Quantitative semi-automated analysis of morphogenesis with single-cell resolution in complex embryos

Claudiu A. Giurumescu[1], Sukryool Kang[2], Thomas A. Planchon[3], Eric Betzig[3], Joshua Bloomekatz[1], Deborah Yelon[1], Pamela Cosman[2] and Andrew D. Chisholm[1,*]

SUMMARY

A quantitative understanding of tissue morphogenesis requires description of the movements of individual cells in space and over time. In transparent embryos, such as *C. elegans*, fluorescently labeled nuclei can be imaged in three-dimensional time-lapse (4D) movies and automatically tracked through early cleavage divisions up to ~350 nuclei. A similar analysis of later stages of *C. elegans* development has been challenging owing to the increased error rates of automated tracking of large numbers of densely packed nuclei. We present Nucleitracker4D, a freely available software solution for tracking nuclei in complex embryos that integrates automated tracking of nuclei in local searches with manual curation. Using these methods, we have been able to track >99% of all nuclei generated in the *C. elegans* embryo. Our analysis reveals that ventral enclosure of the epidermis is accompanied by complex coordinated migration of the neuronal substrate. We can efficiently track large numbers of migrating nuclei in 4D movies of zebrafish cardiac morphogenesis, suggesting that this approach is generally useful in situations in which the number, packing or dynamics of nuclei present challenges for automated tracking.

KEY WORDS: Cell lineage, Nuclear tracking, Computer image analysis, Cell migration, *C. elegans*, Zebrafish

INTRODUCTION

Tracking individual cells in developing embryos or tissues is important for understanding the molecular basis of differentiation and tissue morphogenesis at single-cell resolution. Through tracking, one gains access to a wide array of information on cell behavior, including the dynamics of cell division timing, ancestry, migratory paths and gene expression profiles. The quantitative and potentially comprehensive nature of such data is of increasing importance in systems-level analyses of development (Megason and Fraser, 2007). Development of robust and accurate methods for tracking cells is becoming increasingly important for analysis of data sets in neuroscience (Higginbotham et al., 2011), stem cell research (Glauche et al., 2009) and cell biology (Arhel et al., 2006; Meijering et al., 2009). Automated tracking of cells or particles has become widespread in situations in which the signal-to-noise ratio (SNR), speed and sparseness of the objects permit reliable segmentation and tracking. However, many biological samples consist of large numbers of closely packed objects with low SNR, posing challenges for purely automated tracking.

The rapid development and near-transparent nature of *C. elegans* embryos and larvae has long allowed cell divisions to be followed manually using differential interference contrast (DIC) microscopy of live animals (Sulston and Horvitz, 1977; Sulston et al., 1983). Cell lineages can be traced using tools such as Simi Biocell (Schnabel et al., 1997). However, the low SNR of DIC images has precluded automated tracking in all but the earliest stages (Hamahashi et al., 2005). Localized fluorescent labels such as histone-GFP fusions, with their higher SNR, have recently made computer-based nuclear identification and tracking possible. Automated segmentation and tracking of histone-GFP-labeled nuclei in early *C. elegans* embryos has greatly facilitated analysis of early embryonic cell lineages (Bao et al., 2006; Santella et al., 2010), mutant phenotypes (Boeck et al., 2011), and gene expression patterns (Murray et al., 2008). However, tracking nuclei in later embryonic development, here defined as after the 350-cell stage, has been more challenging. The increasing density of nuclei and their semi-synchronous cell divisions lead to increasing error rates in automated analysis of embryos with >350 nuclei. Moreover, the SNR currently possible in conventional confocal 4D imaging is limited by the need to avoid phototoxicity in long time-lapse movies.

To track nuclear movements quantitatively beyond the 350-cell stage and through morphogenetic stages, we developed a new software approach. This approach has two key aspects. First, we seamlessly integrate automated search and manual curation, whereby the user confirms nuclear identifications at each time point prior to searching at the next time point. Thus, each automated search starts with a confirmed set of nuclei, minimizing propagation of false positives and other errors in the lineage tree. Second, as most nuclei do not move more than a predictable distance in the time intervals used in *C. elegans* 4D movies, we confine searches within a defined radius of each previously confirmed nucleus, greatly reducing tracking errors. Our approach is efficient and adaptable: we are able to track nuclei in 4D movies from several microscopy platforms, including a newly developed optical sectioning technology (Planchon et al., 2011), and from nuclear GFP movies of zebrafish development. Our semi-automated approach allows single-cell-level tracking of nuclei in complex samples in which the errors incurred in fully automated lineaging make automatic analysis challenging.

MATERIALS AND METHODS

4D movie acquisition with laser scanning confocal microscopy

In the *C. elegans* movies reported here we used the HIS-72::GFP (*zuIs178 V*) strain RW10029 (Ooi et al., 2006) to label nuclei. The RW10029 strain

[1]Division of Biological Sciences, Section of Cell and Developmental Biology, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. [2]Department of Electrical and Computer Engineering, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. [3]Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA.

*Author for correspondence (chisholm@ucsd.edu)

was fully viable and fertile over several months of imaging. To acquire 4D data sets using laser scanning confocal microscopy, we followed previously published approaches (Murray et al., 2006). We used Zeiss LSM510 or LSM700 instruments equipped with a 100× NA 1.46 oil immersion objective. During acquisition, we employ a large pinhole size to improve the fluorescence signal, and a pixel dwell time of 1.26 μs to reduce laser exposure. We acquire z-stacks of 64 × 35 × 30 μm³ with a voxel resolution of 0.125 × 0.125 × 0.85 μm³ every minute for 480 minutes of development (Fig. 1A), starting at the 4- to 6-cell-stage embryo. We selected these imaging parameters to ensure embryo viability while maintaining image quality suitable for automated analysis. Different sublineages show different HIS-72::GFP fluorescence intensity, with body muscle D, endodermal E, and germline P₄ sublineages having the lowest signal (not shown). Because the nuclear concentration and intensity of HIS-72::GFP increases through cleavage divisions (notably ~90 minutes after the 4-cell stage), we adjusted imaging parameters to maintain a constant SNR over the time course of acquisition (Fig. 1B). We acquired 4D data sets from 18 embryos imaged on a Zeiss LSM510 and two embryos imaged on a Zeiss LSM700. Phototoxicity due to fluorescence imaging was strongly temperature dependent: embryos above 24°C arrested during time-lapse confocal imaging, whereas embryos below 24°C were fully viable under identical conditions. We developed a Peltier-based cooling device to regulate specimen temperature, controlled by the confocal software.

Imaging using Bessel beam microscopy

We also acquired *C. elegans* 4D movies using Bessel beam plane illumination microscopy (Planchon et al., 2011). Plane illumination microscopy is an optical microscopy method that provides optical sectioning (permitting 3D imaging); its main advantages are its low phototoxicity and its fast imaging capability. Plane illumination microscopy using Bessel beam excitation also improves the axial resolution, allowing imaging of full development while maintaining sample viability. Embryos were attached to polylysine-coated coverslips and oriented vertically in an imaging chamber containing M9 medium at 20-22°C. Volumes of 66 × 48 × 50 μm³ with a voxel resolution of 0.133 × 0.133 × 0.5 μm³ were acquired every minute for 480 minutes using the Bessel beam in linear excitation mode and 0.8 NA excitation and detection objectives. Five HIS-72::GFP embryos were recorded on the Bessel beam microscope.

Zebrafish nuclear GFP 4D movies

Movies of zebrafish endocardial development were obtained using transgenic line *Tg(fli1a:negfp)* (Roman et al., 2002). All animal work followed approved protocols of the University of California San Diego Institutional Animal Care and Use Committee. We mounted transgenic embryos on coverslip bottom dishes (MatTek) in AquaPor low-melting temperature agarose (National Diagnostics) and cultured them in E3 media (Nüsslein-Volhard and Dahm, 2002). We performed time-lapse imaging at 28.5°C using a Leica SP5 confocal microscope with a 40× water immersion objective, capturing 40 optical slices at 1-minute intervals for 20 minutes. After imaging, we removed embryos from the imaging dishes and cultured them overnight to confirm wild-type heart morphology.

Tracking software and data analysis

Our cell tracking software, NucleiTracker4D, is implemented in MATLAB (MathWorks, Natick, MA, USA) and is freely available as an open source project at Sourceforge (https://sourceforge.net/projects/nucleitracker4d/). A user manual is provided on Sourceforge. We compute SNR in 4D data sets as follows. First, each stack is rendered isometric by replicating missing z-slices. Using the user-accepted nuclear positions, we treat all pixels inside the spherical mask defining nuclei as signal, and those outside as noise. We compute the mean signal over all the nuclei and mean noise over the remaining stack space. The SNR in dB is:

$$SNR(t) = 10\log_{10}\left(\frac{mean\ Signal(t)}{mean\ Noise(t)}\right)^2, \qquad (1)$$

where *t* is the stack time. Other data analysis and statistical methods were implemented in MATLAB.

RESULTS
A local search approach to tracking nuclei in 4D data sets

Our goal was to develop methods for tracking of all nuclei in the *C. elegans* embryo through morphogenetic stages, using 4D movies of embryos in which all nuclei have been labeled with histone::GFP (see Materials and methods) (Fig. 1C; supplementary material Movie 1). We developed a combination of automated tracking and manual curation to generate 4D models of the nuclei that can be easily analyzed computationally (Fig. 1C). Tracking of cell nuclei in a developing embryo involves several steps. First, nuclei must be recognized at a given time point, *t*. Second, nuclei identified at the next time point, *t*+1, must be linked to nuclei at time *t*. In the developing *C. elegans* embryo, nuclei at *t* have four possible fates at *t*+1: movement within the search radius; movement over a longer distance; division; or death. As cell deaths are easily recognized (see below), the crux of the problem is to distinguish whether a nucleus in a new position at time *t*+1 is a pre-existing nucleus that has moved beyond the threshold distance, or
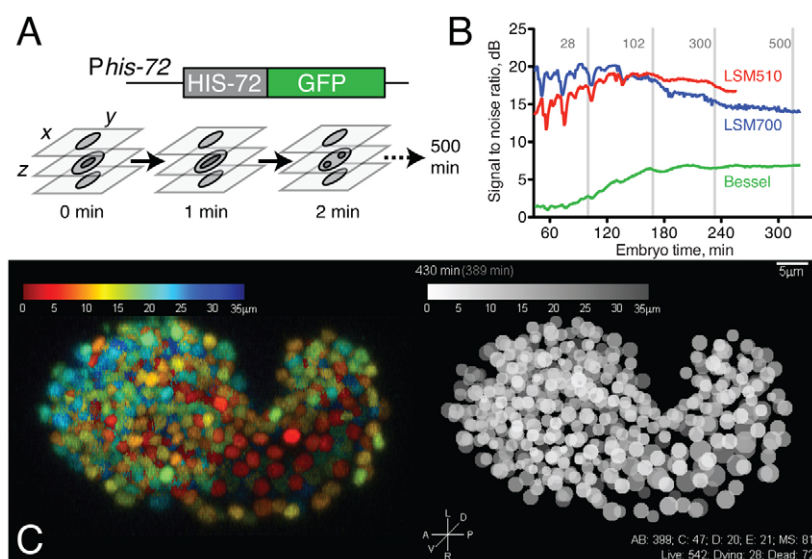

Fig. 1. 4D data acquisition and image properties. (**A**) Schematic of acquisition setup. We use laser scanning confocal microscopy to acquire z-stacks every minute for 5 hours of embryo development followed by z-stacks every two minutes for 3 hours at 20-24°C. Acquisition of each stack takes 12-15 seconds at 512 × 275 × 35 voxels. (**B**) Comparison of signal-to-noise ratio (SNR) of raw data from 4D movies taken using Zeiss LSM510, Zeiss LSM700 and Bessel beam microscopy. Image acquisition was optimized for image quality compatible with viability; under these parameters, the LSM510 and LSM700 data sets showed comparable SNR. (**C**) Comparison of confocal image (color depth code, projection of LSM700 z-stack, 430 minutes) with digital representation (gray scale depth code).

a daughter nucleus generated by division. Our approach does not involve de novo segmentation of the image into nuclei. Rather, we begin with the user defining the nuclei in the starting image, and confirming or correcting nuclear identifications at subsequent time points. Tracking the nuclei of a new sample embryo starts with the user naming the nuclei at the first time frame. We typically start at the 4- to 8-cell stage; at subsequent times, $t+1$, nuclei are propagated automatically based on their position at time t. To find nuclei at time $t+1$ we use only the local fluorescence information within a defined volume of a curated nucleus at time t (Fig. 2A).

Among the many possible segmentation methods for finding nuclear positions, such as active contours or gradient flow (Dufour et al., 2005; Li et al., 2007), we found that the maximum of the convolution of raw data with a spherical mask was most robust. This approach works best when the GFP signal of the nuclei is of uniform intensity and nuclear signals do not merge as a result of crowding. Because real data are noisy and nuclei become very packed at late stages, we decide nuclear position based on convolution of the raw signal with a spherical mask of radius $0.75R$, where R is the predefined radius of the nucleus (see below), capturing features of the nucleus that are less influenced by overall image noise. Given a curated nuclear position at time t, we define a cube with edges of length $1.25R$ around this position, convolve the time $t+1$ raw signal within this box with the spherical mask, and find the new position where the convolved signal is maximum (Fig. 2A). If more than one position within the search space shows a maximum convolution value, we choose the one closest to the position at time t. If the distance between the position at t and at $t+1$ is less than R, the new position is automatically accepted, and the user simply confirms that the automatic tracking is correct. Otherwise, the user is prompted for manual curation of the nucleus. Use of a search box of size proportional to nuclear radius was motivated by the observation that very early in development, nuclei are large and move around over long distances. At later stages, when nuclear radii are smaller, nuclei are more packed, hence their movement is constrained. In late development, large nuclei (e.g. E

or P_4 progeny) are also constrained; therefore, starting at ~250 minutes, we set the search space for nuclei in the C, E and P_4 sublineages as $\min[\text{radius}, \max(0.5\ \mu m, 1.25\delta_{t,t+1})]$, where $\delta_{t,t+1}$ is the displacement of that nucleus from the previous time point. These and other tracking parameters can be optimized depending on the particular sample. The tracking algorithm seeks a single nearby object at $t+1$ based on information at t. Thus, for cells that divided in the time between t and $t+1$, the tracking algorithm will only identify one daughter nucleus. The other daughter can be curated manually, or identified using a 'cell division detection' algorithm that has been described elsewhere (Kang et al., 2012). Most cell divisions in the data sets presented here were curated manually.

Effects of nuclear brightness, nuclear radius and cell death

Our convolution maximum approach is most successful when the nuclear fluorescence signal is uniform within a nucleus and neighboring nuclei have similar intensity. This approach is slightly less reliable when fluorescence intensity differs significantly between neighboring nuclei. For the HIS-72::GFP marker used here, nuclei in the D, P_4 and E lineages show fainter fluorescence than those of the AB, MS or C lineages. Thus, after convolution, the position of a faint nucleus is shifted towards one of the brighter neighboring nuclei if the latter is close enough (Fig. 2B). However, as these faint nuclear signals account for only 42 of all 558 live nuclei, and only some of these 42 reside near more intense AB or MS nuclei, mis-positioning accounts for <5% of curation events.

Although previous nuclear tracking methods search for an appropriate nuclear radius during segmentation (Bao et al., 2006), we found that the efficiency of our convolution maximum approach was increased by pre-specifying nuclear radii. In wild-type *C. elegans* embryos, nuclear radius decreases during development in a predictable way such that nuclei at any given round of division of a blastomere have similar radii (e.g. all AB[16] cells have similar radii) (Fig. 2C). The local search algorithm for tracking is
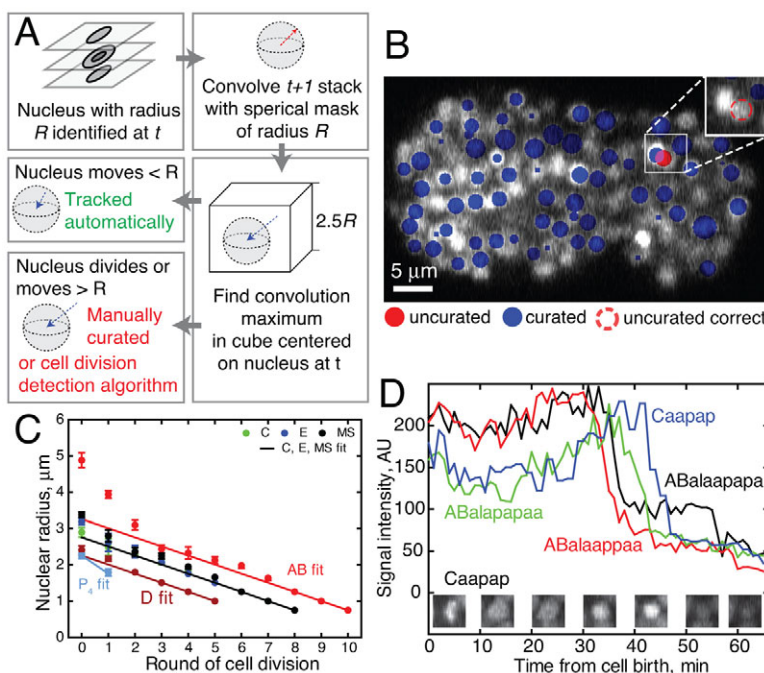

Fig. 2. Automated nuclear tracking using local search.
(**A**) Overview of the tracking algorithm in NucleiTracker4D. The tracking component searches the neighboring space of each nucleus identified at t for new convolution signal maxima. Only nuclei not satisfying the automatic tracking condition (e.g. displacement>R) are prompted for manual curation.
(**B**) An example of a tracking error requiring manual curation. Here, an AB nucleus was mis-tracked towards one of its MS neighbors owing to differing brightness among sublineages. The user corrected position is marked with a dashed red line.
(**C**) Variation in nuclear radii in the major embryonic sublineages. The linear fit is used as a model of nuclear radius in NucleiTracker4D. For early cells (AB, C, E, MS) $n\geq5$ (data from five embryos with a nucleus present in at least one time point); for all other cells, $16\leq n\leq32$, from two embryos with nuclei present for at least two time points. Where more than eight descendants (after third round of division) are present at one time point, cells quantified were picked at random. Error bars represent s.e.m. (**D**) Apoptotic cells undergo characteristic changes in HIS-72::GFP fluorescence; see supplementary material Movie 3. Nuclei of dying cells are assigned a radius according to the model until signal intensity decreases sharply corresponding to DNA condensation (~25 minutes after cell birth). At this point the dying cell nuclear radius is plotted as a minimal radius (0.5 μm). After another ~20 minutes, the HIS-72::GFP signal decreases as the cell is engulfed.

insensitive to slight decreases (but not to increases) in the nuclear radius, i.e. the predicted nucleus could have a radius slightly smaller by as much as 0.25 μm than the actual radius. The additional computation required to obtain a precise nuclear radius is therefore unnecessary for tracking. In mutants, or in non-*C. elegans* samples (see below), the nuclear radius model can be defined to allow more or less variation in nuclear size. Cells undergoing programmed cell death (apoptosis) display a distinctive pattern of histone-GFP expression in 4D movies (Fig. 2D; supplementary material Movie 2). Cells fated to undergo apoptosis are smaller than their sisters, although their nuclei are similar in size (Sulston et al., 1983; Hatzold and Conradt, 2008). Our procedure follows dying nuclei at their initial radius and represents them as a red circle of constant radius. When fluorescence of the dying cell drops below a threshold that would not interfere with the tracking of neighboring cells, the cell is marked as dead and is removed from the digital image.

An integrated graphical user interface for tracking, curation and visualization

Using the graphical user interface (GUI), users can specify the combination of algorithms to use for tracking and cell division and the cell naming convention to be used. If no tracking or cell division algorithms are specified, the program defaults to local search for tracking. The GUI relies on two sources of information for its operation at $t+1$: the user-curated list of nuclei at time t and the estimated list of nuclei at time $t+1$. The latter is derived from the former using the tracking and cell division algorithms specified by the user. The GUI prompts for user curation any objects that are not deemed correct by the algorithms; the user can also curate minor mistakes in automation, such as re-centering a virtual nucleus to a better center, so that the next step of tracking starts with a more reliable position. As our semi-automated approach is different in conception from fully automated tracking, a direct comparison of error rates is not meaningful. However, a comparison of the error rates of the automated part of our software with those of the initial and revised versions of Starrynite (Bao et al., 2006; Santella et al., 2010) indicates that it performs with comparable accuracy (supplementary material Fig. S1).

The key advantage to our approach is that error propagation is minimized by repeated curation as the number and packing of nuclei increase beyond 350 cells. At lower cell numbers or packing than 350 cells, our approach is slightly slower than the Starrynite algorithms, but yields a lineage that is completely accurate as judged by the user. A full 4D movie (300 *z*-stacks at 1 stack/minute, then 60 stacks taken every two minutes) has a cumulative total of 105,000 nuclear objects. About 22% of these accrue up to the 350-cell stage. Given the reliability of automated tracking in the early stages, one day of curation is sufficient to reach the 350-cell stage (270 minutes) (Sulston et al., 1983). As later stages require more user curation, it is possible to obtain an accurate lineaged embryo at a rate of 4000 nuclei tracking steps/day within three weeks. We implement several ways to visualize the resulting 4D data sets. Projections from selected views can be displayed with nuclei represented as depth-coded blobs (Fig. 3A-C; supplementary material Movies 1-3) or as flow-field trajectories (supplementary material Movies 5, 9). Trajectories can be quantified using standard metrics of displacement (Costa et al., 2005). Using these tools, the successive ventralward flows of ABpxp cells during gastrulation cleft closure (Fig. 3D), and of ABpxa cells during epidermal enclosure (Fig. 3E,F; supplementary material Movie 6) are dramatically visualized. High-

resolution versions of the supplementary movies, along with interactive Flash-based movies and the underlying data sets, are available at http://132.239.70.11/~wormlab/SupplementalMovies/. Supplementary material Fig. S1 shows 4D information for a single embryo (1330 nuclei tracked to 478 minutes; 99,000 nuclear positions).

Cellular neighborhoods and the effects of compression

Our analysis confirms that the *C. elegans* embryonic lineage is highly invariant in the wild type; we found only a single instance of an extra division in the E lineage among the 25 wild-type embryos analyzed (20 analyzed to 350-cell stage, five to comma stage). Timing of cell divisions was also highly reproducible, with cell cycles displaying standard deviations of ~4-7% of the mean (not shown), consistent with previous observations (Bao et al., 2008). Despite the invariance of *C. elegans* embryonic lineages and division timing, previous studies have reported greater variability in the relative positions of cells during embryogenesis (Schnabel et al., 1997; Schnabel et al., 2006). This variation might in part be due to differing degrees of mechanical compression of embryos during imaging (Hench et al., 2009). We therefore compared the development of embryos imaged under slight compression (bead mount, laser scanning microscopy confocal imaging) with that of embryos imaged in the uncompressed state using a novel structured illumination technology, Bessel beam imaging (Planchon et al., 2011). Bessel beam microscopy uses scanning of an annular beam projected to the rear pupil of the excitation objective, which itself is orthogonal to a second detection objective. In Bessel beam imaging, the specimen is mounted on a vertically oriented polylysine-coated slide within a buffer chamber, reducing or eliminating compression. Bessel beams permit illumination with extremely thin light sheets, resulting in increased *xyz* resolution with non-invasive levels of illumination. Bessel beam illumination allowed high resolution 4D movies of histone-GFP labeled *C. elegans* embryos to be acquired through embryonic development (supplementary material Movie 5).

Early in gastrulation, a subset of nuclei undergo coordinated rotation as a loose grouping, counterclockwise around the long axis as viewed from the posterior. We confirmed this overall rotation in compressed embryos (supplementary material Fig. S2A, Movie 4) and found that rotation was reduced but not eliminated in uncompressed (Bessel-imaged) embryos (supplementary material Fig. S2B). The gastrulation stage rotation process, therefore, is not purely a consequence of mechanical compression. To quantitatively analyze the effects of compression on cellular contacts, we defined nuclear nearest neighbors (NNs; supplementary material Fig. S3) using Delaunay triangulation. Delaunay triangulation entails selecting sets of four nuclear 'points' such that no other point is contained within the tetrahedron defined by these four points. Thus, each nuclear point will share the edges of the tetragons made up by itself and its nearest neighbors. An equivalent geometrical approach, Voronoi decomposition, has been applied to embryos of up to 150 cells (Hench et al., 2009), and our results are broadly consistent with that analysis (e.g. supplementary material Fig. S3B). Although this metric of 'instantaneous nearest neighbors' (INNs) might overestimate actual cell-cell contacts, it captures potential local interactions without imposing an arbitrary distance threshold. We define lifetime NNs (LNNs) as the cumulative number of INNs that a nucleus has over its lifetime; to normalize for the effects of differing cell lifetimes, we divide LNNs by the cell lifetime to yield the 'LNN rate'. LNNs of a particular cell are
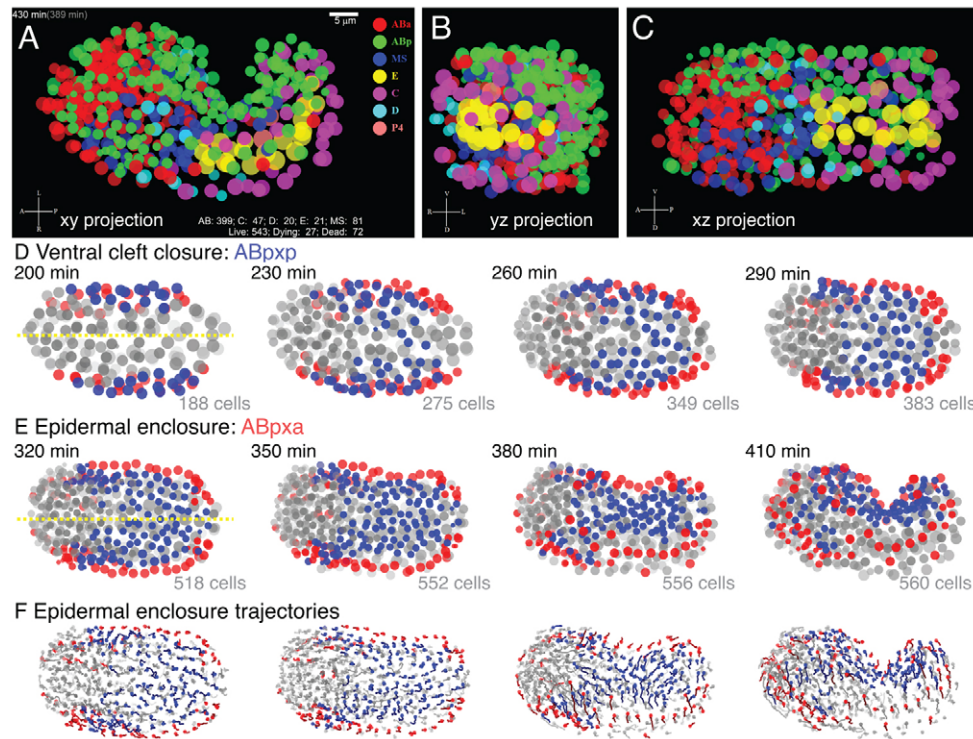
Fig. 3. Visualization of late stage *C. elegans* embryonic morphogenesis with single cell resolution. Visualization of morphogenetic movements from digital 4D movies, focusing on epidermal enclosure. Nuclei are depicted as semitransparent circles (A-E) or as trajectory flow fields (F). (**A-C**) Orthogonal views of 430-minute embryo; see also supplementary material Movies 1, 2. Time stamps at upper left denote actual movie time (gray, in parentheses) and estimated time since first cleavage (movie time + 91 minutes). The counters on lower right side show numbers of surviving descendants of each sublineage, total number of live nuclei, dying nuclei (cells undergoing apoptosis but still visible) and dead nuclei (no longer visible). All views are of the same data set from the ventral surface (*xy* view, A), posterior side (*yz*, B) and left lateral side (*xz*, C). Cells are color coded according to lineage origin. (**D,E**) ABp descendants undergo successive epiboly-like collective movements towards the ventral midline (yellow) during gastrulation cleft closure (200-290 minutes, D) and epidermal enclosure (320-410 minutes, E). ABpxp descendants (blue) are predominantly involved in ventral cleft closure; ABpxa descendants (red) are predominantly involved in epidermal enclosure. Images are every 30 minutes; cell counts include live and dying cells. (**F**) Flow field of nuclear trajectories during epidermal enclosure; ABpxx lineage coding as in D,E; 15-minute trajectories.

determined by its position in the embryo, its lifetime and whether it or its neighbors actively migrate. For example, cells with the most LNNs include internal endodermal cells, such as Eala, with 87.6 ± 6.3 LNNs over its lifetime of 83 minutes (Fig. 4A,B; supplementary material Movie 6), accumulated at a rate of 1.06 ± 0.09 NNs min^{-1}. By contrast, an exterior cell, such as ABprpppap, encounters 29.9 ± 3.7 NNs over its lifetime of 55 minutes, at a rate of 0.55 ± 0.07 NNs min^{-1}. Cells that migrate over long distances encounter more LNNs than do cells that remain within the same neighborhood. For example, the somatic gonadal precursors Z1 and Z4, which undergo long-range migrations from the head to the mid body, accumulate LNNs faster than cells in comparatively static neighborhoods, such as Eprpp (Fig. 4C,D). Finally, only cells that are NNs for a specified minimum time are counted, and are multiplied by the number of frames in which they are NNs, giving a measure of 'persistent NNs' (PNNs) per cell. Internal cells, such as the Eaxa cells, have many PNNs (17-18, as determined by confocal data sets). However, this is largely accounted for by their lifetime; when PNNs are normalized to lifetime, the PNN rate of Eaxa is low (~0.2 PNNs min^{-1}). PNNs were weakly but significantly anti-correlated with migration: cells that migrate effectively tend to have fewer PNNs (supplementary material Fig. S3A).

Next, we addressed whether compression affected cellular neighborhoods by comparing the INNs and LNNs of all cells that were completely imaged in confocal and Bessel embryos. The global mean of INNs increased through development, with an essentially identical profile in both imaging conditions (supplementary material Fig. S3B). LNNs for each nucleus were also highly correlated between compressed and uncompressed embryos (r=0.79) (Fig. 4B), as was the rate of accumulation of LNNs (not shown). To address whether the composition of these neighborhoods varied significantly between compressed and uncompressed data sets, we analyzed the common nearest neighbors (CNNs, common to confocal embryos, common to Bessel embryos, and common to both) (supplementary material Fig. S3C). CNNs of confocal-imaged embryos showed high pairwise correlation (Spearman r=0.87-0.89). Bessel-imaged embryos were slightly less correlated with each other (r=0.79-0.84) and significantly less correlated with confocal-imaged embryos (r=0.67-0.73) (supplementary material Fig. S3D). Compressed embryos displayed higher reproducibility in neighborhoods at most time points (supplementary material Fig. S3E). In conclusion, uncompressed embryos show slightly more embryo-to-embryo variation in their cellular neighborhoods than do compressed embryos, suggesting that compression constrains embryonic nuclei to a more predictable geometry.
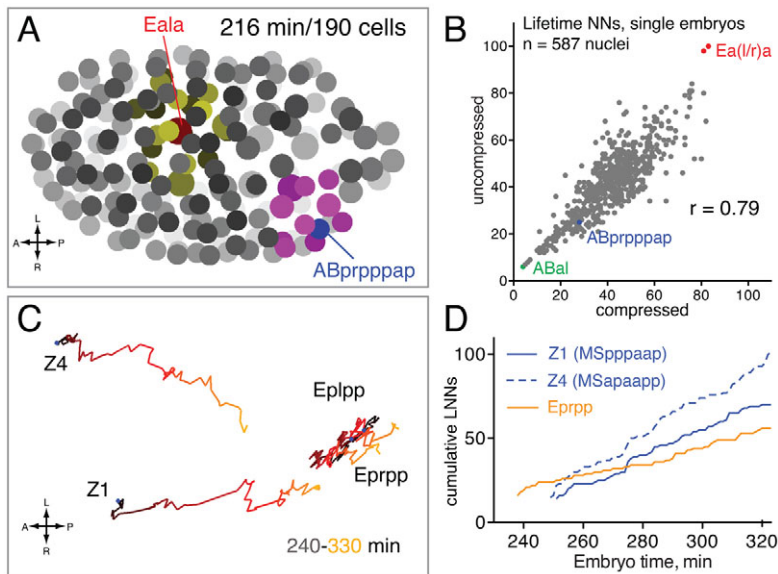
Fig. 4. Cellular neighborhoods and the effect of embryonic compression. (A) Embryo at 192-cell stage with Eala (red) and ABprpppap (blue) nuclei highlighted. Eala is in the middle of the embryo and has 18 instantaneous nearest neighbors (INNs, yellow) as defined by Delaunay triangulation. ABprpppap is located at the surface of the embryo and has ten INNs (magenta); see supplementary material Movie 6. **(B)** Correlation of lifetime nearest neighbors (LNNs) for 587 embryonic cells from third to ninth round of division, between single uncompressed (Bessel imaging) and compressed (LSM imaging) embryos. Cells highlighted include those with the highest LNN (Eala, Eara, red), lowest LNN (ABal, ABar, green) and the surface nucleus ABprpppap (blue). **(C)** Trajectories of long-range migrating cells (Z1, Z4) compared with non-migrating endodermal cells; data from Bessel imaging. **(D)** Migrating cells accumulate LNNs (blue) faster than do non-migrating cells (yellow).

Coordinated ventral neuroblast movements during ventral cleft closure and epidermal enclosure

Our primary interest in developing our nuclear tracking tool was to examine morphogenetic movements in the late embryo, a period in which the number and density of nuclei precluded the use of fully automated tracking algorithms. Here, we present the first single-cell resolution description of the movements of neural precursors in the mid-body following gastrulation and up to epidermal enclosure. After gastrulation, the coordinated ingression of the ventral mesoderm creates a cleft on the ventral surface that closes from posterior to anterior (Sulston et al., 1983). Cleft closure involves the convergent movement of multiple ventral neuroblasts (VNBs) leading to the formation of a VNB monolayer between 290 and 350 minutes (ABpxp-derived cells; Fig. 3D,E, blue; supplementary material Movie 7); this sheet subsequently provides the substrate for epidermal enclosure movements (ABpxa-derived; Fig. 3D,E, red). The ventral blast cells mostly generate neurons, but also make some epidermal and muscle cells; for simplicity, we refer to them as ventral neuroblasts.

VNB movements in cleft closure

Immediately flanking the posterior ventral cleft are three pairs of AB[128] cells: ABp(l/r)pappa, ABp(l/r)pappp and ABp(l/r)ppppp. After ~200 minutes, these precursors move to converge with their left-right counterparts, beginning with the posterior pair ABp(l/r)ppppp. Between 225 and 232 minutes, these six cells divide in the eighth round of AB divisions to generate the cleft-closing VNBs (Fig. 5A, red cells; supplementary material Movie 8); for brevity, we refer to these VNBs as A1-6L and A1-6R; see supplementary material Table S1 for a full list of VNB abbreviations. We found that other VNBs could be grouped into anteroposterior sets that underwent coordinated movements during cleft closure and epidermal enclosure. More lateral to the A VNBs are rows of VNBs that we term the B (yellow), C (blue), and D (dark blue) neuroblasts; these do not undergo extensive movements until epidermal enclosure. The different VNB groups might have distinct molecular identities; for example, CeNeuroD/CND-1 is expressed in the C, D and E VNB groups but not in the majority of A and B VNBs (Murray et al., 2012).

Eph signaling is required for the coordinated movement of A VNBs

Several pathways have been implicated in closure of the gastrulation cleft, including signaling via the Eph receptor VAB-1 (George et al., 1998). To define more precisely the role of VAB-1 in VNB movements, we tracked VNB movements in *vab-1* null mutant embryos and compared them with those in the wild type. We find that in *vab-1* mutants the parents of the A VNBs display dramatic defects in closure. These defects resulted from delays and stalling at specific points rather than an overall slower migration rate (Fig. 5B,C). Thus, VAB-1 signaling might be involved in interactions between the VNBs and specific cells in their environment, rather than in general adhesion or motility.

Substrate movements during epidermal enclosure

Following cleft closure, the VNBs remain fairly static until epidermal enclosure begins at ~350 minutes. During epidermal enclosure, a striking set of coordinated movements rearrange the progeny of the VNBs into a longitudinal row of ventral midline ganglia and neurons (supplementary material Movie 9). The B, C and D rows of VNBs converge on the ventral midline, concurrent with the displacement of A VNB progeny to the anterior and posterior. The anterior A VNBs give rise to the excretory cell and neurons in the retrovesicular ganglion; the posterior A VNBs generate the neuron PVT and a mixture of rectal, valve and muscle cells. The E VNBs (Fig. 5, cyan) undergo strikingly rapid long-range movements during early epidermal enclosure (Fig. 6A). Ventral cord motor neurons, such as the DDs, arise from C-rank VNBs, and intercalate at the ventral midline such that left-hand (odd numbered) neurons migrate anteriorly to right-hand (even) neurons, and migrate in concert with closure of the ventral pocket by epidermal P cells (Fig. 6B). Finally, a large number of anterior (head) neurons also undergo intricate migrations concomitant with late epidermal enclosure; for example, a group of approximately ten anterior ventral neurons and support cells form bilateral streams that converge towards the anterior ventral midline from 330 to 380 minutes (Fig. 6C). In conclusion, our analysis reveals the highly dynamic nature of the neuronal substrate during epidermal enclosure, suggesting close coordination of substrate and epidermal movement.
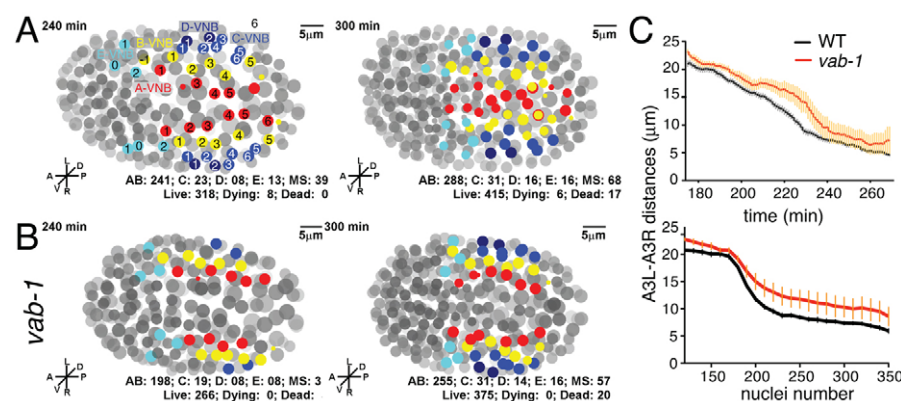
Fig. 5. Ventral neuroblast (VNB) migrations during ventral cleft closure and the role of VAB-1/EphR signaling. (**A**) Positions of VNBs during ventral cleft closure at 240 and 300 minutes. VNBs are named as described in the text and numbered from anterior to posterior. For a list of VNB identities, see supplementary material Table S1. (**B**) VNB positions in *vab-1(e2027)* at 240 and 300 minutes. (**C**) Quantification of midline migration (internuclear distance) for the A3 VNBs in wild-type and *vab-1* embryos (*n*=6 per genotype), plotted versus time or nuclear number. Error bars represent s.e.m.

Tracking patterns of cell movement during zebrafish endocardial morphogenesis

To address the generality of our local search and tracking algorithm, we tested it on 4D movies of GFP-labeled nuclei in the developing zebrafish heart. Transgenic expression of nuclear localized *EGFP* driven by the *fli1a* promoter allows labeling of nuclei during endocardial tube formation (Roman et al., 2002). Endocardial morphogenesis in zebrafish begins with the migration of endocardial progenitors from bilateral positions in the anterior lateral plate mesoderm to the midline at ~14 hours post-fertilization (hpf) (Bussmann et al., 2007). Once at the midline, endocardial cells translocate in a posterior direction between 17 and 20 hpf. Endocardial cells then undergo a leftward migration at ~20 hpf to form a ventrally extending tube by 22 hpf. The patterns of cell movement underlying these morphogenetic changes remain poorly understood.

We imaged endocardial movement beginning at 17 hpf, immediately after the fusion of the bilateral populations at the ventral midline and at the beginning of its posterior movement (Fig. 7A; supplementary material Movie 10). Despite the ellipsoidal morphology of the endocardial nuclei, we found that our local search algorithm performed well with few modifications, using a spherical convolution mask and a search box based on a constant nuclear radius. Nuclei move extensively during this stage (mean speed 1.7 µm/second, 1-minute intervals), up to 50 µm during the 20-minute movie. Nuclei were

also tightly packed; as no endocardial nuclei divided in this short movie, most curation events involved resolving potential collisions of nuclei. Using our tracking approach, we were able to visualize the migration of >130 endocardial nuclei over a 20-minute period (Fig. 7A).

Our analysis revealed the utility of NucleiTracker4D software for visualization of a diversity of patterns of individual endocardial nuclei movement. For example, we easily observed the rapid posterior migration of most endocardial nuclei at the midline (Fig. 7B), consistent with prior studies (Bussmann et al., 2007). At the same time, we were able to observe nuclei moving in the anterior direction (Fig. 7B), as well as nuclei with net leftward or rightward trajectories (Fig. 7C). Notably, quantification of overall displacement vectors in the depicted example revealed a net leftward flow of cells around a central core of cells that moved slightly rightward (Fig. 7C). Coupling individual cell tracks with Delaunay triangulation allowed us to evaluate the degree of correlation between the movement patterns of neighboring endocardial nuclei (Fig. 7D,E). For instance, within the example shown, neighboring nuclei showed significant correlations in their trajectories and speeds (Fig. 7D). However, some endocardial nuclei moved in the opposite direction from their nearest neighbors, suggesting that they actively migrate rather than being passively swept along by other cells (Fig. 7E). Thus, this approach provides opportunities for future studies that might lead to a deeper understanding of
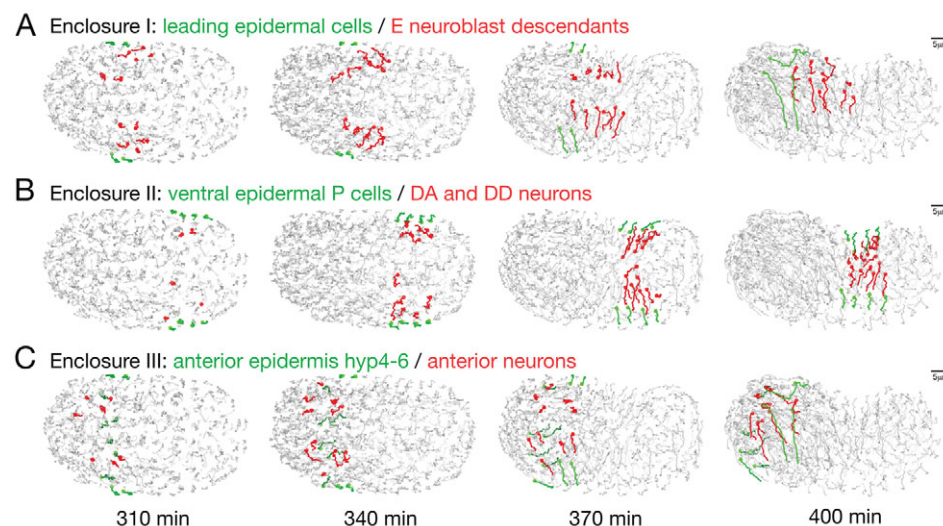

Fig. 6. Collective migration of substrate and epidermis during epidermal enclosure. (**A**) Correlated migrations of E VNB descendants (red) and leading epidermal cells [ABp(l/r)aappa(p/a), green] before and during early ventral enclosure. Trajectories are plotted as 30-minute tails, with other cells in grayscale depth code; circles at end of trajectories. (**B**) Correlated trajectories of the ventral epidermis (P cells 1-8, green) and underlying neurons (DA, DD neurons, red) during ventral pocket closure. (**C**) Migration of anterior epidermis (hyp6,5,4, green) during enclosure of the anterior head correlates with anterior flow of a large set of neurons from 300 to 400 minutes; a subset of anteriorly migrating neurons and support cells are highlighted in red (RMDD, RIP, URAV, ILsoV, ILshV). Scale bars: 5 µm.
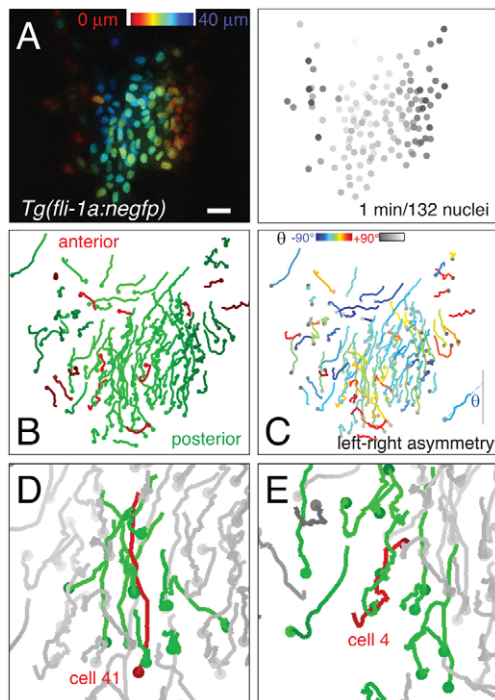
Fig. 7. Asymmetric migration patterns of zebrafish endocardial nuclei. (**A**) Left: Endocardial progenitors labeled with *Tg(fli1a:negfp)*. Dorsal view, posterior is down. Scale bar: 10 μm. Right: Frame from a single confocal movie of 1-minute intervals, together with the corresponding frame of the 132 individual nuclei tracked using NucleiTracker4D. Grayscale depth code reveals dorsal angioblasts (dark gray) at lateral positions. (**B**) Overall posterior movement of endocardial nuclei; trajectories are color coded according to their net movement along the anterior-posterior axis. (**C**) Left-right asymmetry of endocardial migration; trajectories are color coded by net angle θ relative to the anterior-posterior axis. In this movie, 88 cells moved leftward and 44 cells rightward. (**D**) Example of correlated trajectories of a target cell (cell 41) and its nearest neighbors. (**E**) Example of target cell (cell 4) moving in the opposite direction to its neighbors.

migration patterns underlying morphogenetic movements during endocardial tube formation.

DISCUSSION

We have presented a straightforward solution to the problem of tracking large numbers of nuclei in a complex and dynamic embryo. By integrating automated nuclear tracking with user curation we achieve both speed and accuracy. To our knowledge, the resulting data sets are the first complete lineages of single *C. elegans* embryos through morphogenetic stages. The accuracy of the original manual lineage analyses is a testament both to the power of direct observation and to the extreme invariance of the *C. elegans* embryonic lineage. However, our approach does not simply trace genealogical relations but allows quantitative analysis of many other previously inaccessible morphogenetic parameters.

Additional refinement of our approach is clearly possible. For example, as closely packed nuclei tend to move less, the size of the search space for nuclei could be adapted to suit local nuclear packing. Likewise, the use of a spherical mask may not always be optimal; an ellipsoidal mask could be used in cases in which nuclei are more variable in proportion. The open source nature of our software should allow modular additions and improvements by the community. However, the major limitation in achieving accurate automatic tracking from long 4D data sets is the need to compromise image quality (both in *xyzt* resolution and in SNR) to preserve specimen viability. Recent developments in 4D microscopy might be able to overcome this limit (Fischer et al., 2011; Wu et al., 2011). We find that the low light levels possible using Bessel beam illumination (Planchon et al., 2011) allow far higher resolution 4D movies to be obtained, without detectable phototoxicity. Unlike most other microscopy platforms, in which the embryo is slightly compressed between a cover slip and a slide, in the Bessel microscope the embryo is attached only on one side to a polylysine cover slip, and is otherwise suspended, uncompressed, in the buffer chamber. The higher resolution of nuclear structures possible with Bessel microscopy means that nuclei no longer appear as simple spherical blobs, but adopt a wider range of shapes. Although in theory such variability in nuclear form could make use of a spherical convolution mask less valid, in practice the improved resolution means that nuclear tracking accuracy is comparable to that in our LSM confocal data sets. The lack of compression possible in Bessel beam imaging might be important in certain experiments. However, uncompressed embryos can assume a wide variety of orientations, making visual inspection of embryonic processes occasionally challenging. We find that the slight compression of bead mounting allows embryos to develop in stereotyped ways that largely preserve the neighborhood relationships of nuclei in uncompressed embryos. For many purposes, these technical advantages of compression may be very useful.

The dynamic rearrangements of cells in the *C. elegans* embryo remain relatively poorly understood, with the exception of the ingression movements of gastrulation (Lee and Goldstein, 2003; Lee et al., 2006; Harrell and Goldstein, 2011). Following gastrulation, the most prominent morphogenetic movements are those of ventral cleft closure and subsequent epidermal enclosure. Our analysis has defined the individual cell movements that contribute to gastrulation cleft closure. The subsequent rearrangement of substrate cells during epidermal enclosure has not previously been described. Our analysis indicates that epidermal movements are precisely and extensively correlated with movements in the underlying neuronal substrate. Initial movement of the epidermal leading cells correlates with the rapid migrations of the E rank neuroblasts and their progeny. Closure of the ventral pocket is accompanied by extensive convergence and intercalation of the B, C and D rank neuroblasts, which generate the ventral nerve cord and ventral ganglia. Indeed, descendants of the A3, B3 and C3 VNBs have recently been shown to play specific roles in bridging the ventral pocket P9/10 cells during enclosure (Ikegami et al., 2012). Finally, we find that enclosure of the anterior epidermis is accompanied by extensive anterior flow of neurons and neuronal support cells. Interestingly, substrate movements play an active force-generating role in epithelial spreading during *Drosophila* dorsal closure (Solon et al., 2009). It will be important to address the functional and mechanical relationship of substrate and epidermal movements: are they independent, dependent or interdependent?

Our nuclear tracking algorithm has yielded, to our knowledge, the first complete descriptions of cellular positions during *C. elegans* epidermal enclosure. Although tracking nuclei does not give a complete picture of cell contacts or movements, it is important for interpreting other kinds of dynamic data, such as those involving cell surfaces or junctions. Despite recent advances,

automated segmentation of surface or junction markers remains challenging (Blanchard et al., 2009; Aigouy et al., 2010; Fernandez et al., 2010; Luengo-Oroz et al., 2011); nuclear information can provide a framework for such efforts. By combining the nuclear tracking information presented here with data on cell contacts, it might eventually be possible to generate predictive models of the cellular dynamics and mechanical forces at play during *C. elegans* embryonic morphogenesis.

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.086256/-/DC1

References
Aigouy, B., Farhadifar, R., Staple, D. B., Sagner, A., Röper, J. C., Jülicher, F. and Eaton, S. (2010). Cell flow reorients the axis of planar polarity in the wing epithelium of Drosophila. *Cell* **142**, 773-786.

Arhel, N., Genovesio, A., Kim, K. A., Miko, S., Perret, E., Olivo-Marin, J. C., Shorte, S. and Charneau, P. (2006). Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nat. Methods* **3**, 817-824.

Bao, Z., Murray, J. I., Boyle, T., Ooi, S. L., Sandel, M. J. and Waterston, R. H. (2006). Automated cell lineage tracing in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA* **103**, 2707-2712.

Bao, Z., Zhao, Z., Boyle, T. J., Murray, J. I. and Waterston, R. H. (2008). Control of cell cycle timing during C. elegans embryogenesis. *Dev. Biol.* **318**, 65-72.

Blanchard, G. B., Kabla, A. J., Schultz, N. L., Butler, L. C., Sanson, B., Gorfinkiel, N., Mahadevan, L. and Adams, R. J. (2009). Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation. *Nat. Methods* **6**, 458-464.

Boeck, M. E., Boyle, T., Bao, Z., Murray, J., Mericle, B. and Waterston, R. (2011). Specific roles for the GATA transcription factors end-1 and end-3 during C. elegans E-lineage development. *Dev. Biol.* **358**, 345-355.

Bussmann, J., Bakkers, J. and Schulte-Merker, S. (2007). Early endocardial morphogenesis requires Scl/Tal1. *PLoS Genet.* **3**, e140.

Costa, L. D. F., Cintra, L. C. and Schubert, D. (2005). An integrated approach to the characterization of cell movement. *Cytometry* **68A**, 92-100.

Dufour, A., Shinin, V., Tajbakhsh, S., Guillén-Aghion, N., Olivo-Marin, J. C. and Zimmer, C. (2005). Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *IEEE Trans. Image Process.* **14**, 1396-1410.

Fernandez, R., Das, P., Mirabet, V., Moscardi, E., Traas, J., Verdeil, J. L., Malandain, G. and Godin, C. (2010). Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. *Nat. Methods* **7**, 547-553.

Fischer, R. S., Wu, Y., Kanchanawong, P., Shroff, H. and Waterman, C. M. (2011). Microscopy in 3D: a biologist's toolbox. *Trends Cell Biol.* **21**, 682-691.

George, S. E., Simokat, K., Hardin, J. and Chisholm, A. D. (1998). The VAB-1 Eph receptor tyrosine kinase functions in neural and epithelial morphogenesis in C. elegans. *Cell* **92**, 633-643.

Glauche, I., Lorenz, R., Hasenclever, D. and Roeder, I. (2009). A novel view on stem cell development: analysing the shape of cellular genealogies. *Cell Prolif.* **42**, 248-263.

Hamahashi, S., Onami, S. and Kitano, H. (2005). Detection of nuclei in 4D Nomarski DIC microscope images of early Caenorhabditis elegans embryos using local image entropy and object tracking. *BMC Bioinformatics* **6**, 125.

Harrell, J. R. and Goldstein, B. (2011). Internalization of multiple cells during C. elegans gastrulation depends on common cytoskeletal mechanisms but different cell polarity and cell fate regulators. *Dev. Biol.* **350**, 1-12.

Hatzold, J. and Conradt, B. (2008). Control of apoptosis by asymmetric cell division. *PLoS Biol.* **6**, e84.

Hench, J., Henriksson, J., Lüppert, M. and Bürglin, T. R. (2009). Spatio-temporal reference model of Caenorhabditis elegans embryogenesis with cell contact maps. *Dev. Biol.* **333**, 1-13.

Higginbotham, H., Yokota, Y. and Anton, E. S. (2011). Strategies for analyzing neuronal progenitor development and neuronal migration in the developing cerebral cortex. *Cereb. Cortex* **21**, 1465-1474.

Ikegami, R., Simokat, K., Zheng, H., Brown, L., Garriga, G., Hardin, J. and Culotti, J. (2012). Semaphorin and Eph receptor signaling guide a series of cell movements for ventral enclosure in C. elegans. *Curr. Biol.* **22**, 1-11.

Kang, S., Giurumescu, C. A., Chisholm, A. D. and Cosman, P. (2012). Automated nuclei tracking in C. elegans based on spherical model fitting with multiple target tracking. In *Image Analysis and Interpretation (SSIAI)*, pp. 17-20. Santa Fe, NM: IEEE.

Lee, J. Y. and Goldstein, B. (2003). Mechanisms of cell positioning during C. elegans gastrulation. *Development* **130**, 307-320.

Lee, J. Y., Marston, D. J., Walston, T., Hardin, J., Halberstadt, A. and Goldstein, B. (2006). Wnt/frizzled signaling controls C. elegans gastrulation by activating actomyosin contractility. *Curr. Biol.* **16**, 1986-1997.

Li, G., Liu, T., Tarokh, A., Nie, J., Guo, L., Mara, A., Holley, S. and Wong, S. T. (2007). 3D cell nuclei segmentation based on gradient flow tracking. *BMC Cell Biol.* **8**, 40.

Luengo-Oroz, M. A., Ledesma-Carbayo, M. J., Peyriéras, N. and Santos, A. (2011). Image analysis for understanding embryo development: a bridge from microscopy to biological insights. *Curr. Opin. Genet. Dev.* **21**, 630-637.

Megason, S. G. and Fraser, S. E. (2007). Imaging in systems biology. *Cell* **130**, 784-795.

Meijering, E., Dzyubachyk, O., Smal, I. and van Cappellen, W. A. (2009). Tracking in cell and developmental biology. *Semin. Cell Dev. Biol.* **20**, 894-902.

Murray, J. I., Bao, Z., Boyle, T. J. and Waterston, R. H. (2006). The lineaging of fluorescently-labeled Caenorhabditis elegans embryos with StarryNite and AceTree. *Nat. Protoc.* **1**, 1468-1476.

Murray, J. I., Bao, Z., Boyle, T. J., Boeck, M. E., Mericle, B. L., Nicholas, T. J., Zhao, Z., Sandel, M. J. and Waterston, R. H. (2008). Automated analysis of embryonic gene expression with cellular resolution in C. elegans. *Nat. Methods* **5**, 703-709.

Murray, J. I., Boyle, T. J., Preston, E., Vafeados, D., Mericle, B., Weisdepp, P., Zhao, Z., Bao, Z., Boeck, M. and Waterston, R. H. (2012). Multidimensional regulation of gene expression in the C. elegans embryo. *Genome Res.* **22**, 1282-1294.

Nüsslein-Volhard, C. and Dahm, R. (2002). *Zebrafish: A Practical Approach*. New York: Oxford University Press.

Ooi, S. L., Priess, J. R. and Henikoff, S. (2006). Histone H3.3 variant dynamics in the germline of Caenorhabditis elegans. *PLoS Genet.* **2**, e97.

Planchon, T. A., Gao, L., Milkie, D. E., Davidson, M. W., Galbraith, J. A., Galbraith, C. G. and Betzig, E. (2011). Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. *Nat. Methods* **8**, 417-423.

Roman, B. L., Pham, V. N., Lawson, N. D., Kulik, M., Childs, S., Lekven, A. C., Garrity, D. M., Moon, R. T., Fishman, M. C., Lechleider, R. J. et al. (2002). Disruption of acvrl1 increases endothelial cell number in zebrafish cranial vessels. *Development* **129**, 3009-3019.

Santella, A., Du, Z., Nowotschin, S., Hadjantonakis, A. K. and Bao, Z. (2010). A hybrid blob-slice model for accurate and efficient detection of fluorescence labeled nuclei in 3D. *BMC Bioinformatics* **11**, 580.

Schnabel, R., Hutter, H., Moerman, D. and Schnabel, H. (1997). Assessing normal embryogenesis in Caenorhabditis elegans using a 4D microscope: variability of development and regional specification. *Dev. Biol.* **184**, 234-265.

Schnabel, R., Bischoff, M., Hintze, A., Schulz, A. K., Hejnol, A., Meinhardt, H. and Hutter, H. (2006). Global cell sorting in the C. elegans embryo defines a new mechanism for pattern formation. *Dev. Biol.* **294**, 418-431.

Solon, J., Kaya-Copur, A., Colombelli, J. and Brunner, D. (2009). Pulsed forces timed by a ratchet-like mechanism drive directed tissue movement during dorsal closure. *Cell* **137**, 1331-1342.

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev. Biol.* **56**, 110-156.

Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* **100**, 64-119.

Wu, Y., Ghitani, A., Christensen, R., Santella, A., Du, Z., Rondeau, G., Bao, Z., Colón-Ramos, D. and Shroff, H. (2011). Inverted selective plane illumination microscopy (iSPIM) enables coupled cell identity lineaging and neurodevelopmental imaging in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA* **108**, 17708-17713.
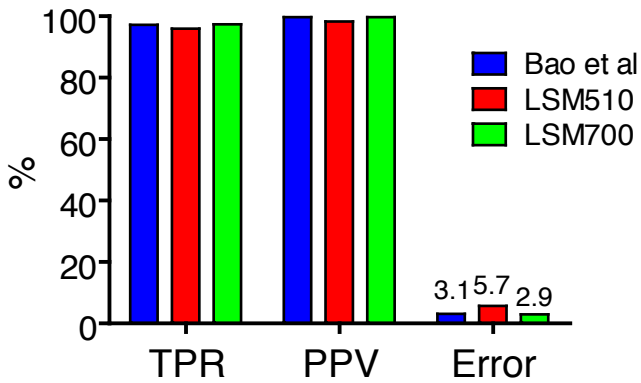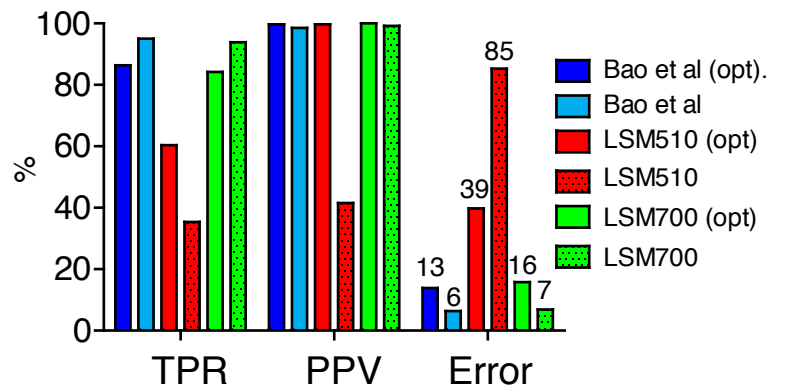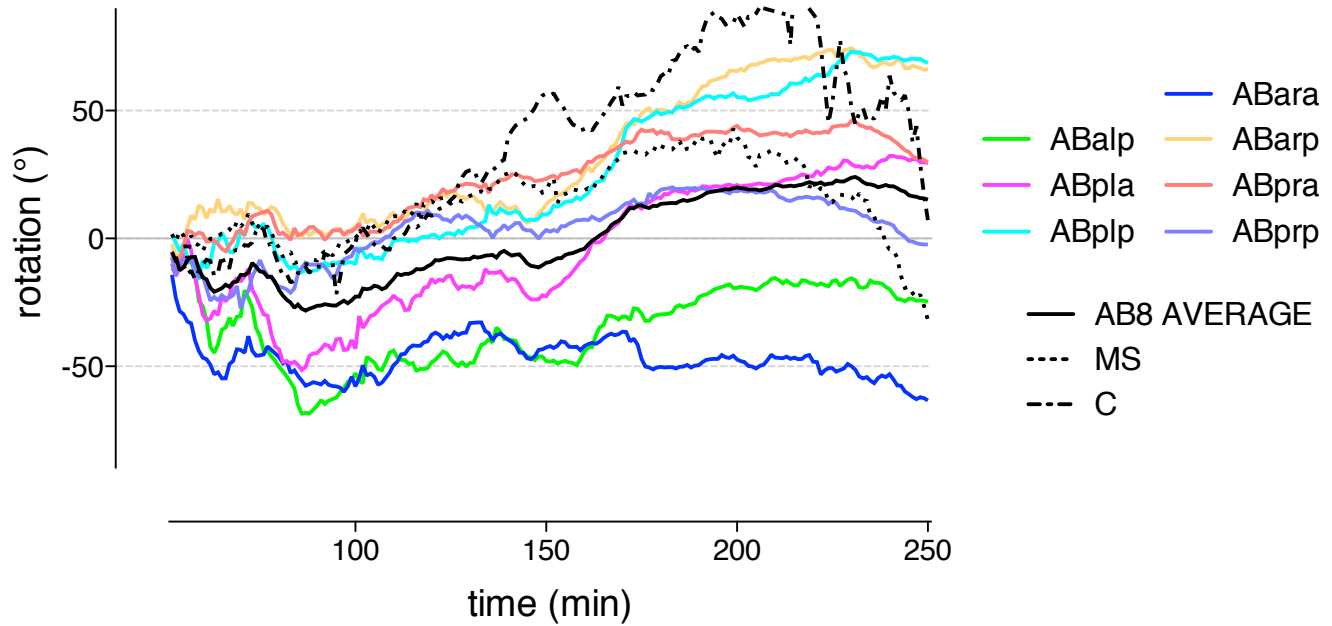
Fig. S1. Error rates in NucleiTracker 4D and other automated lineage software. (**A**) The confusion matrix and error metrics. (**B**) NucleiTracker4D performance. We used NucleiTracker4D to track nuclei over 99 to 115 minutes from the four-cell stage in three confocal data sets: the original data set from Bao et al. (Bao et al., 2006) (blue); one Zeiss LSM 510 movie (red); one dataset obtained from Zeiss LSM 700 (green) using NucleiTracker4D. NucleiTracker4D displays both a high true positive rate (TPR) and positive predicted value (PPV) across all data sets, resulting in very low cumulative error. (**C**) Starrynite v2 (Santella et al., 2010) performance. We first optimized the program parameters for each data set using curated data for the first 50 minutes (opt) and then ran the program with these optimized parameters up to 99 minutes.
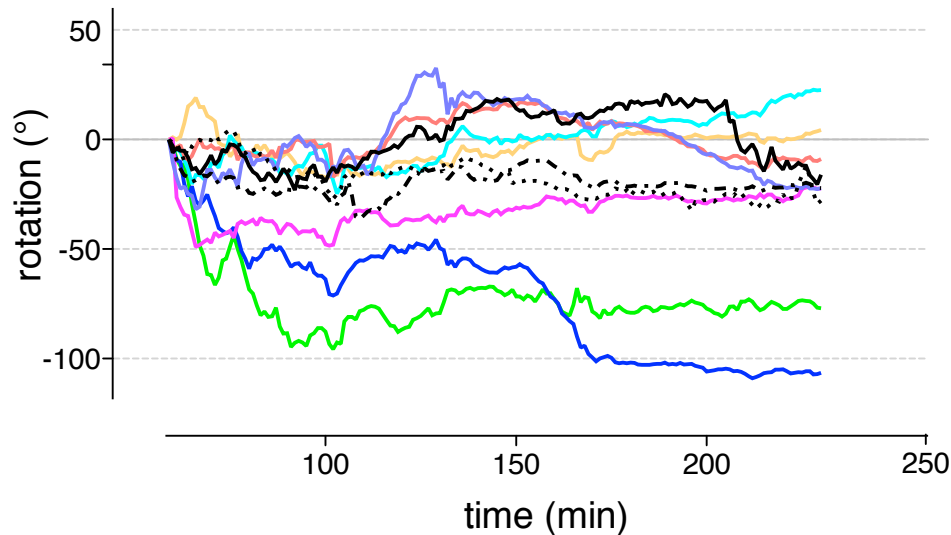
Fig. S2. The effect of compression on embryonic rotational movement and nuclear displacement. (**A**) The relative rotation of a given precursor and its descendants, expressed as degrees around the long axis of the embryo as viewed from the anterior pole. Angle of the precursor at 52 minutes is set as 0°. The angle of the center of gravity of the descendants is plotted over time, as is the average of the AB⁸ groups (excluding ABala, which lies at the anterior pole and movements of which are not well captured by this metric). Note the clockwise rotation of most cell groups except ABalp and ABara between 150 and 180 minutes. In general, posterior AB cells rotate more than anterior AB cells. AB⁸, MS and C color codes as in Movie 4. (**B**) Analysis of embryonic rotations in uncompressed embryos. Unlike confocal imaging, Bessel imaging does not constrain the embryo to a stereotypical orientation. We therefore rotated and aligned Bessel data sets to match the long axes and left-right symmetry of confocal imaged data sets, prior to analysis of embryonic rotation.
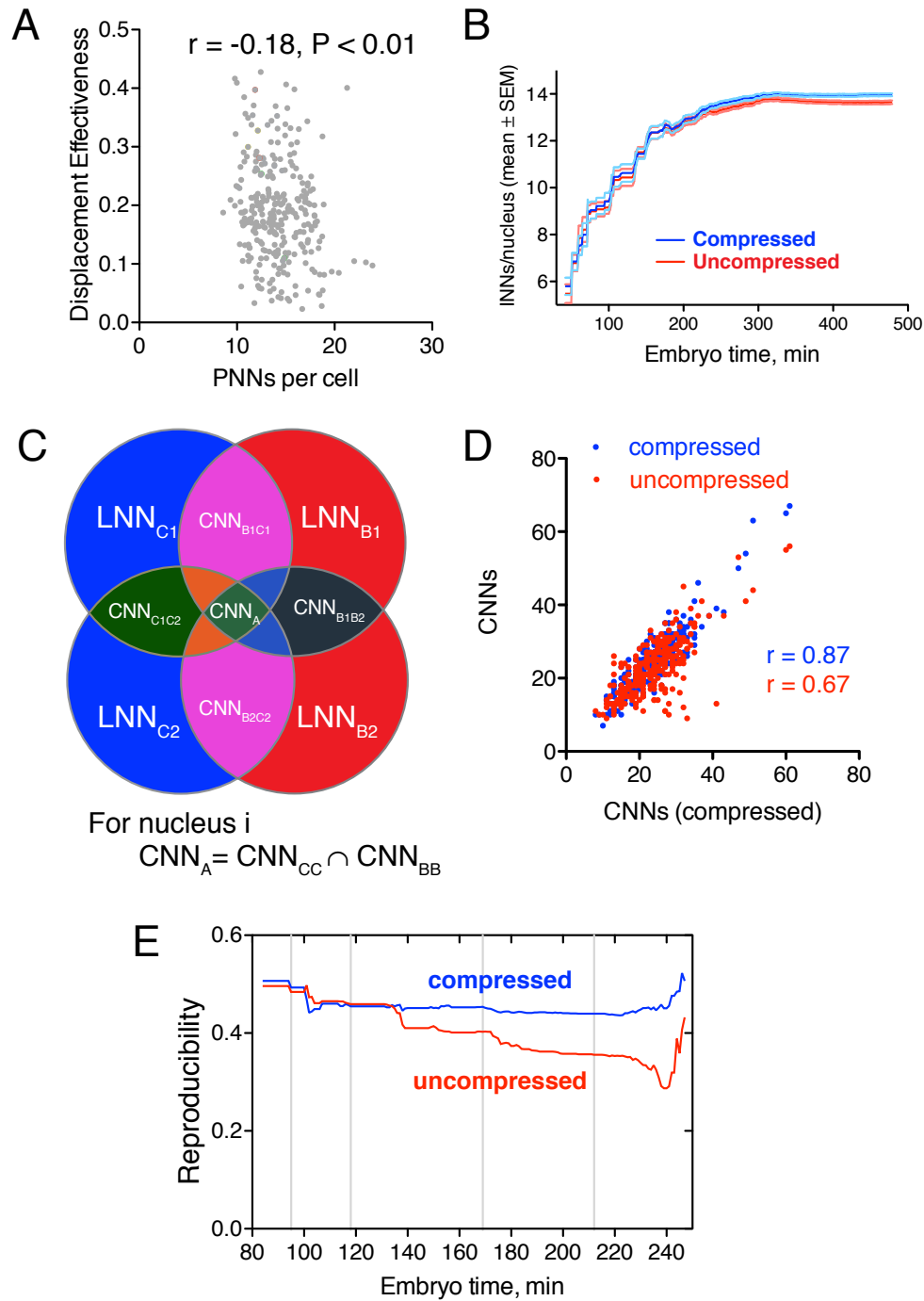
For nucleus i

$$CNN_A = CNN_{CC} \cap CNN_{BB}$$

Fig S3. Nuclear nearest neighbors (NNs) and the effect of compression. (**A**) Negative correlation of persistent NNs with displacement effectiveness in a single confocal-imaged embryo. The net displacement of a nucleus throughout its lifetime is the vector from its position at birth to its position prior to division. The total displacement is the sum of stepwise migrations made by a nucleus throughout its lifetime. The displacement effectiveness (DE) is defined as the ratio of net to total displacement. (**B**) Evolution of the global mean number of instantaneous nearest neighbors (INNs) through embryonic development for wild-type compressed (blue) and wild-type uncompressed (Bessel imaging, red) embryos. At each time point the number of nearest neighbors of all nuclei is averaged for the entire embryo. Error bars indicate s.e.m. across the nuclei present in the embryo at a particular time. Compressed embryos display slightly higher mean INNs at late time points; c.f. Figure 5C of Hench et al. (Hench et al., 2009). (**C**) Set description of common nearest neighbors for a specific nucleus. The set of common nearest neighbors (CNNs) is defined as the LNNs for a specific nucleus consistently found in all data sets. A nucleus must be neighboring for at least a threshold minimum time to be counted as a NN. For each nucleus we make all possible pairwise comparisons among embryos imaged in confocal (C1, C2...*n*=3) or Bessel (B1, B2...*n*=3) conditions. (**D**) Typical correlations of CNN values (threshold=2 minutes) between individual compressed (blue) and uncompressed (red) embryos. Compressed embryos are more highly correlated with each other than are uncompressed embryos. (**E**) Uncompressed (Bessel) imaged embryos display lower embryo-embryo reproducibility in nearest neighborhoods compared with confocal imaged embryos. Reproducibility is defined as the ratio (CNN/LNN), averaged over the whole embryo at each time point.

Table S1. Shorthand names for ventral neuroblasts from ~220 to 280 minutes

E, D right	C right	B right	A right	A left	B left	C left	E, D left
E0R ABarappapa RMDDR/x,SMBDR		**B-1R** ABprpapaaa SMDDR/AIYR			**B-1L** ABplpapaaa SMDDL/AIYL		**E0L** ABalpapapa RMDDL/x,SMBDL
E1R ABprpaaaap exc socket G1/DB3		**B0R** ABprpapaaap SIADR/SIBVR			**B0L** ABplpapaaap SIADL/SIBVL		E1L ABplpaaaap exc duct/DB1
E2R ABarappapp DB2/SMBVR	**C1R** ABprppaapa x/RIMR	**B1R** ABprpapapa excglR/AVKR	**A1R** ABprpappaa RIH/AVL	**A1L** ABplpappaa RMEV/exc	**B1L** ABplpapapa AVKL/excglL	**C1L** ABplppaapa x/RIML	**E2L** ABalpapapp X/SMBVL
	C2R ABprppaapp AIAR/DB7	**B2R** ABprpapapp SIAVR/DA8	**A2R** ABprpappap RIS/DB4	**A2L** ABplpappap x	**B2L** ABplpapapp SIAVL/DB5	**C2L** ABplppaapp AIAL/DB6	
D1R ABprppapaa SABVR, RIFR/DA1	**C3R** ABprppappa RIGR/DD2	**B3R** ABprppppaa PVPR/repVR	**A3R** ABprpapppa RIR/AVG	**A3L** ABplpapppa K/K'	**B3L** ABplppppaa PVPL/repVL	**C3L** ABplppappa RIGL/DD1	**D1L** ABplppapaa SABVL,RIFL/SABD
D2R ABprppapap DA3/DA5	**C4R** ABprpppaaa Y/DA7	**B4R** ABprppppap B/DVA	**A4R** ABprpapppp virl/virR	**A4L** ABplpapppp PVT/repD	**B4L** ABplppppap U/F	**C4L** ABplpppaaa DA9/DA6	**D2L** ABplppapap DA2/DA4
	C5R ABprppappp DD4/DD6	**B5R** ABprpppapa PHshR/hyp	**A5R** ABprpppppa bm/sph	**A5L** ABplpppppa mu intL/an dep	**B5L** ABplpppapa PHshL/hyp	**C5L** ABplppappp DD3/DD5	
	C6R ABprpppaap PVCR,LUAR/ PHAR	**(B6R)** ABprpppapp x	**A6R** ABprpppppp spike/hyp10	**A6L** ABplpppppp spike/hyp10	**(B6L)** ABplpppapp x	**C6L** ABplpppaap PVCL/LUAL/ PHAL	