between systems. Tables 1 and 2 list the settings used with a Zeiss LSM 510 microscope for GFP and RFP lineage channel imaging.

See Troubleshooting.

- i. Collect images with $\leq 1\text{-}\mu\text{m}$ z-spacing and ≤ 2 min between time points.
- ii. Use low excitation light intensity at the top of the embryo for both channels, increasing excitation to higher exposure at the bottom of the embryo to offset loss from scatter, absorption, and aberration.
- iii. Begin with relatively high exposure in early time points with reduced exposure in later time points; see Tables 1 and 2 for approximate reduction ratios.

If the particular microscope allows, some of the increases in exposure at early time points can be offset with reductions in imaging resolution (e.g., line skipping on a point-scanning microscope or increasing the pinhole diameter to increase sensitivity).

- iv. Use exposure levels that are constant with time (although still increasing with depth) for the expression channel to ensure that increases and decreases in expression intensity with time are meaningful.

Automated Lineage Analysis

4. Transfer the collected images to their storage location. Convert to tif (tagged image file) format. Name the files using the naming convention used by StarryNite.

A useful convention is to first create a directory with the same name as the series, then create subdirectories named "tif/" and "tifR/" containing the lineage channel and the reporter channel images, respectively (in monochrome 8-bit tif format). Each image should have a name such as "081505_t056_p14.tif" (this would be plane 14 of time point 56 of the image series named "081505").

5. Perform automated lineaging using StarryNite.

TABLE 1. Example of two-color lineaging settings for Zeiss LSM 510 and RW10029 (GFP) lineaging

Block	Time	488-nm laser top $\rightarrow$ bottom	488-nm laser settings top $\rightarrow$ bottom pinhole (AU) ^a	Line skip ^b
1	To end of four-cell stage	0.3% $\rightarrow$ 1.3%	3	6
2	70 min	0.3% $\rightarrow$ 1.3%	3	6
3	40 min	0.25% $\rightarrow$ 1.15%	2.5	5
4	30 min	0.25% $\rightarrow$ 1.0%	2	5
5	30 min	0.2% $\rightarrow$ 0.88%	1.8	4
6	180 min	0.2% $\rightarrow$ 0.75%	1.6	4
7	10 h, every 30 min	0.2% $\rightarrow$ 0.75%	1.6	4

The following 488-nm laser settings are constant with time: Gain (top $\rightarrow$ bottom) = 1000 $\rightarrow$ 1100; line averaging = 2; scan speed = 8; resolution 712(x) $\times$ 512(y) $\times$ 31(z); x-y pixels = 0.09 μm , z-planes separated by 1 μm .

543-nm laser settings for imaging mCherry reporters^c: Gain (top $\rightarrow$ bottom) 1100 $\rightarrow$ 1150; line averaging = 2; scan speed, line skip, pixel resolution, as for the GFP channel; pinhole = 2 AU. Laser power = 5% $\rightarrow$ 25% (this can be increased to 100% for very dim reporters).

^aAiry units.

^bLine skip means scan every *n*th horizontal line and interpolate linearly. Thus, with a *y* resolution of 512 pixels and a line skip of 4, only 128 lines are scanned. Even using a line skip of 6, *z* resolution is the limiting factor in resolving nuclei.

^cIn theory, a longer-wavelength laser such as 561 nm would more efficiently excite mCherry. However, in practice, phototoxicity at 543 nm is low enough that this is only minimally important (unless laser power is limiting).

TABLE 2. Example of two-color lineaging settings for Zeiss LSM 510 and RW10226 (mCherry) lineaging

Block	Time	543-nm laser top → bottom	543-nm laser settings top → bottom pinhole (AU) ^a	Line skip ^b
1	To end of four-cell stage	6% → 30%	3	6
2	70 min	6% → 30%	3	6
3	40 min	5% → 24%	2.5	5
4	30 min	4% → 20%	2	5
5	30 min	3% → 17%	2	4
6	30 min	2.5% → 14%	1.8	4
7	150 min	2% → 11%	1.6	4
7	10 h, every 30 min	2% → 11%	1.6	4

The following 543-nm laser settings are constant with time: Gain (top → bottom) = 1000 → 1100; line averaging = 2; scan speed = 8; resolution 712(x) × 512(y) × 31(z); x,y pixels = 0.09 μm, z-planes separated by 1 μm.

488-nm laser settings for imaging GFP reporters^c: Gain (top → bottom) 1100 → 1150; line averaging = 2; scan speed, line skip, pixel resolution, as for mCherry channel; pinhole = 2 AU; laser power = 0.2% → 1% (this can be increased to 100% for very dim reporters).

Some of the settings above will not transfer to CCD-based platforms such as spinning disk (e.g., line skip, variable pinhole). The increased signal-to-noise ratio of the detectors in these systems should offset this, allowing viable imaging. In general, StarryNite will accurately identify nuclei if they can be easily identified and can be distinguished by eye.

^aAiry units.

^bLine skip means scan every *n*th horizontal line and interpolate linearly. Thus, with a *y* resolution of 512 pixels and a line skip of 4, only 128 lines are scanned. Even using a line skip of 6, *z* resolution is the limiting factor in resolving nuclei.

^cIn theory, a longer wavelength laser such as 561 nm would more efficiently excite mCherry. However, in practice, phototoxicity at 543 nm is low enough that this is only minimally important (unless laser power is limiting).

- i. From the command line, access the directory containing the software. Enter “java -jar SNLauncher.jar.”

The best way to run StarryNite (automated lineage analysis) is to use the Java helper program SNLauncher.jar (Fig. 1).

- ii. On the interface, select a typical image, the location of the StarryNite binary, a StarryNite parameter file, and a series name or directory.

The image can be any of the appropriately named images from Step 4, and a sample parameters file is included with the software distribution. If you specify the full path of the series directory, SNLauncher will create another subdirectory called “data/” (parallel to the “tif” and “tifR” directories) that will contain the StarryNite and AceTree output files.

- iii. Using “Edit params,” adjust the time_start and time_end values appropriately. Click “Update,” and then, click “StarryNite.”

Typically, it is useful to first optimize the parameters on a small number of time points (10–20).

FIGURE 1. SNLauncher interface. To fill in each file, use the “Browse” buttons. Select “Edit params” to change parameters, “StarryNite” to run StarryNite, and “AceTree” to view the results in AceTree.

- iv. When StarryNite is finished (should be <1 min for 10 to 20 time points), use the “AceTree” button to open the series in AceTree. Navigate the images in AceTree (as described in Steps 6–9) to determine whether the annotations are generally correct.

If there are many annotations (i.e., circles in the AceTree image window) that are missing (false negatives) or extra annotations (false positives), you will need to optimize the StarryNite parameters.

See Troubleshooting.

- v. Close AceTree. Adjust the end time to include all of the time points. Click ‘StarryNite’ again. On a typical new desktop computer in 2009, StarryNite should take <30 min to trace nuclei through comma (>560-cell) stage for an image series with time points every minute.
- vi. When run through SNLauncher, the output of StarryNite will be two files in a directory named “dats”: “<seriesname>.xml” and “<seriesname>.zip.” The .xml file is an AceTree configuration file and specifies the location of images, the name of the zip file, the number of time points, etc. The .zip file contains the nuclei files produced by StarryNite (Bao et al. 2006) in a format readable by AceTree. These files can be moved to a different directory (such as a directory for each particular image series) as long as they are kept together.

Viewing the Lineage in AceTree

6. Run AceTree from the command line by entering “java -jar -mx500m AceTree.jar,” or by double-clicking on a Windows machine with Java installed.

This should bring up the basic AceTree menu (Fig. 2).

7. To open the embryo to be analyzed, run “Open Config File” from the File menu.

This should open the image window. A list of identified lineages (in a format called “JTree”) should appear in the AceTree console.

8. To navigate the lineage with the JTree, open and close different lineages by clicking on the switches to their left.

- i. Left-click on a cell name in the JTree to bring up the image corresponding to that cell’s first time point; right-click to bring up the last time point.
- ii. Use the left/right/up/down arrow keys to navigate the images forward and back in time and up and down in z.
- iii. Right click on a cell in the image window to select that cell, and turn on tracking (or click the “Track” button on the AceTree console).

Tracking causes the display to switch z-planes if the cell being tracked moves out of the plane in a different time point.

9. View the lineage for a given root.

- i. Select “Ancestral Tree” from the Trees menu.

This will bring up a colored tree (Fig. 2, bottom). The colors correspond to the expression levels of the cells if expression has been extracted (see Step 24).

- ii. In this display, left-click on a branch to select and to display the cell at the time in which it was clicked, or right-click to bring up the last time point for that cell.

Basic Curating/Editing of the Lineage in AceTree

The basic process for editing in AceTree is to identify possible errors based on the tree topology or positions of nuclei, to display the relevant cells (see Steps 8 and 9), and to use editing tools (see Steps 10–16) to make the appropriate changes. Lineaging errors include four fundamental types, which are summarized in the Discussion, along with the methods for fixing them. Note that even very complex errors are still just combinations of the four types and, hence, can be fixed by making one change at a time.

10. To begin editing, open the image series in AceTree. Select “Edit Tools” from the Edit menu. This will bring up two new windows (“Nuclei Relink” and “Add One”). It will also enable the basic editing features and keyboard shortcuts. The basic operations are listed here. In some

FIGURE 2. AceTree windows arranged for editing. The title bar of the image window (top left) shows that the image is at time point 223, plane 11 (t223-p11.tif). The circles show the StarryNite output (annotated nuclei). The currently selected cell is “Earaa” as evidenced by the white circle in the image, and the name “Earaa” highlighted in the JTree (upper right). Console options are HideA/ShowA (turn on/off cell names on image), HideC/ShowC (turn on/off cell circles on image), Up/Down (z control), Prev/Next (time control), Clear (remove accumulated names), Track (turn on/off cell tracking with time), Sister (show location of selected cell’s sister), Copy (disabled) and All/G/R/N (toggle between lineaging, reporter, combined, and no images). (Lower window) AceTree “Ancestral Tree” view. The root and the end time can be specified to focus on particular parts of the tree. If reporter signal has been extracted, “minRed” and “maxRed” specify the upper and lower limits of expression intensity levels to display. In this example, a red signal is present and is displayed with values ≤ -500 as gray, values between -500 and 5000 on a green $\rightarrow$ red gradient, and values >5000 as saturated red. The editing control windows “Add single nuclei” and “Nuclei Editor” (top center) provide editing functionality. Selecting “apply/rebuild” in “Nuclei Editor” would link the cell Earaa at time point 223 to the cell MSppapp at time point 224, resulting in an annotated division of Earaa and the annotated death of MSppapp at time point 223.

cases, there are buttons in the editing console that are redundant with the keyboard shortcuts. For these cases, only the keyboard shortcut is listed because the buttons are equivalent and self-explanatory.

11. Navigate the images forward/back in time with the left/right arrows or up/down in z with the up/down arrows.
12. Select the cell of interest.
 - i. Select a cell at a particular time point by clicking that cell in any tree or go to that cell’s last time point by right-clicking on it.
 - ii. Select an annotated cell in the image by right-clicking on its circle.

The circle will turn white, meaning it is selected and is tracking.

13. Edit the cells as necessary to correct lineaging errors:

- i. Remove the currently selected cell with the “Delete” key, or click “killCells.” The “kill-Cells” dialog allows one to delete several time points or all future time points for the selected cell.

This is useful for persistent false positives such as dust flecks outside of the embryo.

- ii. Move the currently selected cell in the *x* and/or *y* planes with Ctrl–(up/down/left/right).
- iii. Move the currently selected cell up or down in *z* with Shift–(up/down).
- iv. Make the currently selected cell larger or smaller with Shift–(left/right).

14. To create a new cell, middle-mouse-click on the image.

15. To link the currently selected cell to another cell at a different time point:

- i. Select the cell (right-click on the cell) at the beginning of the relink (typically a false negative or missed division error). Click “set early cell” in the “relink” dialog.

Alternatively, use keyboard shortcut F1.

- ii. Select the cell at the end of the relink (e.g., the first time point the cell is found again after being missed for several time points). Click set late cell (or use the F2 key) in the relink dialog.

Alternatively, you can add a cell (middle-click; see Step 14) at the appropriate position and time and can use “set late cell.”

- iii. Use “apply” (F3 key) or “apply/rebuild” (F4 key) to apply the change. “apply/rebuild” or “rebuild” (F5 key) will create a new tree with the accumulated changes for display in the various tree displays and lists.

Some good editing practices are to limit relinks to 5 min or less (for longer relinks, add intermediate cells) and to single time points for cells with large movement (such as dividing cells). When relinking, check that you are relinking to the first time point a cell reappears because relinking to a later point creates phantom cells in the embryo that can be hard to identify and can interfere with proper reporter quantification (the start and end times of a cell are displayed in the AceTree console window).

16. To save the edited results, select “Save Nuclei” from the File menu. Save as a .zip file.

Advanced Curating/Editing of the Lineage in AceTree

The latest version of the editing program AceTree has many new features designed to make editing more efficient that were not present in the previously released version (Boyle et al. 2006). It is important to be able to identify errors on the tree and to edit them, and these tools can allow you to do most editing from lists. Several advanced editing tools are available in the editing menu. For most of these tools, the result of the tool is a list of candidate errors. By clicking on each item in the list, the appropriate position in the images will be displayed, allowing you to decide whether the annotation is correct and how to fix it if necessary. These tools are described below in an appropriate order for their use.

17. Make sure that the initial lineage (to ~50 cells) is correct by using standard editing (see Steps 10–16). Identify a target editing end time.

For the 350-cell stage, we use the point when E has 16 descendants.

18. Remove deep false positives.

- i. Identify a *z*-plane that is below all real nuclei (e.g., by looking at a late time point).
- ii. Select “Kill Deep Nucs” from the Edit menu.
- iii. Specify the appropriate *z*-plane.

“Estimate” will tell you how many nuclei are below that plane.

- iv. Use “Kill’em” to delete all of those nuclei.

19. Use “Lazarus” from the Edit menu to fix short-duration false negatives.

Lazarus finds cell deaths (likely false negatives) for which a new cell appears (by movement or division) in nearly the same position after a small number of time points.

- i. Use “setParms” to specify the number of time points (<10 min is usually good) and the three distance parameters (measured in pixels, i.e., how close the new nuclei should be to the old position, and how far it must have moved to get there if it is a new daughter or not).

The defaults are likely good enough for most purposes.

- ii. Specify an end time.

- iii. Click “Deaths.”

This generates a list of “deaths” that will be fixed.

- iv. Click “linkEm.”

- v. Click “rebuild” to apply these changes.

- vi. Repeat Steps 19iii–v.

The changes made in the first iteration will create new deaths to fix. After iterating several times, no new “deaths” will be reported.

20. Use “Siamese” in the Edit menu to identify potentially incorrect division cells, and manually edit them.

Siamese uses three criteria to identify suspect divisions: A cell lifetime less than a specified cutoff, a center of gravity of the two daughter nuclei that is far from the parent position, and an asymmetric division in which one daughter moves far and the other daughter moves very little.

- i. Use “setParms” to edit distance and lifetime parameters.

- ii. Specify a start and end time.

- iii. Click “Divisions” to generate a list of candidate division errors.

- iv. Examine each candidate, and make any necessary changes.

21. Use “Deaths and Adjacencies” to check “Cell Deaths.”

Mostly these are false negatives before the 350-cell stage.

- i. Set an end time.

- ii. Click “Deaths” to generate a list.

- iii. Check, and fix each item.

- iv. Click “Jumps” in the same window to generate a list of nuclei that move an excessively large distance (measured in cell diameters).

22. Check the tree to make sure the division pattern matches that of Sulston (if it is a wild-type embryo) or is biologically plausible.

- i. Select “Ancestral Tree” from the Trees menu.

- ii. Select “Juvenescence” from the Edit menu to identify candidate nonbiological division events (e.g., cells that divide with a lifetime <0.8 times or >2.5 times that of their parent).

- iii. Select “Orientations” from the Edit menu to identify divisions in which the division angle (relative to the anteroposterior, dorsoventral, and left–right axes) differs by >60° from the average position of the wild-type embryos.

23. To save the edited results, select “Save Nuclei” from the File menu. Save as a .zip file.

Reporter Signal Extraction

24. Quantify reporter signals using AceBatch.

AceBatch requires the reporter images to be in a directory named “tifR/” that, parallel to the “tif/” directory, contains the lineaging images.

- i. To run AceBatch, enter “java -cp acebatch2.jar RedExtractor1<.xml filename and path>
<end time>.”

FIGURE 3. Extraction and display of reporter signal. (A) Background subtraction. Nuclei within the inner sphere are used to calculate raw signal, and the outer shell is used to calculate background (with cutoffs if using blot). (B) Example of a three-dimensional output produced by the “3D2 View” tool. In this case, expressing cells are colored from green to red based on their expression level, but they could instead be colored by lineage identity. (C) Example of a lineage-based expression display. Full lineage is shown in black, and expression is colored red (as produced by the “V Ancestral Tree” tool).

For example, entering “`java -cp acebatch2.jar RedExtractor1 /nfs/waterston/murray/ 102405pha4/ 102405pha4.xml 200`” would extract reporter signals for time points 1–200 for a series named “102405_pha4” stored in the directory “/nfs/waterston/murray/ 102405_pha4/.” The resulting data would be written over the existing .zip file for that series and would be available the next time the series is opened in AceTree.

- ii. In AceTree, open the resulting file by opening the original .xml file using “Open Config File” from the File menu.

The reporter signal for the selected nucleus will be displayed at the end of the line starting with the “current index.”

- iii. Using “Options” in the File menu, select one of the four types of background subtraction (Fig. 3A).

For brightly expressed nuclear-localized reporters, the choice of background subtraction method has very little impact on the observed pattern.

- *none*: average raw signal within each nucleus
- *global*: none minus a standard correction (25,000 intensity units)

For reporters with leaky nuclear localization or cytoplasmic localization, global can be a better option because it is not impacted by cytoplasmic signal.

- *local*: none minus local background (the average intensity of pixels in a shell between *kmedium* (1.2 in the example) and *klarge* (2.0) times the radius of the nucleus

local is more sensitive than blot but results in more false positives (i.e., visually nonexpressing nuclei with positive expression values).

- *blot*: same as *local*, except that pixels within other nearby nuclei are excluded from the background calculation. *kblot* gives a multiplier used when excluding neighboring nuclei (for *kblot*, no pixels within 1.2 radii of the neighboring nucleus are included).

The default corrects signals using blot, which gives the most reproducible measurements of replicate strains for tightly nuclear localized reporters.

Viewing

AceTree offers several options for viewing the expression data.

25. Select “Ancestral Tree” from the Trees menu to display expression data between the specified min/max intensity levels using a red → green color scale.
26. Select “V Ancestral Tree” from the Trees menu to bring up a menu allowing the creation of a tree using a black → arbitrary color scale.

This option allows one to specify root cell, end time, minimum and maximum intensities for the color scale, y spacing, line width, and hue (values range from 0 to 1, where 0 = red, 0.33 = green, and 0.66 = blue).

 - i. Select “show” to display the tree on the screen.
 - ii. Select “print” to export a Postscript file of the tree for use in figure making (Fig. 3C).
 - iii. Middle-click on a branch of the V ancestral tree on the screen to bring up a plot of expression versus time for that branch and its ancestors.
27. Select “3D2 View” in the View menu to view a three-dimensional projection of the nuclei at the current time point (Fig. 3B). Use the Properties tab to customize the display (e.g., showing cells by expression, how to color cells from different lineages, etc.).

Export

For further quantitative description of expression patterns, it is useful to export the data to a computer-readable format. Once the red signal is extracted, AceBatch command lines can be used for this. These generate comma-separated value (csv) files that can be opened in Microsoft Excel or can be used for further analysis. The csv files are placed in the same directory as the .xml file.

28. Enter the command “java -cp acebatch2.jar RedExcel2 <path/seriesname.xml> <endtime>” to generate files (“CAseriesname.csv” and “CDseriesname.csv”) in which each row is a particular cell and the columns contain different data values (e.g., position, reporter signal with different correction methods, etc.).

The CA and CD files differ in how they treat individual time points for the same cell. For the CA file, all values for the same cell are averaged, and there is one row per cell (~700 rows for a 350-cell series). For the CD file, each time point is presented separately, and there is one row per cell time point (10,000–20,000 rows for a 350-cell series). The CA file contains all of the data for a particular image series in an easily understood format.

29. Enter the command “java -cp acebatch2.jar RedExcel1 <path/seriesname.xml> <endtime>” to generate a file (named “Sseriesname.csv”) in which each line represents a terminal cell and each column is a time point. Each data point is the *blot*-corrected reporter signal for that cell or its ancestor at the corresponding time point.

As a result, early columns contain duplicated data because many terminal cells share the same early ancestor.

TROUBLESHOOTING

Problem (Step 3): Embryos do not survive imaging.

Solution: Consider the following:

1. Check that the mounting conditions are compatible with viability: Did nonimaged embryos on the same mount hatch? If not, there could be too many embryos, or there could be bacterial growth in the mounting medium. Either of these can cause anoxia, which will cause all embryos on the slide to arrest. Try mounting fewer embryos (five or less), and be careful to avoid transferring bacteria from the worm plate to the mount. It can also help to thaw a fresh tube of bead slurry and to add fresh antibiotics.
2. If the imaging process itself causes loss of viability, try reducing the excitation intensity and increasing the gain (on a photomultiplier tube detector system), or increasing the time

between images. If the images look pretty, it is unlikely the embryo will survive. The image quality needs only to be sufficient to identify nuclei and to quantify reporter signal.

Problem (Step 5.iv): There are too many false positives in the StarryNite output.

Solution: Try increasing the filtering stringency in the StarryNite parameters (contained in the Parameter file). A good strategy is to increase “nuc_density_cutoff” and “noise_factor” by increments of ~ 0.1 until “unsatisfactory” becomes “satisfactory.” We have recently discovered that much higher noise_factor values (in the range of 5–10) can be needed for images collected on other systems such as the Leica SP5. Boyle et al. (2006) and Murray et al. (2006) include detailed descriptions of all parameters in the StarryNite parameter files, but it is usually sufficient to change just the parameters listed here. Note that “nuc_density_cutoff” and “noise_factor” have three values that are used for <50 cells, 51 to 180 cells, and >180 cells, and that these can be tuned independently.

Problem (Step 5.iv): There are too many false negatives in StarryNite output.

Solution: Try reducing “max_weight_cutoff” to 0.1 or 0.05. If there are still too many false negatives, reduce “nuc_density_cutoff” and “noise_factor” in increments of 0.1.

Problem (Step 5.iv): Nuclei identified by StarryNite are the wrong size.

Solution: Check that the “nuc_size” matches the diameter of the initial nuclei (in pixels). Fragmented nuclei can also reflect oversegmentation; try reducing “noise_factor” or increasing “noise_fraction” (e.g., from 0.05 to 0.1).

Problem (Step 5.iv): Cells are not named properly.

Solution: AceTree requires that the x -axes of the images correspond to the anteroposterior axis of the embryo for proper naming. For microscopes for which the collection angle cannot be changed, a rotating stage insert is helpful. Alternatively, the images can be rotated after collection using an imaging program such as ImageJ.

AceTree attempts, by default, to identify the four-cell stage using the following rules: (1) Identify the four-cell stage (i.e., the time point when four cells are present); (2) identify the short-axis cells (EMS and ABp) and the long-axis cells (ABa and P2) by their positions: ABa and P2 are closest to the left/right sides of the images, and EMS and ABp are closest to the top and bottom of the images; (3) distinguish between these cells based on their division time: ABa divides before P2, and ABp divides before EMS. The program then assigns these names and an axis to the embryo. If this process fails, the cells will have numbered names such as Nuc1, Nuc2, etc. Sometimes the automatic naming process is thrown off by small errors that can be corrected easily (e.g., an early false positive or false negative makes the division pattern incorrect, or a false positive blocks correct identification of the four-cell stage). In other cases (e.g., a series starting at the six- or eight-cell stage), it might be necessary to name the founder cells and to set the axis manually. To do this, add a line to the .xml file specifying the axis:

```
<axis axis = “ADL”/>
```

The axis can be one of ADL, AVR, PDR, or PVL. The letters state, in order, the embryonic axes corresponding with left, up, and higher z -planes, so an ADL series has anterior (ABa) to the left, dorsal (ABp) at the top of the image, and left at high z -planes. Open this .xml file in AceTree, open the “Editing Tools,” select each cell, type its correct name in the Force Name box, and click “Force Name.” This can be performed at any stage, although the cells will only link to the root (P0), allowing full tree display if the names start at or before the eight-cell stage.

DISCUSSION

The nematode *Caenorhabditis elegans* develops through an invariant pattern of cell divisions and fate determination decisions (Sulston and Horvitz 1977; Sulston et al. 1983). The development of 4D

imaging methods and efficient software tools for manual lineage tracing from these images (Schnabel et al. 1997) made it possible to analyze the lineage for reporter genes or mutants. This work emphasized the power of using lineage alterations as a detailed phenotype for genetic analyses. However, it was still too inefficient for large-scale studies. To make lineage analysis more efficient, we developed a method for automated lineage tracing and expression mapping of *C. elegans* embryos expressing fluorescent protein–histone fusions (Bao et al. 2006; Murray et al. 2006, 2008). This method produces full lineages through the 350-cell stage with 1–2 h of manual lineage curation (compared with several days required for manual lineaging to this stage). It is also useful in phenotyping mutants. For example, we used these methods to analyze the lineage-specific control of cell cycle timing during embryogenesis (Bao et al. 2008).

Developing Imaging Parameters

Because of variations in optical paths, detectors, and laser intensities, it is necessary to develop detailed imaging parameters specifically for your microscope, but the ratios of exposure intensity from high to low plane and early to late time points should be consistent across platforms. As an example, the settings for a Zeiss LSM 510 are listed in Table 1 and Table 2 for the standard GFP and mCherry lineaging strains RW10029 and RW10226, respectively. To optimize the settings for a new microscope, the best strategy is to attempt to replicate the image quality in the images available at <http://waterston.gs.washington.edu/downloads/081505-images.zip>. This set includes sample images collected from 31 z-planes, 1 μm apart, taken once per minute through the 350-cell stage of development (195 min). If embryos imaged with these settings develop normally (as judged by timing of morphogenesis and hatching and larval morphology), it might be possible to improve image quality by increasing exposure times and reducing detector gain or by using a smaller pinhole.

Lineage tracing is possible with images that have quite a bit of noise. If you are used to collecting images suitable for cell biology journal figures, you might have difficulty forcing yourself to take suitably noisy images (e.g., Fig. 2). Increasing the interval between time points as high as 2 min results in adequate lineage tracing; the error rate for division matching increases, but this is offset partially by a reduced number of false negatives because of the smaller number of time points. The embryos are comparatively resistant to excitation light of 543 nm or longer (e.g., RFP), so it is most useful to tune the GFP settings for viability, regardless of whether this is the reporter or the lineaging channel.

Types of Lineaging Errors

- **False negative.** This appears as a cell death or a missed division in the lineage tree. In the image, this appears as a nucleus (i.e., region of GFP or RFP signal) with no circle around it. Identify the time point before the cell death (or missed division), and relink forward to the appropriate cell at a later time point (or newly create a cell to relink to). Note that usually this later cell (i.e., in which the false negative reappears) will be attached to the tree arbitrarily and, thus, will have an arbitrary (incorrect) name.
- **False positive.** This appears as an extra branch on the tree. In the images, this appears as a circle around an area that does not appear to be a nucleus (noise) or as a nucleus that contains two or more circles: Use `killCells` to remove the extra cell.
- **Mispositioned nucleus.** This appears as a nucleus with a circle that is in the wrong position (e.g., displaced to one side), or is larger or smaller than what appears to be the correct size: Move to an appropriate position, and resize if necessary.
- **Incorrect link.** These might not be visible on the tree but are usually coupled to other errors (false positives or negatives) that are. Incorrect links can be seen in the image window by tracking a nucleus over time; the tracked nucleus is indicated by a white circle. If the white circle jumps from one nucleus to another, this is a bad link. Relink the correct cells to each other.

Uses of the Method

The method described above will provide detailed expression information for any GFP or RFP reporter expressed by the 350-cell stage of embryogenesis. For reporters expressed in a modest number of cells after the 350-cell stage but before the onset of movement, it is possible to curate only the lineages leading to the expressing cells. One difficulty of this is in determining which sublineages to edit. A good general strategy is to visually trace backward from some of the expressing cells to an earlier stage in which their ancestors can be identified, then to edit these lineages to the onset of expression. This process can then be repeated for other expressing cells of interest. The same editing techniques described above can be used, although it is important to follow the path manually to each expressing cell from the last fully edited time point. This is necessary because partial editing removes one of the best checks of editing quality—the fact that silent errors, which do not affect the topology of the lineage tree, typically do result in errors in other parts of the lineage tree.

One use of this method is to identify expressing cells in a particular movie. This is more challenging than it sounds, because the expression values are quantitative intensity measurements and include a background noise level. With the settings described here, the noise range for nonexpressing cells ranges from -2000 to 0 units (blot corrected). Typically, a cell with average expression >0 is highly statistically significant. This can lead to high sensitivity: A brightly expressed reporter can reach levels $>100,000$ units in some cells but can also be expressed at levels <5000 in others. The significance of these differences depends on the particular situation. One important factor that compounds the noise in comparing cells within an embryo is the effect of z -position. The average effect of z -position on intensity is estimated to be $\sim 3\%$ per micrometer when using gradient imaging parameters or roughly twofold from the top of the embryo to the bottom. Thus, the observation that one group of cells is twice as bright as another might not be significant if the dimmer cells are on the bottom of the embryo.

Note that this method assumes the reporter signal is localized to the nucleus (either by nuclear localization signal fusion or by protein fusion to a nuclear-localized protein). If analyzing a cytoplasmic signal, the best approach is to trace the lineage and to manually check the identity of the expressing cells rather than relying on the quantification method described here. However, for nuclear reporters, this procedure provides the first full-animal method to quantitatively measure expression with high sensitivity and cellular resolution.

ACKNOWLEDGMENTS

We thank members of the Waterston and Bao laboratories for their help in thinking about and developing these methods, especially R. Waterston and T. Boyle for comments on the manuscript. J.I.M. is supported by funds from the National Institutes of Health (NIH) (GM083145), and Z.B. is supported by funds from the NIH (HG004643).

REFERENCES

- Bao Z, Murray JI. 2011. Mounting *Caenorhabditis elegans* embryos for live imaging of embryogenesis. *Cold Spring Harb Protoc* doi: 10.1101/pdb.prot065599.
- Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. 2006. Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc Natl Acad Sci* 103: 2707–2712.
- Bao Z, Zhao Z, Boyle TJ, Murray JI, Waterston RH. 2008. Control of cell cycle timing during *C. elegans* embryogenesis. *Dev Biol* 318: 65–72.
- Boyle TJ, Bao Z, Murray JI, Araya CL, Waterston RH. 2006. AceTree: A tool for visual analysis of *Caenorhabditis elegans* embryogenesis. *BMC Bioinformatics* 7: 275.
- Hope IA., ed. 1999. *C. elegans: A practical approach*. Oxford University Press, Oxford.
- Murray JI, Bao Z, Boyle TJ, Waterston RH. 2006. The lineaging of fluorescently-labeled *Caenorhabditis elegans* embryos with StarryNite and AceTree. *Nat Protoc* 1: 1468–1476.
- Murray JI, Bao Z, Boyle TJ, Boeck ME, Mericle BL, Nicholas TJ, Zhao Z, Sandel MJ, Waterston RH. 2008. Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat Methods* 5: 703–709.
- Schnabel R, Hutter H, Moerman D, Schnabel H. 1997. Assessing normal embryogenesis in *Caenorhabditis elegans* using a 4D microscope: Variability of development and regional specification. *Dev Biol* 184: 234–265.
- Stiernagle T. 2006. Maintenance of *C. elegans*. *WormBook*: 1–11.
- Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* 56: 110–156.
- Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* 100: 64–119.

Cold Spring Harbor Protocols

Automated Lineage and Expression Profiling in Live *Caenorhabditis elegans* Embryos

John Isaac Murray and Zhirong Bao

Cold Spring Harb Protoc; doi: 10.1101/pdb.prot070615

Email Alerting Service

Receive free email alerts when new articles cite this article - [click here](#).

Subject Categories

Browse articles on similar topics from *Cold Spring Harbor Protocols*.

- [C. elegans](#) (45 articles)
- [Cell Imaging](#) (525 articles)
- [Developmental Biology](#) (728 articles)
- [Fluorescent Proteins](#) (259 articles)
- [Image Analysis](#) (124 articles)
- [Imaging Development](#) (255 articles)
- [Use of Reporter Genes](#) (124 articles)
- [Video Imaging / Time Lapse Imaging](#) (171 articles)

To subscribe to *Cold Spring Harbor Protocols* go to:
<http://cshprotocols.cshlp.org/subscriptions>