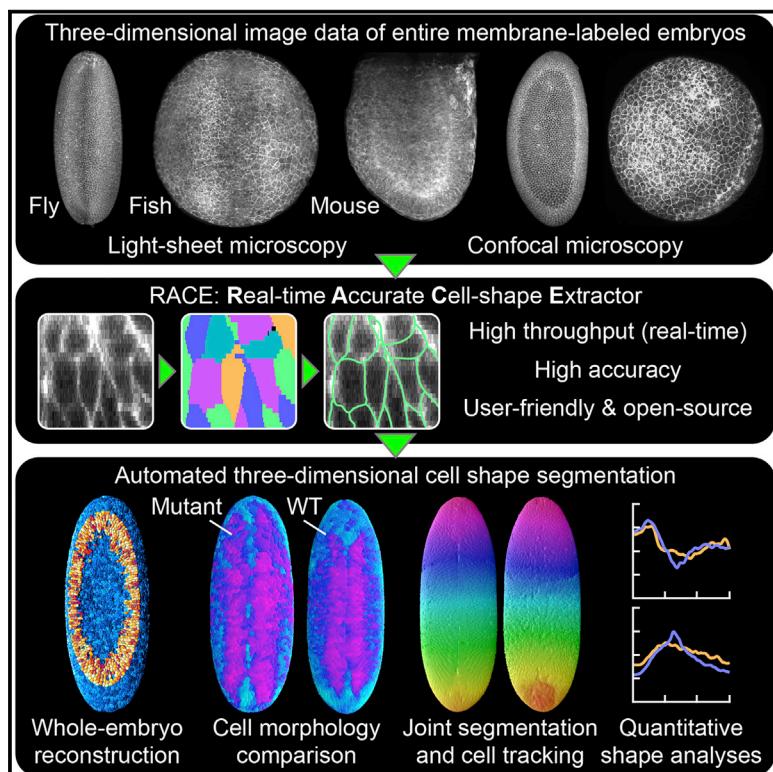


Developmental Cell

Real-Time Three-Dimensional Cell Segmentation in Large-Scale Microscopy Data of Developing Embryos

Graphical Abstract



Authors

Johannes Stegmaier, Fernando Amat, William C. Lemon, ..., George Teodoro, Ralf Mikut, Philipp J. Keller

Correspondence

johannes.stegmaier@kit.edu (J.S.), kellerp@janelia.hhmi.org (P.J.K.)

In Brief

Stegmaier et al. present the open-source software RACE for automated three-dimensional cell segmentation in large-scale microscopy images. RACE rapidly reconstructs cell shapes in entire developing *Drosophila*, zebrafish, and mouse embryos imaged with confocal or light-sheet microscopy and quantifies cell-shape dynamics and phenotypic differences in wild-type and mutant embryos.

Highlights

- RACE framework automates 3D cell segmentation in entire fly, fish, and mouse embryos
- RACE is 55–330 times faster and 2–5 times more accurate than state-of-the-art methods
- RACE quantifies differences in cellular dynamics in wild-type and mutant embryos
- RACE + TGMM enable joint segmentation and cell tracking in entire developing embryos

Real-Time Three-Dimensional Cell Segmentation in Large-Scale Microscopy Data of Developing Embryos

Johannes Stegmaier,^{1,2,*} Fernando Amat,¹ William C. Lemon,¹ Katie McDole,¹ Yinan Wan,¹ George Teodoro,³ Ralf Mikut,² and Philipp J. Keller^{1,*}

¹Howard Hughes Medical Institute, Janelia Research Campus, 19700 Helix Drive, Ashburn, VA 20147, USA

²Karlsruhe Institute of Technology, Institute for Applied Computer Science, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

³Department of Computer Science, University of Brasilia, Campus Darcy Ribeiro, Brasília, CEP 70910-900, Brazil

*Correspondence: johannes.stegmaier@kit.edu (J.S.), kellerp@janelia.hhmi.org (P.J.K.)

<http://dx.doi.org/10.1016/j.devcel.2015.12.028>

SUMMARY

We present the Real-time Accurate Cell-shape Extractor (RACE), a high-throughput image analysis framework for automated three-dimensional cell segmentation in large-scale images. RACE is 55–330 times faster and 2–5 times more accurate than state-of-the-art methods. We demonstrate the generality of RACE by extracting cell-shape information from entire *Drosophila*, zebrafish, and mouse embryos imaged with confocal and light-sheet microscopes. Using RACE, we automatically reconstructed cellular-resolution tissue anisotropy maps across developing *Drosophila* embryos and quantified differences in cell-shape dynamics in wild-type and mutant embryos. We furthermore integrated RACE with our framework for automated cell lineage tracing and performed joint segmentation and cell tracking in entire *Drosophila* embryos. RACE processed these terabyte-sized datasets on a single computer within 1.4 days. RACE is easy to use, as it requires adjustment of only three parameters, takes full advantage of state-of-the-art multi-core processors and graphics cards, and is available as open-source software for Windows, Linux, and Mac OS.

INTRODUCTION

Fluorescence microscopy is an essential tool for live imaging in developmental biology. Recent technological advances have made it possible to follow the development of tissues, organs, and entire embryos at the single-cell level and with high temporal resolution for up to several days. The large amounts of multi-dimensional image data produced by such experiments demand automated computational methods to accurately and rapidly convert the raw images into biologically interpretable information (Keller, 2013; Pantazis and Supatto, 2014). Quantitative analyses of the shapes of cells captured in these recordings are crucial for

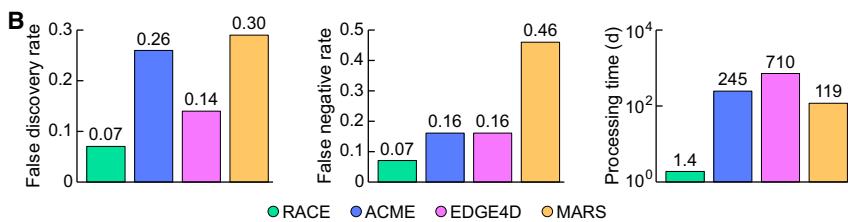
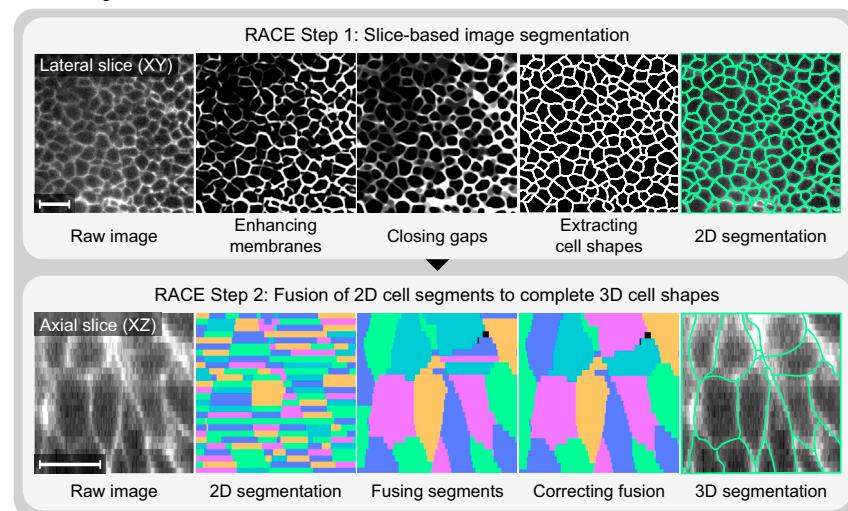
understanding cell-cell interactions, changes in cell morphology during cell migration and differentiation, morphogenesis of tissues and entire embryos, and phenotypic alterations in mutants (Lecuit and Le Goff, 2007; Oates et al., 2009).

Several powerful methods for cell-shape reconstruction have been developed (Fernandez et al., 2010; Khan et al., 2014; Mosaliganti et al., 2012). However, rapid advances in high-performance light microscopy in recent years have outpaced the development of computational methods required to interpret the massive amounts of image data produced by these new microscopes. Limited scalability, resource-intensive computations, and high processing time requirements of state-of-the-art methods for cell-shape segmentation are prohibitive in the analysis of datasets routinely produced by next-generation live imaging methods: for example, processing the terabytes of image data acquired in a single day by a typical light-sheet microscope would take at least several months on a high-end workstation, as shown below. Moreover, the most accurate existing computational approaches are also the slowest and most memory demanding.

To overcome these limitations and greatly improve both speed and accuracy, we developed a new image analysis framework named RACE (Real-time Accurate Cell-shape Extractor), which is capable of rapidly extracting cell shapes from large-scale image data from a variety of fluorescence microscopes and biological model organisms. Inspired by slice-based segmentation approaches for large-scale electron microscopy used in the field of connectomics (Funke et al., 2012; Liu et al., 2014), our method extracts high-quality two-dimensional (2D) segments from individual image slices first, then rapidly merges them to complete three-dimensional (3D) cell shapes using discrete combinatorial optimization techniques. Segmentation quality is further improved using post-processing heuristics. We systematically optimized our computational framework for both speed and accuracy by developing highly efficient algorithms for all computation-intensive processing operators that take full advantage of modern multi-core processors (CPUs) and graphics processors (GPUs).

We demonstrate the high speed, high accuracy, and generality of our method by segmenting cell shapes in fruit fly, zebrafish, and mouse embryos recorded with light-sheet and confocal fluorescence microscopes. We showcase the importance of these

A RACE segmentation framework



the TGMM framework (Amat et al., 2014) and for incorporating manual cell annotations created with the CATMAID framework (Saalfeld et al., 2009). Scale bar, 20 μ m.

(B) Compared with state-of-the-art membrane segmentation algorithms (Fernandez et al., 2010; Khan et al., 2014; Mosaliganti et al., 2012), RACE provides improved segmentation quality while reducing computation time by up to two orders of magnitude. Average false discovery (left) and false negative rates (middle) were calculated on multiple model organisms and microscopes using manually annotated ground truth images (Experimental Procedures). Processing time (right) is shown for a time-lapse image dataset of *Drosophila* embryonic development, comprising 3,836 time points (830 gigabytes). RACE processing time is based on the reconstruction of the complete dataset, whereas ACME, EDGE4D, and MARS benchmarks are based on extrapolation from an early developmental time point, and represent underestimates of the required total processing time.

strengths by analyzing typical terabyte-sized light-sheet live imaging datasets. Using RACE, we reconstructed, for the first time, cellular-resolution tissue anisotropy maps of fast tissue invagination processes in early *Drosophila* development at the whole-embryo level. RACE also facilitated a fully automated, quantitative comparison of cell-shape dynamics during *Drosophila* gastrulation in wild-type and *bicoid nanos torso-like* (*bnt*) mutant embryos (Blankenship and Wieschaus, 2001; Nüsslein-Volhard et al., 1987). We furthermore integrated the RACE framework for automated cell-shape segmentation with our TGMM (tracking with Gaussian mixture models) framework for automated cell-lineage reconstruction (Amat et al., 2014) and demonstrate joint 3D cell-shape segmentation and cell tracking in *Drosophila* whole-embryo image datasets. RACE is publicly available as an open-source software package that includes computational tools for efficient manual data curation, thus supporting applications that require fully validated, error-free image segmentation results.

Design

Advances in state-of-the-art live imaging methods, such as light-sheet microscopy, have made it possible to capture the cellular

Figure 1. RACE Framework for Fast and Accurate Cell-Shape Segmentation

(A) The RACE framework performs 3D cell segmentation in two main steps. In the first step (top), raw 3D image data are processed slice by slice, using efficient strategies for enhancing cell membranes and closing remaining gaps in the enhanced membrane signal, followed by the extraction of the membrane outlines themselves. This initial step yields a high-quality 2D segmentation of cell shapes present within each image slice. In the second step (bottom), these 2D cell segments are merged to form complete 3D cell shapes. Note that the second panel from the left shows a cross-section of the 2D cell segments (using a random color code that avoids collisions between neighbors), i.e. each row of pixels corresponds to the results of slice-based watershed segmentation of one image plane from the first processing step (see Figure S3). The merging of cell segments starts at the sites of "seeds" marking the locations of individual cells (see Figure 2). Further segments are then added to these seed segments by evaluating segment similarities in neighboring image slices. Finally, two fusion heuristics are utilized to correct mistakes and further improve 3D segmentation quality. Seed points for the fusion of cell segments can be extracted either directly from the membrane image data or from additional nuclei image data (if available). RACE furthermore provides options for taking advantage of seed points obtained by cell tracking with the CATMAID framework (Saalfeld et al., 2009).

dynamics underlying the development of tissues and even entire embryos at high spatiotemporal resolution and over extended periods of time. Such experiments produce datasets of unprecedented complexity and size, and demand new computational methods capable of extracting biologically relevant information in a robust and automated manner. To efficiently extract 3D cell-shape information from such fluorescence microscopy images, we developed the RACE automated segmentation framework (Figures 1 and S1). Our design goals included high accuracy, high speed, ease of use, and scalability, ensuring that the resulting framework is also suitable for large-scale applications, such as automated reconstruction of cell shapes in whole-embryo image datasets acquired at high spatiotemporal resolution and over long periods of time.

RACE uses two steps to extract 3D cell-shape information from image data of fluorescently labeled cell membranes, and can optionally take advantage of additional cell nuclei markers to further improve segmentation accuracy.

In the first step, RACE rapidly segments the 3D image data slice by slice, thus producing a set of accurate but unconnected 2D cell segments (Figure 1A). This is achieved by first enhancing the membrane signal in the images, such that membranes

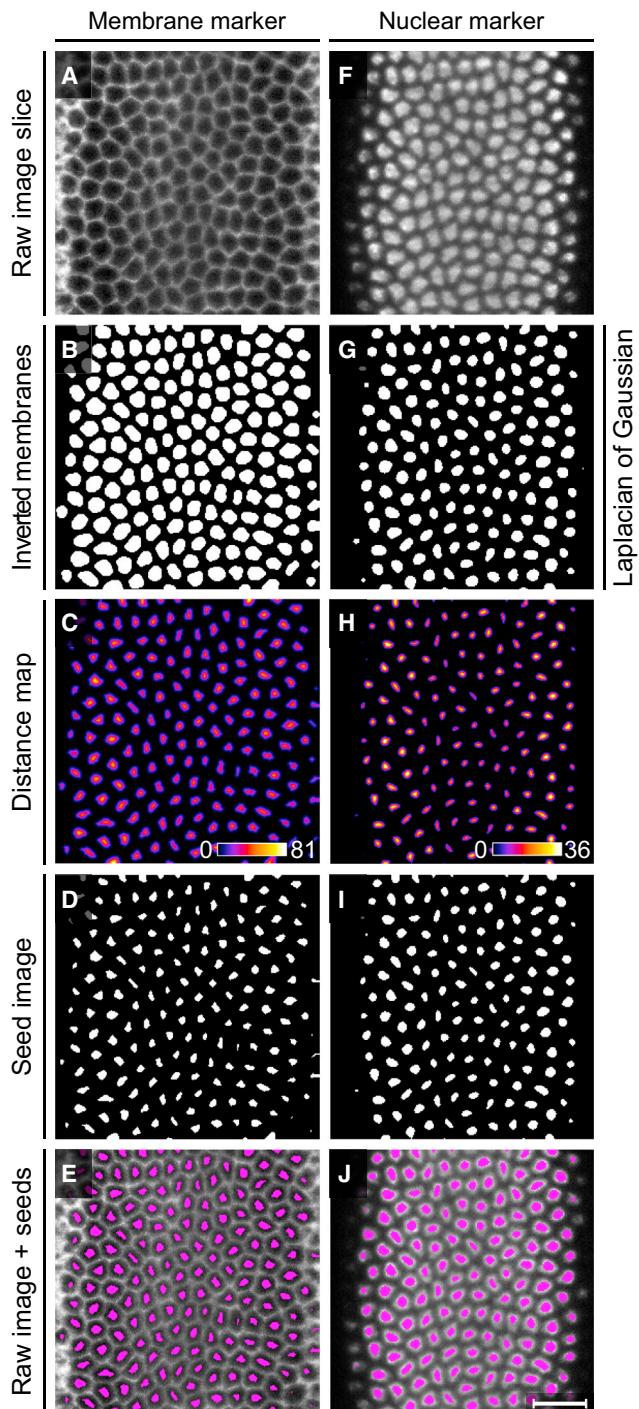


Figure 2. Seed Extraction from Membrane or Nuclei Image Data
 Illustration of key processing steps for seed extraction from a membrane channel (A–E) or a nuclear channel (F–J) of a fluorescently labeled *Drosophila* embryo imaged with SimView microscopy. The enhanced, inverted, and binarized membrane image (B) or a binarized, filtered version of the nuclei image (G, using a Laplacian-of-Gaussian filter) serve as initial cell markers. Erroneously connected objects in the binary images (B and G) can be efficiently separated by calculating the distance of each foreground pixel to the closest background pixel (C and H, using a squared Euclidean distance map) and by extracting small binary regions at all local maxima of this distance map (D and I). Labeling each connected component of (D and I) with a unique identifier

appear as smooth, bright structures, whereas intra- and extracellular, non-membrane regions appear dark (Figures 1A and S2). In some cases, however, holes can persist in the membrane signal, e.g. due to limited image quality or inhomogeneity of the fluorescent label. Therefore, RACE subsequently closes such remaining membrane gaps iteratively, using gray-scale morphology (Figures 1A and S2). The 2D cell segments themselves are then extracted from the enhanced images using a fast and memory-efficient, slice-based implementation of the watershed segmentation algorithm (Figures 1A and S2).

In the second step, RACE combines the 2D cell segments extracted in the first step to complete 3D cell shapes, based on the similarity of segments across image slices (Figures 1A and S3). The fusion of 2D segments is guided by “seeds,” i.e., small, well-separated 3D markers inside each cell, which are automatically computed from the membrane images themselves or, if available, from additional nuclei image data (Figure 2 and Movie S1). To maximize both speed and memory efficiency of the RACE framework, we systematically analyzed the bottlenecks of this processing workflow and developed highly efficient CPU-optimized or GPU-accelerated code modules for all computation-intensive operations (Table S1). A detailed technical description of the computational framework is provided in [Experimental Procedures](#) and [Supplemental Experimental Procedures](#).

RESULTS

Fast and Accurate Cell Segmentation in Fruit Fly, Zebrafish, and Mouse Embryos

To demonstrate the versatility of our method, we used RACE for 3D cell segmentation in live imaging data from three widely used model systems, namely fruit fly, zebrafish, and mouse embryos, recorded with two different imaging modalities, including a light-sheet microscope (Tomer et al., 2012) as well as Carl Zeiss LSM 710 and LSM 510 confocal microscopes. These six classes of image datasets (Figure 3 and Movie S2) cover a wide range of image analyses challenges, owing to heterogeneity in cell shape, cell size, and cell density within and across biological model systems. Moreover, these datasets exhibit substantial differences in image contrast, signal-to-noise ratio, and optical aberrations, which result from differences in sample preparation, microscope specifications, and changes in local optical properties across large multi-cellular specimens. To quantitatively assess the performance of our algorithm, we compared the automatic reconstructions produced by RACE to manually annotated cell-shape data (also referred to as “ground truth annotations” throughout this paper). Our comparison follows the standard performance measures defined by Coelho et al. (2009) and also includes side-by-side performance comparisons with three state-of-the-art segmentation methods: ACME (Mosalignani et al., 2012), EDGE4D (Khan et al., 2014), and MARS (Fernandez et al., 2010).

yields the final seeds that are used to initialize the 3D fusion of segments illustrated in Figure 1A. The last row (E and J) shows the raw image with superimposed seeds. See also [Experimental Procedures](#) and part 1 of [Supplemental Experimental Procedures](#) for algorithmic details of the seed detection. Scale bar, 20 μm.

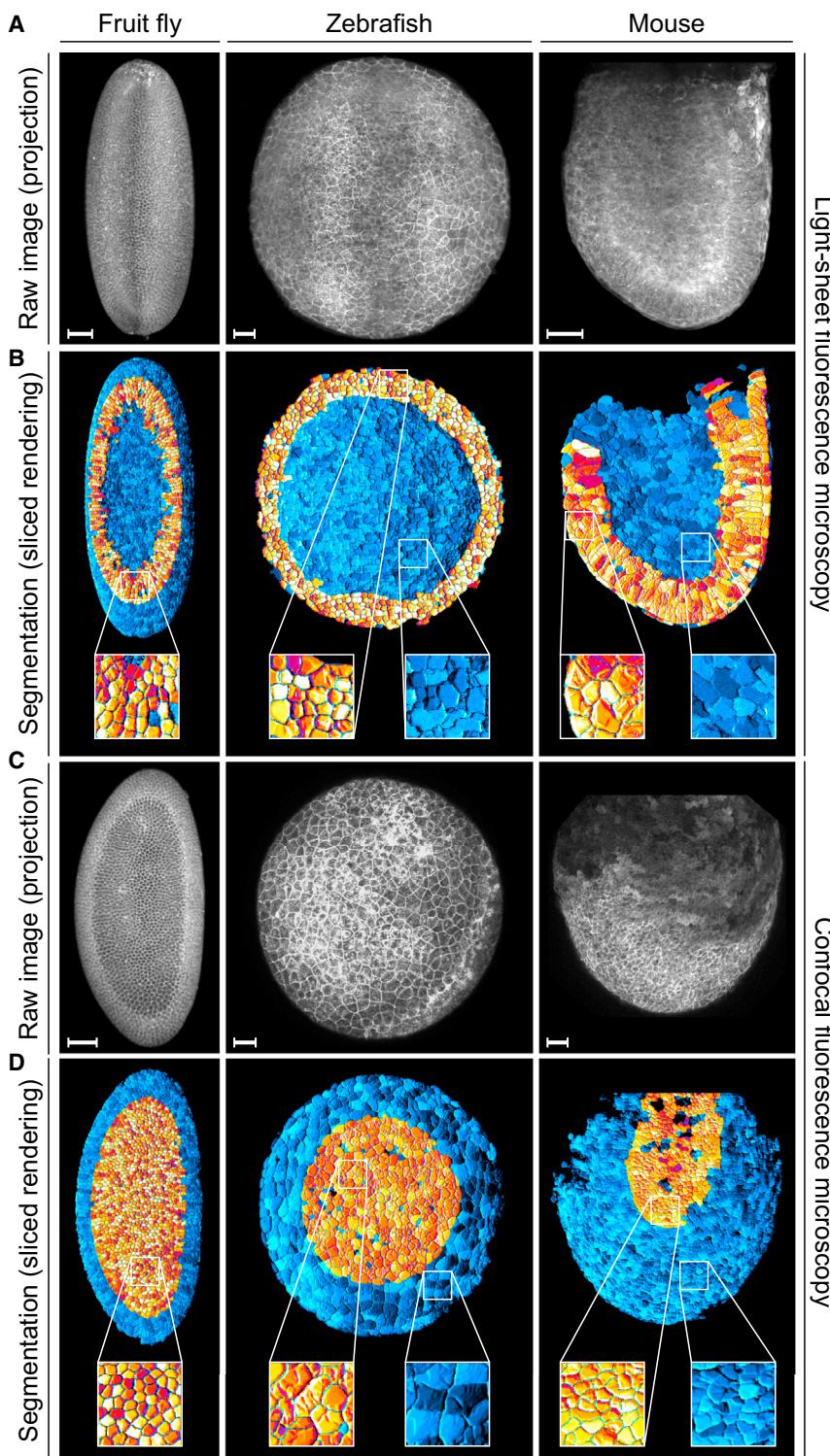


Figure 3. Whole-Embryo Cell Segmentation in Fruit Fly, Zebrafish, and Mouse

(A) Maximum-intensity projections of 3D image datasets of fruit fly, zebrafish, and mouse embryos expressing fluorescent markers labeling all membranes. Image data were acquired with light-sheet fluorescence microscopy (Tomer et al., 2012). The fruit fly embryo was imaged at 3 hr after egg laying (AEL), the zebrafish embryo at 6 hr post fertilization (hpf), and the mouse embryo at E6.5 (Experimental Procedures). Scale bar, 50 μ m.

(B) Segmentation results, visualized as renderings of sliced embryos. Cells in the exposed cross-sections are shown in an orange/red color scheme, and in a cyan/blue color scheme for the rest of the embryo. Insets show enlarged views of the cell segmentation results.

(C) As in (A), but for fruit fly, zebrafish, and mouse embryos imaged with confocal fluorescence microscopes (Zeiss LSM 510 and LSM 710 microscopes). The fruit fly embryo was imaged at 3 hr AEL, the zebrafish embryo at 6 hpf, and the mouse embryo at E7.5 (Experimental Procedures). Scale bar, 50 μ m.

(D) Segmentation results for the image data shown in (C).

tive rates, RACE is 330 times faster and two times more accurate than the most accurate existing method (Khan et al., 2014); compared with the fastest existing approach (Fernandez et al., 2010), RACE is 55 times faster and five times more accurate (Figures 4 and S4, Tables S2 and S3, and parts 1 and 2 of Supplemental Experimental Procedures). Our algorithm scales linearly with dataset size and provides real-time capability by reaching processing speeds exceeding the speed of data acquisition in all datasets (Figures S4C and S4D). For fruit fly, zebrafish, and mouse embryos imaged with SIMView light-sheet microscopy, image acquisition took 40 s, 124 s, and 47 s, respectively, whereas image processing was completed within only 20 s, 44 s, and 31 s, respectively. For fruit fly, zebrafish, and mouse embryos imaged with confocal microscopy, image acquisition took 29 s, 120 s, and 85 s, respectively, whereas image processing was completed within only 9 s, 23 s, and 20 s, respectively. Thus, RACE image processing was on average 3.2-fold faster than the image acquisition

process itself. At the same time, only three parameters of the RACE framework needed to be adjusted across all datasets analyzed in this study, which allows easy and effective application of RACE to new image data (Figure S1 and Table S2).

To evaluate RACE processing speed in large time-lapse datasets recorded with light-sheet microscopy, we measured the

RACE provided the best performance across all algorithms, yielding low false-discovery rates (Figure 1B), low false-negative rates (Figure 1B) and high segmentation quality at the whole-embryo level for all investigated model organisms and microscopes (Figure 3 and Movie S2). Based on these measurements of average processing time, false-discovery rates, and false-nega-

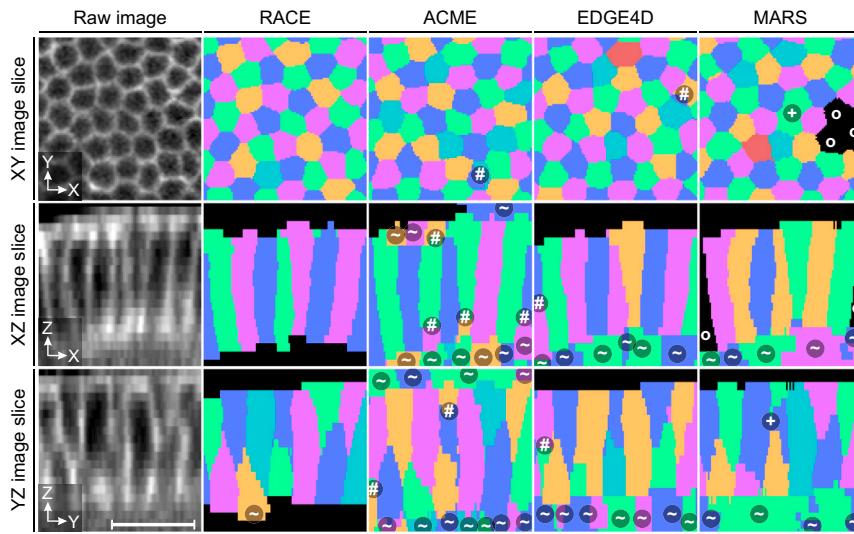


Figure 4. Comparison of Cell-Segmentation Quality in *Drosophila*

Comparison of segmentation quality obtained with RACE, ACME, EDGE4D, and MARS, respectively, in representative image regions of a *Drosophila* embryo. The panels show xy, xz, and yz slices of the raw images as well as the corresponding segmentation label images. Superimposed symbols highlight segmentation errors, including merged (+), split (#), missing (o), and added (~) cells. RACE and EDGE4D provided relatively high segmentation quality. EDGE4D produced a few split cells and some false-positive detections in background regions. ACME and MARS results suffered from leakage into background regions, merged cells, and false-positive detections in background regions. As the size of false-positive objects produced by ACME, MARS, and EDGE4D was comparable with the size of actual cells, purely size-based object filtering was not sufficient to curate the data in these cases.

Scale bar, 20 μ m.

total computation time required to process high-speed, long-term recordings of *Drosophila* embryogenesis. RACE processed typical terabyte-scale time-lapse datasets on a single computer workstation within 1.4 days, whereas existing methods would require 100–700 days for the same task (Figures 1B and Movies S3, S4, and S5). This combination of high processing speed and high segmentation quality opens the door to quantitative image analysis of large developing embryos as demonstrated in the following sections.

Capabilities and Limitations of Fully Automated Image Segmentation with RACE

RACE is designed to extract cell-shape information from image data obtained with various types of fluorescence microscopes and from various biological model organisms in a completely automated manner. As detailed above, this automated workflow produces exceptionally low error rates and achieves data throughput rates suitable for real-time application. To evaluate the potential of RACE in directly providing accurate quantitative information on derived cell-shape features without prior data curation, we systematically compared such features obtained from automated segmentation results with manual ground truth annotations for a variety of cell-shape analyses and organisms. The 22 cell-shape features considered in this comparison include the most commonly used parameters in developmental studies, including measurements of cell volume, centroid position, shape eccentricity, spatial extent, perimeter length, solidity, and cell-shape anisotropy (see [Experimental Procedures](#) for feature descriptions). When comparing automatically extracted features with ground truth annotations at the level of individual cells, we obtained an average deviation of $9.8\% \pm 4.9\%$ across all shape features and model organisms (mean \pm SD, n = 22 features evaluated in three types of datasets; Figures 5 and S5, and Table S3). When performing the same comparison at the level of small, local groups of cells (combining data from on the order of 100 cells), the deviation of automatically extracted cell-shape features and ground truth annotations was further reduced to $2.8\% \pm 2.1\%$ (mean \pm SD, n = 22 features evaluated in three

types of datasets; Figures 5B and S5, and Table S3). This analysis thus confirms that RACE is capable of accurate, automated measurements of cell-shape properties at the single-cell level and that accuracy and robustness of such measurements can be further improved by combining statistics across small groups of cells (Figures 5D and S5D, and Table S3). Importantly, the magnitude of these statistical errors inherent to fully automated analyses with RACE is significantly smaller than relative cell-shape changes occurring during typical morphogenetic processes in embryonic development. Thus, manual data curation is usually not required to resolve cell-shape changes in the course of developmental events or to reliably detect phenotypic differences in corresponding anatomical domains of wild-type and mutant embryos (see section “[Reconstructing Cell-Shape Dynamics during Drosophila Gastrulation](#)”). To demonstrate these key capabilities of RACE, we systematically compared RACE segmentation results with manual ground truth annotations as a function of time throughout *Drosophila* gastrulation as well as in corresponding blastoderm regions of gastrulating wild-type and *bnt* mutant embryos (Figures S6 and S7). We found that RACE and manual cell-shape annotations were in good agreement not only with respect to average values of typical cell-shape parameters; good correspondence was also achieved for the width (or SD) of parameter distributions across local cell populations (Figure S7).

Across all experiments and analyses, the largest deviation between automated segmentation results and ground truth were observed for cell-shape features that rely on good axial segmentation accuracy (e.g. cell volume, cell size along z axis, cell perimeter length and shape anisotropy in xz or yz image cross sections; see Table S3). In these cases, however, manual data annotation itself was frequently problematic (and, in some extreme cases, ambiguous) due to low axial resolution, low image quality as a result of light scattering and aberrations in deeper tissue regions, weak signals from en face membranes in multi-layered tissues, and high background levels arising from autofluorescence. Such limitations in image quality inevitably constrain cell-shape quantifications, irrespective of the

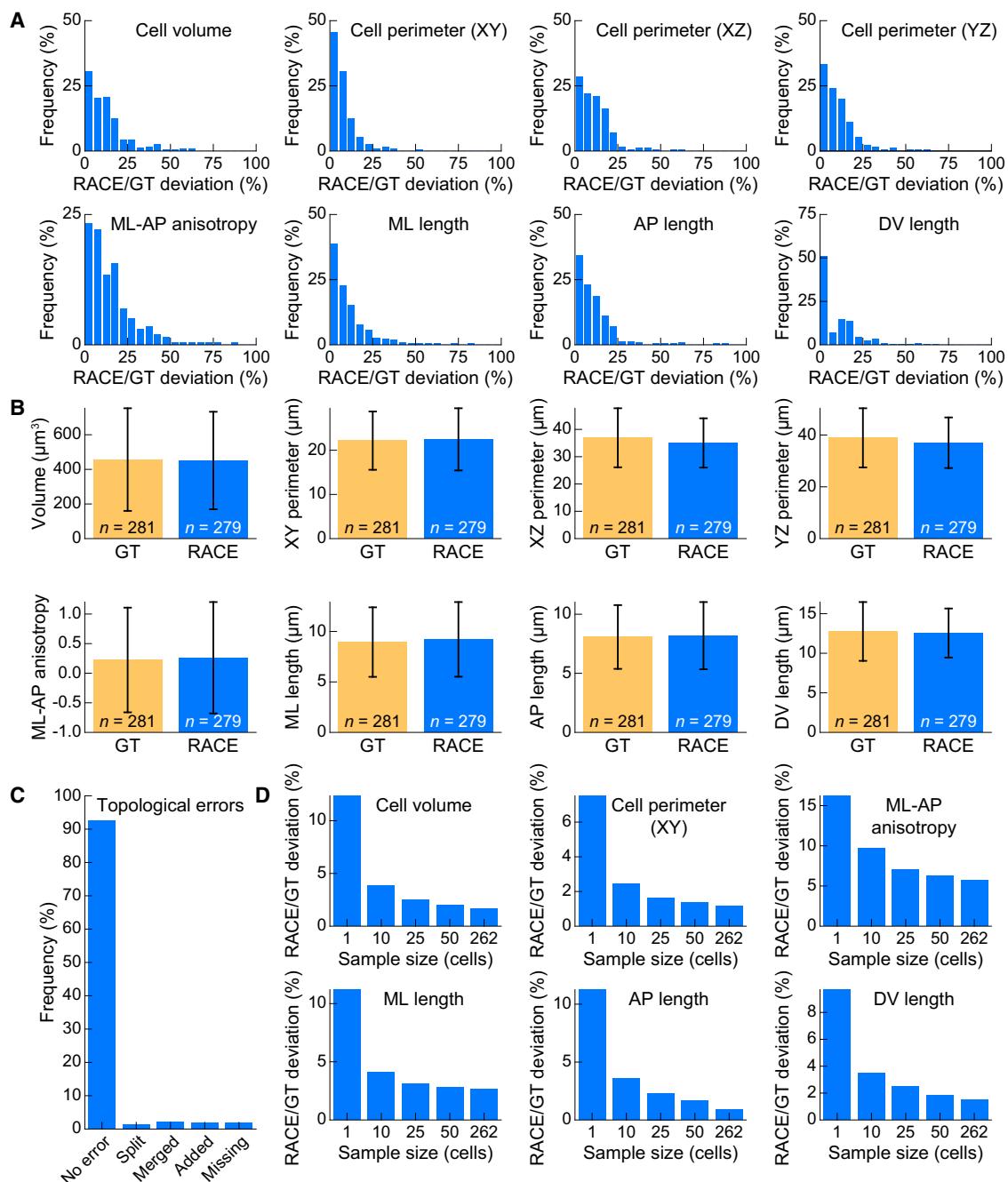


Figure 5. Accuracy of Cell-Shape Information Extracted from SiMView *Drosophila* Images

Comparison of cell-shape features extracted by RACE with results obtained by manual ground truth (GT) annotations in a *Drosophila* embryo. Features include cell volume (number of voxels), cell perimeter (number of surface voxels in xy, xz, and yz planes at the cell centroid location), mediolateral versus anteroposterior cell-shape anisotropy (ML-AP anisotropy), mediolateral (ML) cell size, anteroposterior (AP) cell size, and dorsoventral (DV) cell size. As a measure of RACE accuracy, the absolute deviation of RACE and ground truth results was determined ("RACE/GT deviation," shown in percent) by matching segmented cells with $\geq 50\%$ voxel overlap.

(A) Histogram of the absolute feature value deviation between automatic segmentation results and ground truth annotations, using a bin width of 5%. The deviation of cell-shape feature values is on average below 12% (tab 4 of Table S3).

(B) Bar plots of mean feature values measured across multiple regions of interest for both ground truth and automated RACE cell segmentation results. Error bars correspond to one SD across the analyzed group of cells.

(C) Frequency of topological errors in RACE cell segmentation results, categorized as split, merged, added, and missing cells.

(D) Deviation between RACE and GT annotations for six types of cell-shape parameters, shown as a function of sample size used for estimating average feature values. The special cases $n = 1$ and $n = 262$ show mean deviation of cell-shape features at the single-cell level and averaged across all annotated cells, respectively. For settings in between these cases ($n = 10$, $n = 25$, and $n = 50$), 1,000 groups of matching cell pairs were randomly selected and results were averaged over all groups. Additional features and results are presented in numerical form in tab 4 of Table S3.

performance of the image analysis methodology. We note that recent advances in whole-embryo imaging techniques, such as the development of light-sheet microscopy for isotropic multi-view imaging (Chhetri et al., 2015), overcome previous limitations in axial resolution and thus have the potential to further strengthen manual as well as automated cell-shape analyses at the whole-embryo level.

Finally, we note that the finite error rate of RACE can introduce a need for manual correction in applications with particularly stringent requirements. This concerns analyses in which (1) single cells need to be faithfully tracked over long periods of time while providing error-free information on 3D cell shape along the time axis, or (2) 3D cell shapes of a large group of cells need to be reconstructed at a single time point without errors throughout the data pool. Such applications are currently beyond the capabilities of fully automated computer vision techniques. However, we provide tools for efficient data curation that complement the RACE core framework in such scenarios, as discussed in the section “Efficient Manual Data Curation.”

Reconstructing Cell-Shape Dynamics during *Drosophila* Gastrulation

Accurate information on cell-shape changes during large-scale tissue reorganization is invaluable for the study of the biophysical mechanisms underlying embryogenesis (Bosveld et al., 2012; He et al., 2014; Osterfield et al., 2013; Pare et al., 2014; Tamada et al., 2012; Wang et al., 2012). To evaluate the potential of RACE to yield quantitative information on cell-shape dynamics at the whole-embryo level automatically, we segmented cells in time-lapse datasets of developing *Drosophila* embryos. In our first demonstration, we applied our method to a time-lapse recording of *Drosophila* gastrulation. In this imaging experiment, multiple views of the *Drosophila* embryo were acquired at 3-min intervals using light-sheet microscopy (Tomer et al., 2012). We selected a 1-hr period covering the formation of the ventral furrow, which gives rise to mesoderm and anterior endoderm (Sweeton et al., 1991). Starting at 3 hr after egg laying (AEL), the ventral furrow starts to form as a result of cell-shape changes in a narrow band of cells along the ventral midline. In total, approximately 800 cells are internalized through the ventral furrow. At the same time, the apical surfaces of cells in more lateral locations experience forces directed toward the site of the ventral furrow, which leads to an increase in mediolateral-versus-anteroposterior cell-shape anisotropy on the ventral side of the embryo (Sweeton et al., 1991). We used RACE to automatically segment cell shapes across the embryo during this process and to generate cellular-resolution tissue anisotropy maps as a function of time (Figure 6A and part 4 of [Supplemental Experimental Procedures](#)). In addition, we performed a computational reconstruction of a second dataset covering early embryonic development at high temporal resolution (from 3 to 6 hr AEL at 15-s intervals), and include a video of the resulting cellular-resolution tissue anisotropy maps in [Movie S3](#). This terabyte-sized time-lapse dataset captures the dynamic behavior of more than 5,000 cells over a period of 720 time points, which can be processed by the GPU-accelerated version of RACE within 3.9 hr on a single computer workstation ([Table S1](#)), a task that is not feasible with existing computational methods.

We used the resulting whole-embryo tissue anisotropy maps to quantitatively follow cell-shape dynamics in the vicinity of the ventral midline, and evaluated the accuracy of results provided by RACE by quantitative comparison to manual ground truth annotations. Figure 6B shows a quantitative analysis of cell-shape dynamics in an 80- μm -wide corridor along the ventral midline (region of interest labeled ROI 1 in Figure 6A), which provides information on average anisotropy (left), average cell volume (middle), and average cell size along dorsoventral, mediolateral, and anteroposterior axes (right) as a function of time. Average mediolateral-versus-anteroposterior cell-shape anisotropy rapidly increases for 6 min, starting at the onset of gastrulation at 3.00 hr AEL, and remains at an elevated level of 0.476 ± 0.017 (mean \pm SEM, n = 1,101 cells) for 12 min, before returning to a baseline of 0.111 ± 0.011 (mean \pm SEM, n = 959 cells) over the course of the next 15 min. These changes in shape anisotropy are the result of forces directed toward the ventral furrow, and lead to an increase in average mediolateral cell length from $11.07 \pm 2.95 \mu\text{m}$ (mean \pm SD, n = 1,169 cells) to $17.61 \pm 5.62 \mu\text{m}$ (mean \pm SD, n = 1,131 cells) from 3.00 to 3.25 hr AEL. Thereafter, average mediolateral cell length decreases and converges to the level of average anteroposterior cell length. Average cell height (length along the apico-basal axis) increases prior to the onset of gastrulation and reaches a peak value of $21.04 \pm 7.53 \mu\text{m}$ (mean \pm SD, n = 1,169 cells) at 3.00 hr AEL. Together, these cell-shape changes lead to a steady increase in average volume of cells located at the ventral surface in the vicinity of the ventral furrow, from $534 \pm 199 \mu\text{m}^3$ (mean \pm SD, n = 1,059 cells) at 2.75 hr AEL to $842 \pm 413 \mu\text{m}^3$ (mean \pm SD, n = 854 cells) at 3.50 hr AEL.

To assess the precision of RACE in this fully automated analysis of cellular-resolution tissue anisotropy and to compare the performance with state-of-the-art cell-shape segmentation methods, we manually segmented a representative region of the embryo (region of interest labeled ROI 2 in Figure 6A) at 3.25 hr AEL and processed this annotated region with RACE, ACME, EDGE4D, and MARS. Both qualitatively and quantitatively, RACE provided the highest segmentation quality and the most accurate measurements of tissue anisotropy, producing results most similar to those obtained in our manual ground truth segmentation (Figures 6C and S6). A side-by-side comparison of cell-shape anisotropy profiles along the mediolateral axis of the embryo shows that RACE reduces root-mean-square errors relative to ground truth annotations by more than 2-fold (Figure 6C). Moreover, RACE provided the most accurate quantification of the decrease in cell-shape anisotropy away from the ventral furrow and toward the lateral sides of the embryo (Figure 6C).

Next, we evaluated RACE’s potential in automatically quantifying cell-shape-related phenotypes in mutant embryos. To this end, we imaged and analyzed cell-shape changes during *Drosophila* ventral fold formation side by side in wild-type and *bnt* mutant embryos lacking anteroposterior patterning. Both embryos were imaged with a SiMView light-sheet microscope at 20-s (wild-type) and 30-s (*bnt* mutant) time intervals from 3.00 to 3.75 hr AEL (Movie S3). The image data and computational reconstructions are presented side by side as maximum-intensity projections (Figure 7A) and maps of mediolateral-versus-anteroposterior tissue anisotropy at the cellular level in

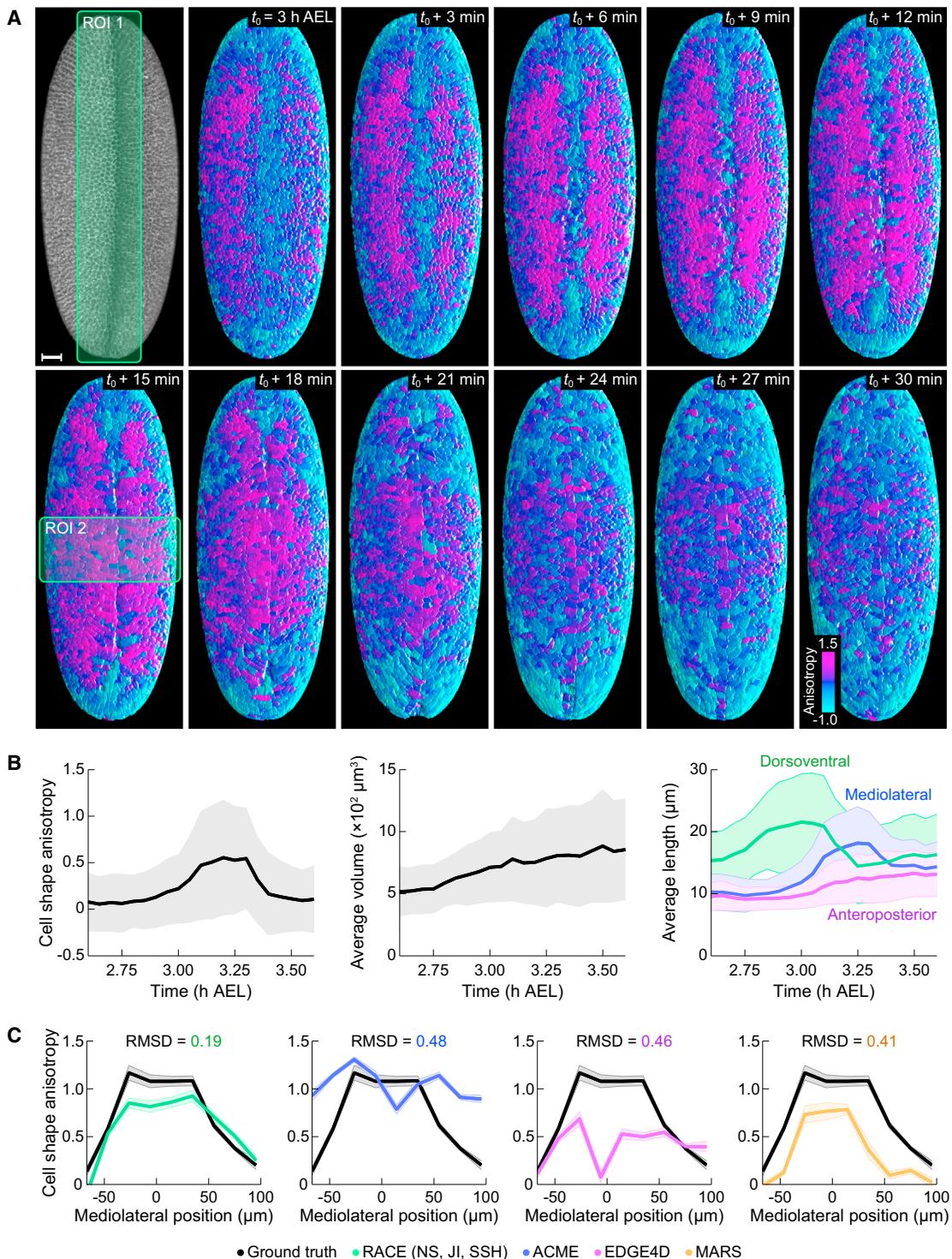


Figure 6. Reconstructing Cell-Shape Anisotropy during *Drosophila* Gastrulation

(A) Cell-shape segmentation and visualization of tissue anisotropy on the ventral side of a *Drosophila* embryo during ventral furrow formation. The embryo expressed a fluorescent marker labeling all cell membranes and was imaged with a SiMView light-sheet microscope at 3-min intervals. The first image of the panel shows a maximum-intensity projection of the image data from the ventral side of the embryo at $t_0 = 3$ hr AEL. The other images show segmentation results over time obtained with RACE. The color code indicates the level of mediolateral-versus-anteroposterior cell-shape anisotropy, and was chosen to highlight in particular cell-shape changes in response to ventral furrow formation: cells elongated along the anteroposterior axis are shown in cyan; uniformly shaped cells are shown in dark blue; and cells elongated along the mediolateral axis are shown in bright magenta. The segmentation data and associated annotation of cell-shape anisotropy reveal local changes in cell shape in the course of large-scale tissue reorganization and epithelial folding, including the formation of the ventral furrow,

(legend continued on next page)

wild-type and *bnt* mutant embryos (Figure 7B; see also part 4 of [Supplemental Experimental Procedures](#)). Manual inspection of the image data shows that shortly after the onset of gastrulation mediolateral-versus-anteroposterior cell-shape anisotropy near the ventral midline rapidly approaches peak levels in wild-type development, whereas corresponding blastoderm cells in the *bnt* mutant exhibit less pronounced mediolateral cell elongation. Mediolateral-antroposterior anisotropy relaxes very slowly toward pre-furrow baseline levels after peak anisotropy levels are reached in the *bnt* mutant embryo. In addition, blastoderm cells have a larger total cell volume and a longer anteroposterior axis in the *bnt* mutant compared with wild-type. These qualitative observations are quantitatively confirmed and further refined by the automated RACE analysis shown in Figure 7C. To assess relative changes in local cell-shape dynamics, we selected a $50 \times 50 \times 60\text{-}\mu\text{m}^3$ region close to the ventral furrow (Figure 7B). Mediolateral-versus-antroposterior cell-shape anisotropy in this region reaches a peak level of 1.00 ± 0.33 (mean \pm SD, n = 26 cells) in the wild-type embryo at 3.31 hr AEL and a peak level of 0.75 ± 0.39 (mean \pm SD, n = 27 cells) in the *bnt* mutant at 3.25 hr AEL. Throughout the analyzed time span, blastoderm cells in the *bnt* mutant are on average $43.0\% \pm 18.6\%$ larger with respect to their total cell volume and also exhibit a $34.0\% \pm 11.3\%$ increased anteroposterior cell size compared with blastoderm cells in the wild-type embryo (mean \pm SD, n = 1,884 and 4,729 cells in *bnt* mutant and wild-type, respectively). Average cell volume increases from $688 \pm 303\text{ }\mu\text{m}^3$ (mean \pm SD, n = 25 cells) at 3.27 hr AEL to $1,461 \pm 793\text{ }\mu\text{m}^3$ (mean \pm SD, n = 24 cells) at 3.73 hr AEL in the *bnt* mutant, whereas average cell volume in the wild-type embryo increases from $523 \pm 234\text{ }\mu\text{m}^3$ (mean \pm SD, n = 41 cells) at 3.24 hr AEL to only $817 \pm 302\text{ }\mu\text{m}^3$ (mean \pm SD, n = 33 cells) at 3.75 hr AEL. Similarly, anteroposterior cell size increases from $10.5 \pm 2.0\text{ }\mu\text{m}$ at 3.34 hr AEL (mean \pm SD, n = 24 cells) to $13.5 \pm 3.7\text{ }\mu\text{m}$ at 3.75 hr AEL (mean \pm SD, n = 24 cells) in the *bnt* embryo, and from $7.9 \pm 2.0\text{ }\mu\text{m}$ at 3.32 hr AEL (mean \pm SD, n = 24 cells) to $10.5 \pm 1.9\text{ }\mu\text{m}$ at 3.71 hr AEL (mean \pm SD, n = 33 cells) in the wild-type embryo (Figure 7C). By contrast, dynamic changes in anteroposterior-versus-dorsoventral cell-shape anisotropy appear relatively closely matched in wild-type and *bnt* mutant embryos (Figure 7C). A side-by-side visualization of the cellular-resolution tissue anisotropy maps for wild-type and *bnt* mutant development are included in [Movie S4](#). Importantly, these cell-shape changes as a function of time and corresponding differences in wild-type and *bnt* mutant cell-shape dynamics are quantitatively revealed by RACE in a completely automated manner, without the need for manual data curation. RACE segmentation errors and associated errors in the estimation of these cell-shape features are considerably

smaller compared not only with phenotypic differences at the cellular level in wild-type and *bnt* mutant development but also with relative changes in cell shape as a function of time: by manually annotating 11 groups of cells and comparing nine cell-shape features extracted by RACE to ground truth data (Figure S7), we found that relative changes in cell-shape features in wild-type embryos were by a factor of 12.7 ± 6.9 larger than RACE measurement errors, changes in cell-shape features in *bnt* mutants were by a factor of 15.0 ± 10.2 larger than RACE measurements errors, and relative differences in cell-shape features between wild-type and *bnt* mutants were by a factor of 8.6 ± 3.0 larger than RACE measurement errors (mean \pm SD, n = 9 features).

Joint Reconstruction of Cell Lineages and Cell Morphology

The RACE framework is capable of segmenting cells but does not track their identity as a function of time. By default, RACE assigns a unique identifier to each segmented cell in an image, but these identifiers are not necessarily consistent over time. To extend this functionality to the joint reconstruction of cell lineages (i.e. cell tracking through cell divisions) and cell-shape information, we integrated the RACE framework with our TGMM algorithm for high-throughput cell lineageing ([Amat et al., 2014](#)) ([Experimental Procedures](#)). We tested this combined methodological framework in a *Drosophila* whole-embryo time-lapse recording. We imaged cell nuclei and cell membranes with a SiMView light-sheet microscope (Figure 8A) and acquired multiple views of the embryo at 20-s intervals over a period of 2 hr, starting at 2.25 hr AEL. Following image acquisition and multi-view image fusion ([Amat et al., 2015](#)), we performed cell tracking with TGMM, which offers 97.3% linkage accuracy for fluorescently labeled cell nuclei ([Amat et al., 2014](#)). Information about segmented nuclei and their temporal associations were then propagated as seed points to the RACE framework. Finally, we segmented cell shapes with RACE using the fluorescently labeled cell membranes (Figure 8A and [Table S3](#)). In this workflow, the identifiers assigned to cell shapes extracted by RACE match the identifiers obtained by TGMM, which facilitates not only cell tracking but also the automated detection of cell divisions and apoptotic events (or generally any factors that lead to the disappearance of cells from the image data). This approach thus enables the user to evaluate cell-shape changes as a function of time at the single-cell level. A visualization of this imaging experiment and the corresponding cell segmentation and cell tracking results is provided in [Movie S5](#).

The combined functionality of RACE and TGMM allowed us to systematically follow cell-shape dynamics across the embryo

the formation of the cephalic furrow, and germband extension. Computation and color mapping of anisotropy levels were performed as described in part 4 of [Supplemental Experimental Procedures](#). Whole-embryo development and cell-shape anisotropy are shown in high spatiotemporal resolution in [Movie S3](#). Scale bar, 25 μm .

- (B) Time course of the average anisotropy level (left), average cell volume (middle), and average cell size along dorsoventral, mediolateral, and anteroposterior axes (right) for cells in the area indicated by the vertically oriented green rectangle ROI 1 in (A). Shaded areas show SD.
- (C) Quantitative comparison of segmentation quality obtained with RACE, ACME, EDGE4D, and MARS with respect to estimation of average tissue anisotropy along the mediolateral axis in the annotated region of interest (horizontally oriented green rectangle ROI 2 in A). Abscissa values indicate mediolateral position of segmented cells relative to the ventral furrow (0 μm). Ordinate values indicate mean anisotropy estimated by binning cells according to mediolateral position, using a bin width of 20 μm . Shaded areas show SEM. Root-mean-square deviation of automatically reconstructed anisotropy profiles versus ground truth annotation is shown above each graph (lower values indicate better performance of the respective segmentation method).

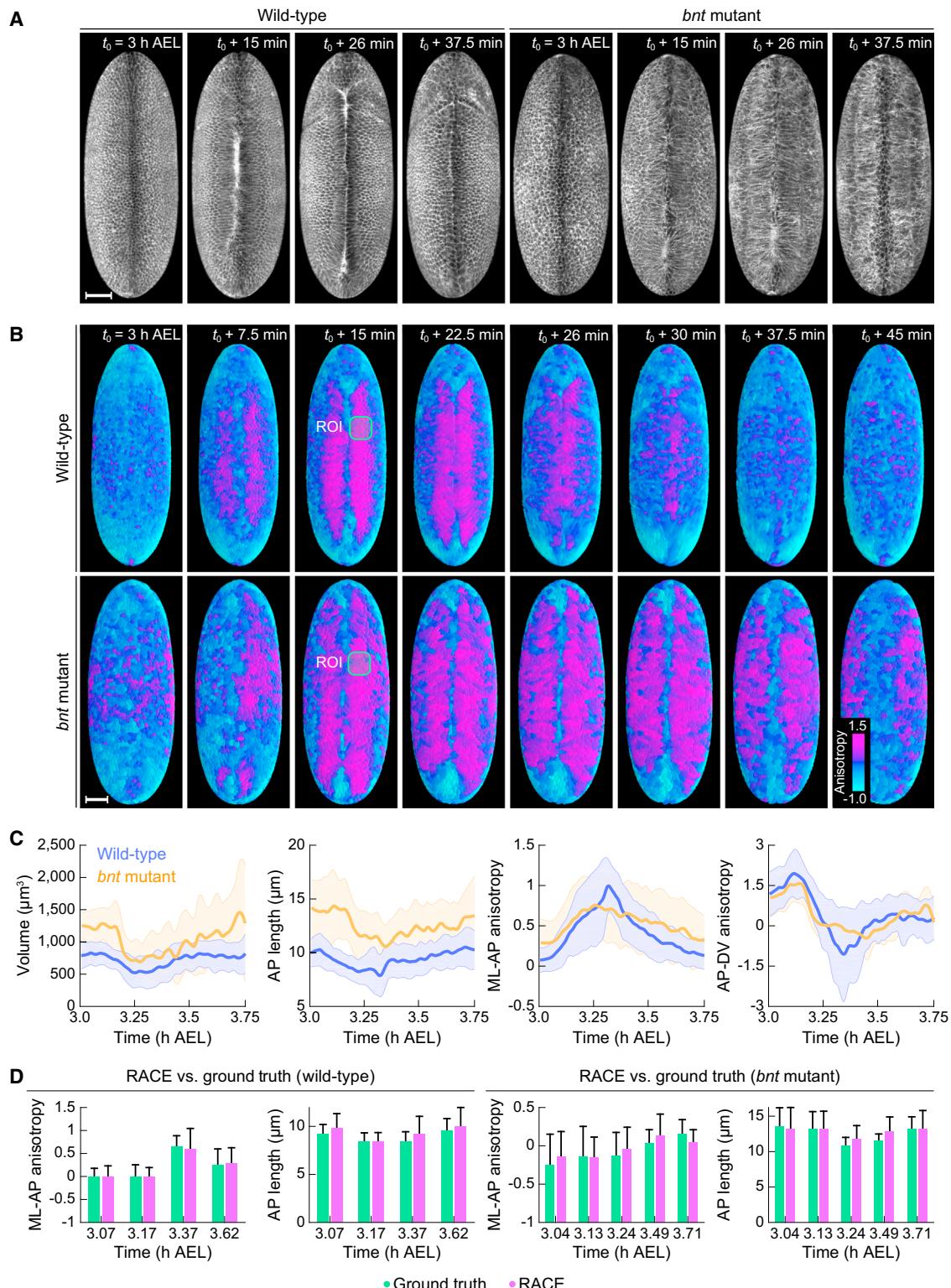


Figure 7. Comparative Analysis of Gastrulation in *Drosophila* Wild-Type and *bnt* Mutant Embryos

(A) Side-by-side comparison of gastrulation in *Drosophila* wild-type and *bnt* mutant embryos expressing fluorescent markers labeling all cell membranes. The image panels show maximum-intensity projections of the ventral half of the embryos, using 3D image data acquired with SiMVView light-sheet microscopy. Whole-embryo imaging was performed at 20-s (wild-type) and 30-s (*bnt* mutant) time intervals from 3.00 to 3.75 hr AEL. A video showing whole-embryo development and cell-shape anisotropy in the *bnt* mutant embryo is included in [Movie S3](#). Scale bar, 50 μ m.

(legend continued on next page)

while keeping track of individual cell identities for up to several hundred time points (Figures 8B and 8C). To illustrate these capabilities, we selected three groups of cells at time point 200 (3.36 hr AEL) of the time-lapse experiment: a population of cells adjacent to the ventral furrow (group 1, n = 89 cells), a dorsal group of cells located directly above the extending germband (group 2, n = 110 cells), and a population of cells on the anterior dorsal side (group 3, n = 77 cells) of the embryo (Figure 8A). As is evident from the image data, cells in these regions undergo substantial changes in cell shape after the onset of gastrulation at 3 hr AEL (Movie S5): in particular, cells in groups 1 and 3 rapidly stretch along the mediolateral axis during ventral furrow formation, whereas cell shapes in group 2 become elongated along the anteroposterior axis as the extending germband approaches their location. The computational reconstruction of cell tracks and cell shapes by RACE/TGMM reveals the underlying cellular dynamics with high spatiotemporal resolution and at the single-cell level (Figure 8B). At the same time, RACE facilitates the extraction of robust statistics on cell-shape changes at the population level: the fully automated computational analysis shown here, which quantitatively follows cell-shape dynamics of in total more than 5,000 cells over 350 time points, enables rapid quantification of cell behavior underlying large-scale morphogenetic events in early development (Figure 8C).

DISCUSSION

RACE Design and Performance

Cell segmentation and analysis of cell-shape changes are key requirements for data mining and data interpretation in many imaging experiments following developmental processes. The RACE segmentation framework presented here provides a solution for fast and accurate reconstruction of cell-shape information from large-scale image datasets. Compared to state-of-the-art methods for cell-shape segmentation, RACE fundamentally improves automated segmentation accuracy and speed while using an efficient algorithmic design that enables high-throughput, large-scale image analysis on a single computer workstation in real time. These advances aid investigations relying on information about cell shape in general, but they also improve performance beyond the critical threshold needed to effectively study cell-shape dynamics at the scale of entire complex embryos. At the same time, we designed RACE to be easy to use by minimizing the need for parameter tuning.

In the following sections we briefly discuss the most important considerations in applying RACE to new image data, provide practical recommendations for optimizing segmentation quality and speed, and assess methodological limitations and pitfalls (see also the RACE user guide and video tutorial provided in Software S1).

High-Throughput and Real-Time Capabilities of the RACE Framework

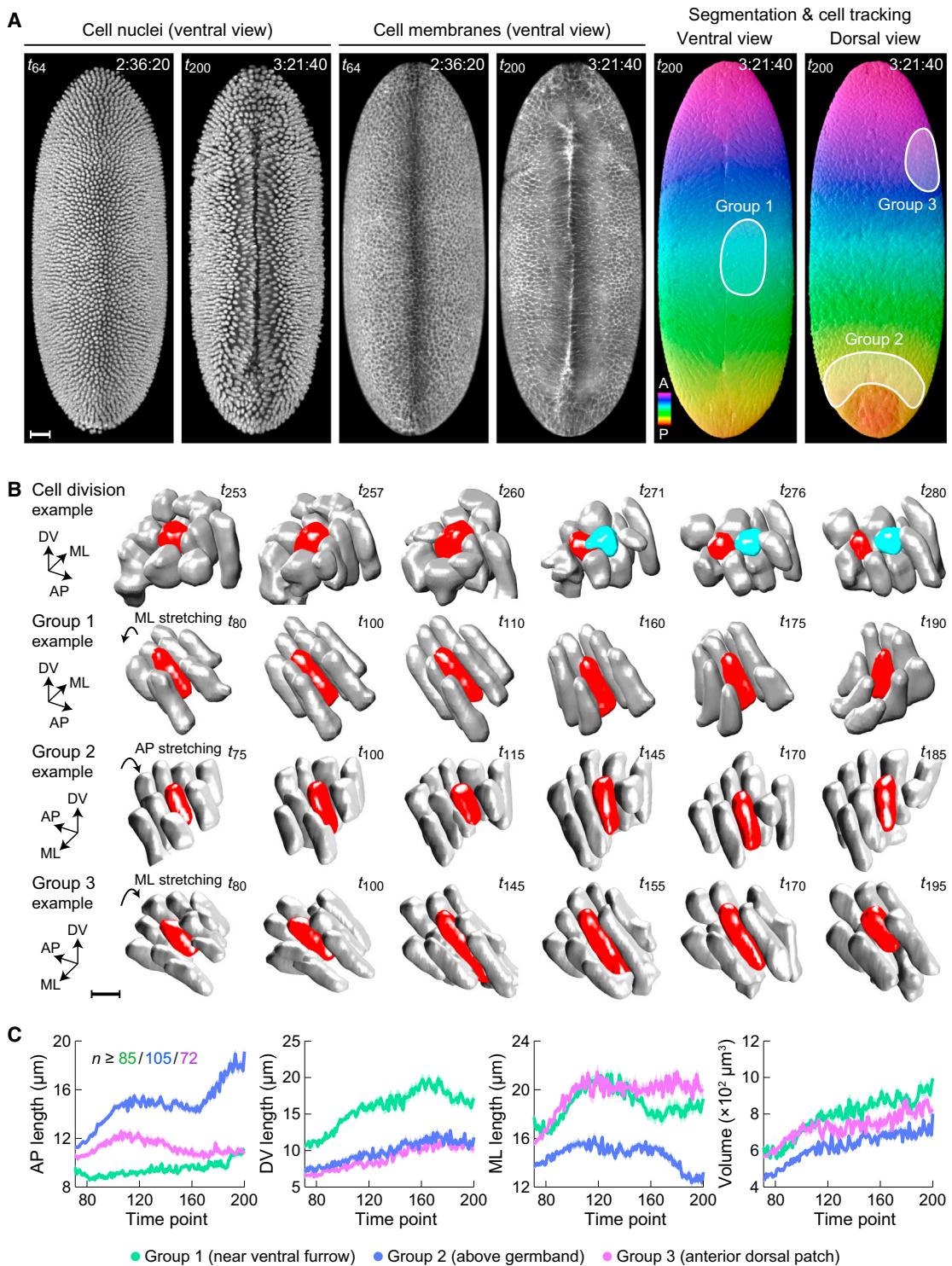
A key design goal in our development of the RACE framework was high computational efficiency, i.e. the realization of high processing speeds while ensuring scalability and minimizing memory consumption. These performance criteria are essential when performing cell-shape segmentation in large multi-cellular organisms consisting of up to tens of thousands of cells or when following cell-shape dynamics in time-lapse image datasets at the terabyte scale. Moreover, the high computational efficiency of RACE essentially aids all applications as it delivers fast and accurate cell segmentation performance at minimal cost to the investigator: all reconstructions demonstrated in this study were performed on a single computer, without the need for expensive and complex computer cluster environments. In addition to a fast CPU-optimized implementation of the RACE framework, we also developed a GPU-accelerated version of our framework, which offers exceptionally high processing speed using a standard, low-cost CUDA-enabled graphics card.

The speed advances introduced in RACE open the door to new types of experiments that require real-time manipulation of individual cells in developing organisms during an ongoing imaging experiment. For instance, RACE could be used to rapidly and automatically classify cells based on shape parameters or cell-shape changes at any time during an imaging experiment. These cells could then be automatically targeted for laser ablation or other types of optical manipulation, such as light-mediated manipulation of gene expression or photoactivation of fluorescent markers. As a foundation for such applications, we confirmed that RACE segmentation speeds exceed image acquisition speeds in all six imaging experiments shown in Figure 3, which cover confocal microscopy and light-sheet microscopy recordings performed with fruit fly, zebrafish, and mouse embryos. Although a certain fraction of cells would likely be incorrectly classified by RACE in such an online setting (as expected based on the finite

(B) Cell-shape segmentation and visualization of tissue anisotropy on the ventral side of wild-type (top) and *bnt* mutant (bottom) embryos during ventral furrow formation. Starting at $t_0 = 3$ hr AEL, segmentation results obtained with RACE are shown for selected time points over a time span of 45 min. As in Figure 6, the color code indicates the level of mediolateral-versus-anteroposterior cell-shape anisotropy and was chosen to highlight cell-shape changes in response to ventral furrow formation. Computation and color mapping of anisotropy levels were performed as described in part 4 of [Supplemental Experimental Procedures](#). Whole-embryo development and cell-shape anisotropy in both wild-type and *bnt* mutant embryo are shown at high spatiotemporal resolution in Movie S4. Scale bar, 50 μ m.

(C) Cell volume, anteroposterior cell length (AP length), mediolateral-versus-anteroposterior cell-shape anisotropy (ML-AP anisotropy), and anteroposterior-versus-dorsoventral cell shapes (AP-DV anisotropy) as a function of time for a region highlighted in (B) close to the ventral midline. Solid lines represent mean values and shaded areas indicate the SD across all cells in the region of interest (ROI). The first time point (t_0) corresponds to 3.0 hr AEL. Cell-shape information is sampled at 20-s and 30-s intervals for wild-type and *bnt* mutant embryos, respectively.

(D) Evaluation of the accuracy of cell-shape features automatically extracted by RACE, using segmentation data from *Drosophila* wild-type (left) and *bnt* mutant (right) embryos for the period 3.00–3.75 hr AEL. At each time point a group of 20–30 cells was manually annotated, and median values (colored bars) as well as SD (error bars) of cell-shape features across the group were extracted and visualized side by side with the respective results obtained from automatic reconstructions performed with RACE. Results for anteroposterior cell length (AP length) and mediolateral-versus-anteroposterior cell-shape anisotropy (ML-AP anisotropy) are shown here. The evaluation of a wider spectrum of cell-shape features is presented in Figure S7.

**Figure 8. Combined Cell-Lineage Reconstruction and Cell-Shape Segmentation in *Drosophila***

(A) Maximum-intensity projections of a *Drosophila* embryo expressing fluorescent markers labeling nuclei (left) and cell membranes (middle). The embryo was imaged every 20 s in both color channels with a SiMVView light-sheet microscope, starting at 2.25 hr AEL. Using an integrated computational framework combining RACE and TGMM functionality, cells were automatically segmented and tracked. In this workflow, each cell segmented by RACE is assigned a unique (and temporally consistent) identifier corresponding to the identifier of the cell track reconstructed by TGMM. The images to the right show rendered ventral and dorsal views of this computational reconstruction at time point 200, shortly after the onset of germband extension. The color code was assigned at the beginning of the time-lapse dataset to mark the anteroposterior location of each cell, and was subsequently propagated in time using information about automatically reconstructed cell tracks and cell divisions. Cell-shape dynamics in three groups of cells near the ventral furrow (group 1), above the germband (group 2), and in an

(legend continued on next page)

segmentation error rate), our statistical analysis presented in tabs 4–6 of [Table S3](#) suggests that RACE extracts most cell-shape parameters with high accuracy, and should thus be fairly reliable in correctly identifying and designating cells as targets for optical manipulation. The investigator could then simply determine manually, after the experiment, which cells were correctly targeted by RACE and select only those cells for further analysis. Importantly, RACE’s ability to segment 3D cell shapes rapidly and with good accuracy would be crucial to enable statistically meaningful real-time manipulation of developing organisms in the first place: in contrast to a human observer who would be unable to rapidly assess 3D cell shapes, e.g. for the tens of thousands of cells in a gastrulating zebrafish embryo, RACE could facilitate the continuous monitoring and prediction of suitable manipulation targets at the high speed required for imaging cellular dynamics at the scale of entire embryos.

Limitations and Strategies for Efficient Manual Data Curation

Our performance analyses show that automated segmentation results obtained with RACE exhibit exceptionally low error rates ([Figures 1B, 4, 5, 6C, S4A, S4B, and S5–S7; Table S1](#); parts 1 and 2 of [Supplemental Experimental Procedures](#)). By comparing automatically extracted cell-shape features with ground truth annotations, we found that manual results and segmentation data obtained by RACE at the single-cell level deviate on average by $9.8\% \pm 4.9\%$. This average deviation can be further reduced to $2.8\% \pm 2.1\%$ when pooling data from small groups of cells (on the order of 100). Although these percentages are usually significantly smaller than relative changes in cell shape occurring in the course of developmental processes, this level of performance may still not be sufficient for applications that demand completely error-free results. An example of a particularly challenging analysis to this end is the faithful tracking and accurate shape characterization of an individual cell in a developing embryo over long periods of time. Depending on the desired time period, the statistical error rate of RACE can certainly be prohibitive in this scenario.

Although time-consuming, manual data curation is a possible solution to this limitation. To establish a convenient and intuitive infrastructure for manual data curation that enables curation even of relatively large datasets in a relatively time-efficient manner, we developed a specialized interface to the open-source software CATMAID, a Web-based manual data annotation tool designed for large-scale microscopy datasets ([Saalfeld et al., 2009](#)). Our data curation strategy takes advantage of the fact that the most severe segmentation errors can be corrected simply by providing an improved, manually defined seed to the

automated segmentation framework. Instead of manually tracing the entire 3D shape of a cell in a time-consuming manner, the user only needs to provide a single 3D point inside the cell to serve as a new seed. RACE then updates the reconstruction results guided by the revised seeding information. The CATMAID-based graphical user interface for data curation allows rapid browsing of image data and segmentation results, and provides the user with tools for editing and correcting seed positions, which are then imported into the RACE segmentation framework. This approach to data curation is especially effective in cases where seeding is ambiguous, e.g. as a result of low temporal sampling, low image quality, or high local cell density. In addition to supporting such data curation efforts, the RACE segmentation framework allows importing tracking results from the TGMM algorithm for automated cell lineageing, and directly links cell tracks and cell morphology data ([Amat et al., 2014](#)).

Distribution and Ease of Use of the RACE Framework

RACE employs a user-friendly design, provides a graphical interface, and does not require the user of the framework to be an expert in image processing. We provide a detailed step-by-step protocol, a troubleshooting guide, and a video tutorial (Boxes 1 and 2 in the RACE user guide and folder “Video_Tutorial” in [Software S1](#)) to help new users familiarize themselves with the RACE segmentation framework. Tuning of framework parameters for optimizing processing results is straightforward: only three parameters had to be adjusted across all examples presented in this work; these parameters have intuitive meaning and are easy to optimize with minimal time investment in manual data inspection ([Figures S1B–S1E](#) and [Table S2](#)). We furthermore provide interfaces to TGMM ([Amat et al., 2014](#)) and CATMAID ([Saalfeld et al., 2009](#)), which allow importing complementary information about cell tracks and manual annotations, respectively. Our software is open-source and compatible with Windows, Linux, and Mac OS X operating systems (<https://bitbucket.org/jstegmaier/race> and [Software S1](#)).

We envision that RACE will facilitate easy access to quantitative information on cell morphology and cell-shape changes in a wide variety of light microscopy image datasets. The ability to perform rapid and accurate cell segmentation even in relatively large image datasets will be invaluable for reconstructing the developmental building plans of complex multi-cellular organisms, quantitatively evaluating differences in wild-type and mutant development, realizing strategies for real-time analysis and manipulation of cell behavior in living biological specimens, and designing and validating computational models of embryonic development.

anterior dorsal patch (group 3) are quantified and visualized in more detail in (B) and (C). A video showing microscopy images and cell segmentation and tracking data side by side over a period of 2 hr (354 time points) is included in [Movie S5](#). Scale bar, 25 μm .

(B) Four examples of automatically segmented and tracked cells (red and cyan) from the dataset visualized in (A), shown in high resolution together with a subset of their neighboring cells (gray) at six time points. These examples include a dividing cell (row 1) as well as cells from the three populations indicated in (A) (rows 2–4). Scale bar, 10 μm .

(C) Average length of cells along anteroposterior, dorsoventral, and mediolateral axes, as well as average cell volume, for the tracked cell groups indicated in (A) as a function of time. Solid lines represent mean values and SEM are indicated by shaded areas (which do not extend significantly beyond the solid lines). The first time point (t_0) corresponds to 2.25 hr AEL, and cell-shape data are sampled at 20-s intervals. Cell counts reflect the minimum number of cells in each group across the time-lapse dataset.

AP, anteroposterior; DV, dorsoventral; ML, mediolateral.

EXPERIMENTAL PROCEDURES

Sample Preparation and Imaging of Zebrafish Embryos Using SiMView Microscopy

Animal experiments described here and in the following sections were conducted according to protocols approved by the Institutional Animal Care and Use Committee of the Howard Hughes Medical Institute, Janelia Research Campus.

SiMView imaging of zebrafish embryos was performed with a transgenic line homozygous for the nuclear label H2B-eGFP and membrane label lyn-tdTOMO under the control of the β -actin promoter. Embryos were fixed with 4% paraformaldehyde (PFA) at 6 hr post fertilization (hpf; shield stage) overnight before imaging. Fixed samples were embedded in 2.0-mm glass capillaries (inner diameter) filled with 1% low-melting-temperature agarose prepared in PBS and mounted vertically in the imaging chamber containing E3 buffer. Animal and vegetal poles of the embryo were facing the microscope's two detection arms. EGFP was excited at 488 nm and fluorescence was detected using 525–50-nm band-pass filters (Semrock). tdTomato was excited at 561 nm and fluorescence was detected using 561-nm long-pass filters (Semrock). Imaging was performed using Nikon 16 \times /0.8 NA objectives and Hamamatsu Orca Flash 4.0 sCMOS cameras. Image stacks of 410 planes encompassing the entire embryo with an axial step size of 2.03 μ m and lateral pixel size of 0.41 μ m were acquired from four complementary (SiMView-type) views.

Sample Preparation and Imaging of Mouse Embryos Using SiMView Microscopy

SiMView imaging of mouse embryos was performed with CAG-TAG1-positive embryos collected from CD-1 females crossed to CAG-TAG1 males (Trichas et al., 2008), which were maintained on a C57BL/6J background through natural mating. Embryos were dissected at E6.5 in PBS with 10% fetal bovine serum and mounted inside Teflon FEP tubes with 50- μ m wall thickness filled with PBS. Nuclear GFP was excited at 488 nm and membrane-localized tdTomato at 561 nm. Image stacks of 149 planes encompassing the entire embryo were acquired with a step size of 2.03 μ m. SiMView optics were configured as described above.

Sample Preparation and Imaging of Fruit Fly Embryos Using SiMView Microscopy

SiMView imaging of *Drosophila* embryos was performed with embryos homozygous for membrane label spider-GFP (a gift from Eric Wieschaus) and nuclear label His2Av-mRFP1 (Bloomington stock #23560). SiMView imaging was also performed with *bnt* triple mutants lacking functional copies of maternally expressed genes *bicoid*, *nanos*, and *torso-like*, which specify the extremes of the embryonic anterior-posterior axis (Blankenship and Wieschaus, 2001). One-fourth of the females in a population of *w⁺;bnt*, resiliss-GFP/TM3 flies (a gift from Eric Wieschaus) produce embryos with no anterior-posterior patterning and with fluorescently labeled membranes. Wild-type and *bnt* mutant embryos were collected by placing female flies on grape-juice agar plates for 1 hr, dechorionating the embryos with 50% bleach solution, and rinsing the dechorionated embryos in tap water.

Embryos were embedded in 1% low-melting-temperature agarose in a 1.5-mm glass capillary (inner diameter). Imaging began during blastoderm stages, about 2 hr AEL. GFP was excited at 488 nm and fluorescence was detected using 525–50-nm band-pass filters (Semrock). mRFP1 was excited at 594 nm and fluorescence was detected using 594-nm long-pass filters (Semrock). Imaging was performed using Zeiss 20 \times /1.0 NA objectives and Hamamatsu Orca Flash 4.0 sCMOS cameras. Four complementary (SiMView-type) views of the embryo were recorded. The *Drosophila* embryo used in the visualizations in Figures 3, 5, 6, S2, and S6, and in part 1 of Movie S1 was imaged using 0.65- μ m z steps and a lateral pixel size of 0.33 μ m. Due to availability of ground truth annotations, the performance assessments shown in Figures 1, 2, and S4, Movie S2, and tab 1 of Table S3 were performed on another embryo that was imaged using 0.81- μ m z steps. In both recordings, complete two-color volumetric images were acquired at 3-min intervals. The *bicoid nanos torso-like* (*bnt*) mutant embryo shown in Figures 7 and S7, Movie S4, and part 2 of Movie S3 was imaged for 531 time points using 2.03- μ m z steps and 30-s time intervals. The wild-type embryos shown in Figures 7, 8, and S7, Movies S4, S5, and part 1 of Movie S3 were imaged for 1,400–4,000 time points using

1.62- μ m z steps. Single-color (membrane-only) volumetric images were acquired at 15-s intervals for Movie S4 and part 1 of Movie S3, and two-color volumetric images were acquired at 20-s intervals for Movie S5.

Sample Preparation and Imaging of Zebrafish Embryos Using Confocal Microscopy

Confocal imaging of zebrafish embryos was performed with embryos collected from an in-cross of the same transgenic line as above. The embryos were dechorionated at 6 hpf and embedded on a coverslip-bottom dish with 1% low-melting-point type VII agarose (Sigma) in filtered system water, with the animal pole facing the coverslip/objective. Imaging was performed with a Zeiss 510 inverted laser-scanning confocal microscope using a Zeiss 20 \times /0.8 NA air objective. H2B-eGFP was excited at 488 nm and lyn-tdTOMO at 561 nm. 152 images (1,024 \times 1,024 pixels each) were acquired with 0.90- μ m z steps and at a lateral pixel size of 0.62 μ m. The z stack encompassed approximately half of the embryo, disregarding the yolk cell.

Sample Preparation and Imaging of Mouse Embryos Using Confocal Microscopy

Confocal imaging of mouse embryos was performed with CAG-TAG1-positive embryos collected as described above. Embryos were dissected at E7.5, fixed for 10 min with 4% PFA, washed with cold PBS, and imaged in a coverglass-bottom dish in PBS on a Zeiss 510 inverted laser-scanning confocal microscope using a Zeiss 20 \times /0.8 NA air objective. The nuclear marker was excited at 488 nm and the membrane marker at 561 nm. An image stack of 148 planes (1,024 \times 1,024 pixels each) was acquired with a step size of 0.79 μ m and at a lateral pixel size of 0.52 μ m. The stack encompassed approximately half of the embryo.

Sample Preparation and Imaging of Fruit Fly Embryos Using Confocal Microscopy

Confocal imaging of *Drosophila* embryos was performed with embryos homozygous for spider-GFP and His2Av-RFP, which were collected and prepared as described above. Dechorionated embryos were embedded in a thin layer of low-melting-temperature agarose in a deep-well microscope slide. The distance between the slide and the coverslip was 225 μ m, which placed the surface of the embryo in contact with the coverglass. The embryo was imaged with a Zeiss 710 laser-scanning microscope using a Zeiss 20 \times /0.8 NA air objective. Membrane-GFP was excited at 488 nm and nuclear-RFP at 594 nm. Sixty images (1,024 \times 1,024 pixels each) were acquired with 0.81- μ m z steps and at a lateral pixel size of 0.38 μ m, using a zoom of 0.8. This produced a z stack encompassing approximately 21% of the embryo.

Algorithmic Design of the RACE Segmentation Framework

To facilitate rapid and accurate cell-shape segmentation in terabyte-scale 3D + time image data, we developed the RACE segmentation framework. The RACE algorithm can be divided into three main parts: (1) slice-based extraction of high-quality 2D segments of cells, (2) detection of seeds located inside the cells, and (3) seed-based fusion of 2D segments to final 3D cell shapes.

In the first step, each 3D image stack containing image data of fluorescently labeled cell membranes (Figures S2A and S2B) is filtered slice by slice with a 5 \times 5 2D median filter to reduce Poisson noise in the images. To enhance the membrane signal and further reduce undesired noise in the images, locally plane-like structures are enhanced using a so-called objectness filter (Antiga, 2007; Frangi et al., 1998). Based on specific properties of the ordered eigenvalues of the Hessian matrix at each pixel location, the filter can be tuned to obtain high responses on plane-like structures, such as cell membranes, and to efficiently suppress undesired structures as well as background noise (Figures 1A and S2C).

To further improve the membrane signal and close remaining gaps caused by inhomogeneous marker expression or other signal disruptions, we used a viscous watershed transform (Fernandez et al., 2010; Vachier and Meyer, 2005). This approach simulates the flooding of the grayscale images with a viscous fluid and is accomplished by using a morphological closing operation with successively increasing radii (Figures 1A and S2D). The maximum radius is determined by the smallest structures that still need to be individually resolved.

The segmentation of the enhanced image stacks is subsequently performed in a slice-by-slice manner, by processing each of the slices of the stack independently and in parallel using a morphological watershed segmentation algorithm (Figures 1A, S2E, and S2F).

In the second step, a set of initial seed points marking the inside of potential candidate cells is identified. These seed points serve the purpose of facilitating subsequent merging of 2D segments to 3D cell shapes. We provide multiple options for seed detection. First, seeds can be directly obtained from the enhanced membrane signal (Figure 2A) by generating a pseudo-nuclear channel from the binarized inverted membrane signal image (Figure 2B). Second, additional image data of fluorescently labeled cell nuclei (if available, Figure 2F) can be used to detect seed points based on a binarized version of the Laplacian-of-Gaussian filtered image data (Figure 2G). Irrespective of the choice of source data, H-maxima of the Euclidean distance map (Maurer et al., 2003) of the binary images are then extracted to separate touching binary regions (Figures 2C–2E, 2H–2J). Each of the extracted regions is assigned a unique integer label that is used to initialize the fusion of 2D cell segments generated by the slice-based watershed transform. Instead of taking advantage of only the centroids of the detected seeds (Stegmaier et al., 2014), we use the entire 3D seed morphology such that multiple slices can be initialized at once with a single seed, thereby accelerating segment fusion. For compatibility with existing software tools, the seed points used for segment fusion can also be imported from other sources using a comma-separated values-based (CSV) table format. For instance, we provide interfaces to the TGMM tracking framework (Amat et al., 2014) and CATMAID databases (Saalfeld et al., 2009) containing manually annotated or corrected seed points.

In the third and final step, we use discrete combinatorial optimization techniques to combine high-quality 2D segments to actual 3D cell shapes (Figures 1A and S3). Using the identified seed points as initial cell markers, 2D segments that intersect with a seed point are assigned the same seed label. Based on the Jaccard index of touching 2D segments in neighboring slices, we then calculate a seeded minimum spanning tree, in which 3D segments are iteratively assembled from intersecting 2D regions. This fusion is performed by starting with the largest regional overlap, i.e. the most similar regions are fused first, and fusion is stopped as soon as the Jaccard index of intersecting segments drops below an intersection threshold. To further improve segmentation results, we implemented two post-processing heuristics that exploit both intersection- and volume-based criteria to mitigate oversegmentation errors that may be caused, for instance, by redundant seed points (part 1 of Supplemental Experimental Procedures, Figures S3C–S3E).

In total only three intensity-dependent parameters need to be adjusted throughout the pipeline: the binarization threshold for the seed detection stage, the height of H-maxima extracted from the Euclidean distance map, and the starting level of the slice-based watershed segmentation algorithm (Figure S1 and Table S2).

We systematically identified all major performance bottlenecks of the RACE prototype we initially developed using ITK libraries, and replaced all performance-critical components with CPU-optimized or GPU-accelerated custom code. In particular, we wrote CUDA-based GPU implementations of the 2D median filter, the eigenvalue-based objectness filter, and the morphological closing operation (Figures S4C and S4D and Table S1). We furthermore developed custom high-performance CPU-optimized implementations of the watershed transform and the H-maxima filter, which are based on OpenCV and form part of the C++ library Nscale (Teodoro et al., 2013).

An explanation of acronyms used to refer to the different versions of the RACE algorithm is provided in tab 7 of Table S3. Further information, a more detailed description of all processing steps, guidance for parameter adjustments, implementation details, and general instructions on how to use the RACE segmentation framework are provided in Supplemental Experimental Procedures.

Segmentation Performance Evaluation

To quantitatively assess RACE segmentation quality, we manually segmented representative image regions in SiMView light-sheet microscopy recordings of *Drosophila* and mouse embryos as well as image data from a *Drosophila* embryo recorded with commercial confocal micro-

scopy. Ground truth annotations of 2D slices were performed with the Fiji plugin TrackEM2 (Cardona et al., 2012) and combined to complete 3D cell shapes using a custom graphical user interface in MATLAB. We compared RACE with three state-of-the-art cell-shape-segmentation algorithms (Fernandez et al., 2010; Khan et al., 2014; Mosaliganti et al., 2012) using the segmentation quality measures proposed by Coelho et al. (2009). Furthermore, topological errors such as added, missing, split, and merged segments were used to calculate precision, recall, and F-scores (Figure S4A and S4B, tabs 1–3 of Table S3, and part 2 of Supplemental Experimental Procedures). As high data throughput is crucial when analyzing terabyte-scale datasets, we also performed a detailed assessment of computation time for all algorithms. Performance was measured both in seconds, to reflect the actual computation time needed in practice, and in voxels per second, to ensure a fair comparison that also considers image size in those cases where images needed to be re-scaled to obtain isotropic sampling (Figures 1B, S4C, and S4D, and part 2 of Supplemental Experimental Procedures).

Based on the segmentation data provided by RACE, we extracted a variety of cell-shape features to quantitatively characterize cell morphology and cell-shape changes. These features include cell volume (number of voxels), centroid location, eccentricity (defining an ellipse with the same second moments as the extracted cell segment), extent (ratio of segment area and bounding box area), perimeter (number of boundary pixels in a reference plane), solidity (fraction of pixels that are part of the convex hull), and cell-shape anisotropy (measure of relative cell size along two orthogonal directions). We compared the features obtained from automatic reconstructions to ground truth annotations for those cells that exhibited at least 50% overlap across automatic and manual annotations. Feature values of matching cell pairs were then used to calculate relative deviations at the single-cell level as well as the level of mean values obtained by averaging over groups of cells (Figures 5, 7C, and S5, and tabs 4–6 of Table S3).

A detailed description of the various algorithms and segmentation quality measures as well as the in-depth side-by-side performance comparison itself are provided in parts 1 and 2 of Supplemental Experimental Procedures, Figures 4, 5, S1B–S1E, and S4–S7, and Tables S1, S2, and S3.

Joint Reconstruction of Cell Morphology and Cell Lineages

To investigate cell-shape changes over time at the single-cell level, we combined the RACE cell segmentation framework with our TGMM algorithm for cell tracking (Amat et al., 2014). In the first step of this combined computational framework, cell tracking is performed using image data of labeled cell nuclei. For this purpose, TGMM performs a low-level image oversegmentation using a watershed transform and hierarchical agglomeration with persistence-based clustering (PBC) that partitions the image into super-voxels. These super-voxels are subsequently combined to form complete cell nuclei tracks using parametric contour evolution with a sequential Gaussian mixture model. A post-processing step using machine-learning classifiers improves cell tracks in difficult cases, for instance to correctly identify and handle cell division events. As previously described, the TGMM framework has two primary open parameters (Amat et al., 2014), a background intensity threshold and a threshold for PBC agglomeration, which were set to 50 and 12, respectively.

To incorporate the cell tracking information obtained with TGMM into the RACE cell-shape segmentation framework, we skipped the seed detection step offered by the RACE algorithm and directly used cell centroids and cell identities identified by the TGMM algorithm for RACE seeding. This approach thus maintains cell tracks and cell lineages reconstructed by TGMM, and utilizes RACE to annotate these cell tracks with cell-shape information. For data visualization (Figure 8 and Movie S5), all cells were assigned a color code that identifies anteroposterior cell position at the first time point of the time-lapse recording. This color code was subsequently propagated forward in time using the automatically reconstructed cell tracks and cell divisions.

Visualizations of individual cells by high-resolution surface rendering were automatically created by extracting iso-surfaces from the segmentation masks generated by RACE, using custom code written in MATLAB. Each triangular mesh was smoothed with a $7 \times 7 \times 3$ kernel to avoid surface edges arising from the anisotropic spatial resolution of the microscope.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, three tables, five movies, and one software file and can be found with this article online at <http://dx.doi.org/10.1016/j.devcel.2015.12.028>.

AUTHOR CONTRIBUTIONS

P.J.K and J.S. conceived of the research with input from F.A. and R.M. J.S., F.A., R.M., and P.J.K. developed the segmentation framework. J.S., F.A., and G.T. implemented the software. W.C.L., Y.W., and K.M. performed the imaging experiments. J.S. and P.J.K. wrote the manuscript with input from all authors. P.J.K. supervised the project.

ACKNOWLEDGMENTS

This work was supported by the Howard Hughes Medical Institute (J.S., F.A., W.L., K.M., Y.W., P.J.K.), the Helmholtz Association in the Program Bio-Interfaces (J.S., R.M.), the Karlsruhe House of Young Scientists (J.S.), and the Brazilian National Council for Scientific and Technological Development (CNPq) (G.T.).

Received: June 8, 2015

Revised: November 28, 2015

Accepted: December 23, 2015

Published: January 25, 2016

REFERENCES

- Amat, F., Lemon, W., Mossing, D.P., McDole, K., Wan, Y., Branson, K., Myers, E.W., and Keller, P.J. (2014). Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. *Nat. Methods* 11, 951–958.
- Amat, F., Hockendorf, B., Wan, Y., Lemon, W.C., McDole, K., and Keller, P.J. (2015). Efficient processing and analysis of large-scale light-sheet microscopy data. *Nat. Protoc.* 10, 1679–1696.
- Antiga, L. (2007). Generalizing vesselness with respect to dimensionality and shape. *Insight J.* 3. <http://hdl.handle.net/1926/576>.
- Blankenship, J.T., and Wieschaus, E. (2001). Two new roles for the *Drosophila* AP patterning system in early morphogenesis. *Development* 128, 5129–5138.
- Bosveld, F., Bonnet, I., Guirao, B., Tili, S., Wang, Z., Petitalot, A., Marchand, R., Bardet, P.L., Marcq, P., Graner, F., et al. (2012). Mechanical control of morphogenesis by Fat/Dachsous/Four-jointed planar cell polarity pathway. *Science* 336, 724–727.
- Cardona, A., Saalfeld, S., Schindelin, J., Arganda-Carreras, I., Preibisch, S., Longair, M., Tomancak, P., Hartenstein, V., and Douglas, R.J. (2012). TrakEM2 software for neural circuit reconstruction. *PLoS One* 7, e38011.
- Chhetri, R.K., Amat, F., Wan, Y., Hockendorf, B., Lemon, W.C., and Keller, P.J. (2015). Whole-animal functional and developmental imaging with isotropic spatial resolution. *Nat. Methods* 12, 1171–1178.
- Coelho, L.P., Shariff, A., and Murphy, R.F. (2009). Nuclear segmentation in microscope cell images: a hand-segmented dataset and comparison of algorithms. In *IEEE International Symposium on Biomedical Imaging: from Nano to Macro (ISBI) (IEEE)*, 518–521.
- 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 lineage tracing at cell resolution. *Nat. Methods* 7, 547–553.
- Frangi, A.F., Niessen, W.J., Vincken, K.L., and Viergever, M.A. (1998). Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Interventions (MICCAI)* (Springer), 130–137.
- Funke, J., Andres, B., Hamprecht, F.A., Cardona, A., and Cook, M. (2012). Efficient automatic 3D-reconstruction of branching neurons from EM data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE)*, 1004–1011.
- He, B., Doubrovinski, K., Polyakov, O., and Wieschaus, E. (2014). Apical constriction drives tissue-scale hydrodynamic flow to mediate cell elongation. *Nature* 508, 392–396.
- Keller, P.J. (2013). Imaging morphogenesis: technological advances and biological insights. *Science* 340, 1234168.
- Khan, Z., Wang, Y.C., Wieschaus, E.F., and Kaschube, M. (2014). Quantitative 4D analyses of epithelial folding during *Drosophila* gastrulation. *Development* 141, 2895–2900.
- Lecuit, T., and Le Goff, L. (2007). Orchestrating size and shape during morphogenesis. *Nature* 450, 189–192.
- Liu, T., Jones, C., Seyedhosseini, M., and Tasdizen, T. (2014). A modular hierarchical approach to 3D electron microscopy image segmentation. *J. Neurosci. Methods* 226, 88–102.
- Maurer, C.R., Jr., Qi, R., and Raghavan, V. (2003). A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.* 25, 265–270.
- Mosaliganti, K.R., Noche, R.R., Xiong, F., Swinburne, I.A., and Megason, S.G. (2012). ACME: automated cell morphology extractor for comprehensive reconstruction of cell membranes. *PLoS Comp. Biol.* 8, e1002780.
- Nüsslein-Volhard, C., Frohnhofer, H.G., and Lehmann, R. (1987). Determination of anteroposterior polarity in *Drosophila*. *Science* 238, 1675–1681.
- Oates, A.C., Gorinkiel, N., Gonzalez-Gaitan, M., and Heisenberg, C.P. (2009). Quantitative approaches in developmental biology. *Nat. Rev. Genet.* 10, 517–530.
- Osterfield, M., Du, X., Schupbach, T., Wieschaus, E., and Shvartsman, S.Y. (2013). Three-dimensional epithelial morphogenesis in the developing *Drosophila* egg. *Dev. Cell* 24, 400–410.
- Pantazis, P., and Supatto, W. (2014). Advances in whole-embryo imaging: a quantitative transition is underway. *Nat. Rev. Mol. Cell Biol.* 15, 327–339.
- Pare, A.C., Vichas, A., Fincher, C.T., Mirman, Z., Farrell, D.L., Mainieri, A., and Zallen, J.A. (2014). A positional Toll receptor code directs convergent extension in *Drosophila*. *Nature* 515, 523–527.
- Saalfeld, S., Cardona, A., Hartenstein, V., and Tomančák, P. (2009). CATMAID: collaborative annotation toolkit for massive amounts of image data. *Bioinformatics* 25, 1984–1986.
- Stegmaier, J., Otte, J.C., Kobitski, A., Bartschat, A., Garcia, A., Nienhaus, G.U., Strähle, U., and Mikut, R. (2014). Fast segmentation of stained nuclei in terabyte-scale, time resolved 3D microscopy image stacks. *PLoS One* 9, e90036.
- Sweeton, D., Parks, S., Costa, M., and Wieschaus, E. (1991). Gastrulation in *Drosophila*: the formation of the ventral furrow and posterior midgut invaginations. *Development* 112, 775–789.
- Tamada, M., Farrell, D.L., and Zallen, J.A. (2012). Abl regulates planar polarized junctional dynamics through beta-catenin tyrosine phosphorylation. *Dev. Cell* 22, 309–319.
- Teodoro, G., Pan, T., Kurc, T.M., Kong, J., Cooper, L.A., Podhorszki, N., Klasky, S., and Saltz, J.H. (2013). High-throughput analysis of large microscopy image datasets on CPU-GPU cluster platforms. *IPDPS 2013*, 103–114.
- Tomer, R., Khairy, K., Amat, F., and Keller, P.J. (2012). Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nat. Methods* 9, 755–763.
- Trichas, G., Begbie, J., and Srinivas, S. (2008). Use of the viral 2A peptide for bicistronic expression in transgenic mice. *BMC Biol.* 6, 40.
- Vachier, C., and Meyer, F. (2005). The viscous watershed transform. *J. Math. Imaging Vis.* 22, 251–267.
- Wang, Y.C., Khan, Z., Kaschube, M., and Wieschaus, E.F. (2012). Differential positioning of adherens junctions is associated with initiation of epithelial folding. *Nature* 484, 390–393.