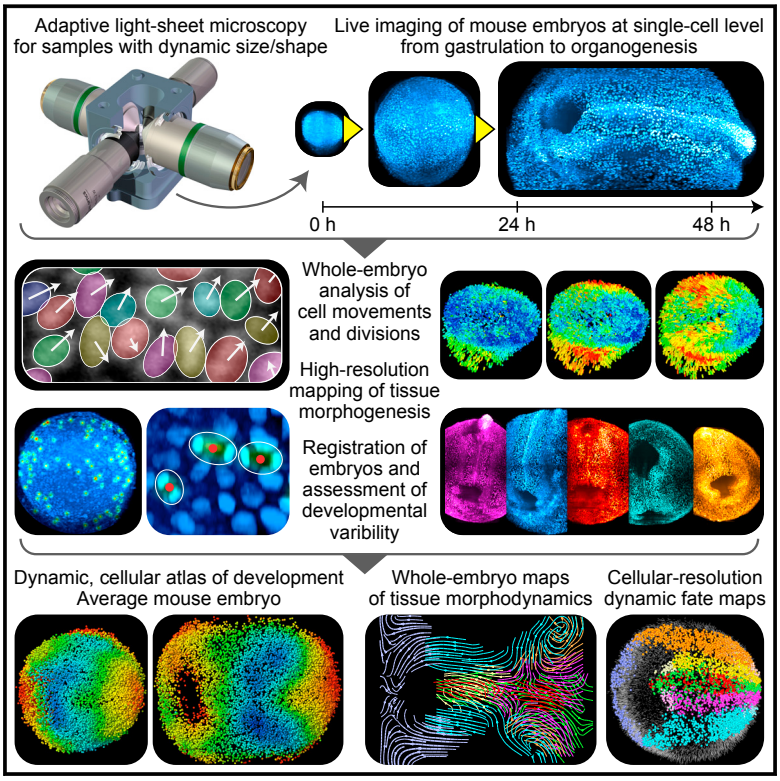
Resource

Cell

In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level

Graphical Abstract



Highlights

- Adaptive light-sheet microscopy captures mouse development at the single-cell level
- We analyzed embryo-wide cell dynamics from gastrulation to early organogenesis
- We reconstructed high-resolution fate maps and maps of tissue morphogenesis
- We created a statistical, dynamic atlas of development from multiple embryos

Authors

Katie McDole, Léo Guignard, Fernando Amat, ..., Srinivas C. Turaga, Kristin Branson, Philipp J. Keller

Correspondence

mcdolek@janelia.hhmi.org (K.M.), guignardl@janelia.hhmi.org (L.G.), kellerp@janelia.hhmi.org (P.J.K.)

In Brief

Adaptive light-sheet microscopy is used to establish a dynamic atlas of post-implantation mouse development at the single-cell level.

McDole et al., 2018, Cell *175*, 1–18
October 18, 2018
https://doi.org/10.1016/j.cell.2018.09.031

CellPress

Please cite this article in press as: McDole et al., *In Toto* Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level, Cell (2018), https://doi.org/10.1016/j.cell.2018.09.031

Resource

Cell

In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level

Katie McDole,[1,*] Léo Guignard,[1,*] Fernando Amat,[1] Andrew Berger,[1] Grégoire Malandain,[2] Loïc A. Royer,[3] Srinivas C. Turaga,[1] Kristin Branson,[1] and Philipp J. Keller[1,4,*]

[1]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

[2]Université Côte d'Azur, Inria, CNRS, I3S, 06900 Sophia Antipolis, France

[3]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[4]Lead Contact

*Correspondence: mcdolek@janelia.hhmi.org (K.M.), guignardl@janelia.hhmi.org (L.G.), kellerp@janelia.hhmi.org (P.J.K.)

https://doi.org/10.1016/j.cell.2018.09.031

SUMMARY

The mouse embryo has long been central to the study of mammalian development; however, elucidating the cell behaviors governing gastrulation and the formation of tissues and organs remains a fundamental challenge. A major obstacle is the lack of live imaging and image analysis technologies capable of systematically following cellular dynamics across the developing embryo. We developed a light-sheet microscope that adapts itself to the dramatic changes in size, shape, and optical properties of the post-implantation mouse embryo and captures its development from gastrulation to early organogenesis at the cellular level. We furthermore developed a computational framework for reconstructing long-term cell tracks, cell divisions, dynamic fate maps, and maps of tissue morphogenesis across the entire embryo. By jointly analyzing cellular dynamics in multiple embryos registered in space and time, we built a dynamic atlas of post-implantation mouse development that, together with our microscopy and computational methods, is provided as a resource.

INTRODUCTION

Mammalian development has been an area of intense study for many decades, but for all our efforts to elucidate the developmental mechanisms that enable the transition from a single cell to a fully formed organism, we have only limited knowledge of the dynamic processes that shape the embryo. The study of mouse embryogenesis, though an ideal candidate to investigate early mammalian development due to its size, relative accessibility, and genetic tractability, remains largely limited to snapshots in time. In particular, the period from gastrulation to organogenesis (E6.5 to E8.5 d.p.c., Figures S1A–S1C), when cells from the pluripotent epiblast migrate out of the primitive streak and begin to differentiate into various tissues and organ systems, is an area of keen interest for stem cell biology and tissue engineering. However, owing to the technical challenges associated with the visualization and quantification of dynamic cell behaviors *in vivo*, it is still largely unknown how cells migrate, how they organize during the formation of tissues, what dynamic behaviors they exhibit, how cells interact with their neighbors, and how stereotyped cell behaviors are across individuals (Herion et al., 2014; Lawson and Pedersen, 1992; Rivera-Pérez and Hadjantonakis, 2014; Tam and Beddington, 1992). The ability to image and computationally reconstruct whole-embryo development at the cellular level would enable the detailed analysis of these morphodynamic events and reveal developmental processes never before visualized *in vivo*.

Standing in the way of this goal, however, is that mouse embryos and their culture present a unique set of challenges. First, the rapid growth, complex optical properties, and often dense cell structures of the embryo, which change during development, present fundamental obstacles for high-resolution imaging. Second, the embryo requires optically scattering, autofluorescent serum to develop properly. Third, due to its rapid expansion in volume, the embryo cannot be mechanically constrained. Finally, mouse embryos are very photosensitive, limiting the amount of light they can tolerate without compromising development. Widely used confocal microscopes, for example, are limited to observing sub-regions of the embryo, with large steps in time and space, and cannot sustain normal development beyond 24 hr due to phototoxicity. To address even the most fundamental questions in post-implantation mouse development will require not only customized solutions, but entirely new imaging strategies.

Recent technological advances, such as the introduction of light-sheet microscopy for developmental imaging (Huisken et al., 2004; Keller et al., 2008), have provided a major opportunity for advancing our understanding in the field. The unique benefits of light-sheet microscopy make it ideal for studying sensitive, developing organisms. A sheet of laser light illuminates the embryo selectively along the detection focal plane an entire section at a time. This scheme greatly reduces the amount of light the embryo is exposed to and allows for significantly higher temporal resolution without compromising viability. Early efforts toward adapting light-sheet microscopy to mouse embryogenesis enabled the imaging of early developmental stages

(Ichikawa et al., 2014; Strnad et al., 2016) and sub-regions of larger, sparsely labeled embryos (Udan et al., 2014). However, methods capable of imaging post-implantation development beyond a short period of time, or even entire embryos at the single-cell level, are still lacking. Here, we developed a light-sheet microscope for long-term imaging of entire mouse embryos from gastrulation to early organogenesis at the high spatial and temporal resolution required to systematically follow single-cell behavior. We designed an adaptive imaging approach, custom optics, and an integrated culturing system to build a multi-view light-sheet microscope that tracks the changing shape, size, and optical properties of the mouse embryo as it grows over 250-fold in volume.

Overcoming these challenges in live imaging is, however, only the first step needed for reconstructing a cellular-resolution, dynamic atlas of development. A single 48 hr light-sheet recording of gastrulation and early organogenesis captures tens of thousands of cells with diverse shapes and dynamic properties over hundreds of time points, as they form an embryo 10–70 times larger than zebrafish or *Drosophila* embryos at comparable stages. To extract biologically meaningful information from images of such complexity and size, we developed a computational framework for the following: (1) automated long-term cell tracking over the full 48 hr period with an average precision of two cell diameters; (2) robust detection of cell divisions across the embryo; (3) construction of high-resolution dynamic fate maps and spatiotemporal maps of tissue morphodynamics; (4) spatiotemporal registration of multiple embryos to enable statistical quantifications of cellular dynamics and variability in developmental processes across individuals; and (5) the construction of an average mouse embryo. We used these methods to reconstruct the development of entire mouse embryos across scales, from single-cell dynamics to whole-embryo morphogenesis, and jointly analyze these data to build a statistical, dynamic atlas of gastrulation and early organogenesis in the mouse embryo. We applied these resources to the analysis of specific cell and tissue behaviors, including migration of primordial germ cells, embryo-wide spatiotemporal patterns and geometric properties of cell divisions, and cellular dynamics during elongation and folding of the neural tube.

While this study focuses on *in toto* imaging, dynamic analysis, and atlas construction for the stages of mouse development covering gastrulation to early organogenesis, we designed our imaging and computational tools to be broadly applicable across organisms and developmental systems, including organoids, stem cell systems, and tissue explants. Moreover, the tools and resources presented here should not only enable the analysis of broader developmental timescales, but also support a wide spectrum of applications—from mutant phenotyping, real-time analysis of cell signaling, morphodynamics, and force measurements to atlas-guided transcriptional analyses, optical manipulations, and perturbations.

RESULTS

Adaptive Multi-view Light-Sheet Microscope for Imaging Mouse Development

We developed a light-sheet microscope capable of culturing and imaging mouse embryos from gastrulation to early organogenesis at high spatiotemporal resolution (Figures 1A–1C, S1D–S1H; Data S1A). The microscope adapts itself to the rapid growth and complex optical properties of the embryo while maintaining viability over days of continuous imaging. The instrument does the following: (1) optimizing spatial resolution by automatically measuring and compensating for the continuously changing optical properties and geometry of the embryo; (2) minimizing optical path lengths in light-scattering serum through the use of custom-designed objectives; (3) continuously tracking the three-dimensional (3D) movements and growth of the embryo; and (4) supporting proper embryonic development, with minimal phototoxicity and a mounting strategy that does not interfere with embryonic growth.

Conceptually, our microscope design advances the principles introduced in simultaneous multi-view (SiMView) light-sheet microscopy (Tomer et al., 2012) and the AutoPilot light-sheet microscopy framework (Royer et al., 2016), which enable the imaging of externally developing organisms such as *Drosophila* and zebrafish. In our present work, we introduced a wide spectrum of improvements required to sustain and image normal mouse embryonic development at high resolution over a period of days (STAR Methods; Figure 1; Videos S1 and S2). While necessary for proper embryo development, the loose mounting strategy required by the mouse embryo complicates time-lapse live imaging. Not only does the embryo position drift slowly over time, but the embryo expands in volume more than 250-fold over 48 hr and dramatically changes in shape and composition. If the system were to be left in its initial configuration, image quality and spatial resolution would not only degrade over time, but the embryo may drift out of the field of view (Video S1B; Table S1). We thus developed an adaptive imaging framework suitable for dynamic specimens that simultaneously compensates for drift, growth, and changing optical properties (Figures 1D and 1E; STAR Methods). This method improves over our earlier AutoPilot system (Royer et al., 2016), which was limited to imaging developing specimens with approximately constant size and shape. We found that the AutoPilot corrects for only 37% of the aberration-induced defocus error in developing mouse embryos and suffers from an average defocus-induced mismatch between light sheets and detection focal planes of 1.90 ± 0.81 μm (n = 680,000 focus measurements across 20 time-lapse experiments) (Table S1; Figure S2). A fully uncorrected microscope produces an even higher average defocus error of 2.99 ± 1.34 μm (same statistics). Existing imaging methods therefore yield aberrated images that lack cellular resolution in more than 50% of the embryo (Figure 1F). In addition, existing light-sheet microscopes lack the ability to track the embryo and expand the imaging volume to match its growth, leading to a loss of 64.4% ± 11.9% of the embryo's volume over 24 hr when using mechanically unconstrained sample mounting (n = 20 time-lapse experiments) (Table S1).

We developed algorithms for image-based tracking of embryo position and mapping of embryo size (Figures 1D and 1E), which achieve the following: (1) ensure that no part of the embryo is lost during imaging, and (2) facilitate building a dynamic geometrical model of the embryo that automatically defines reference locations for aberration measurements throughout the embryo. These algorithms perform real-time processing of the time-lapse

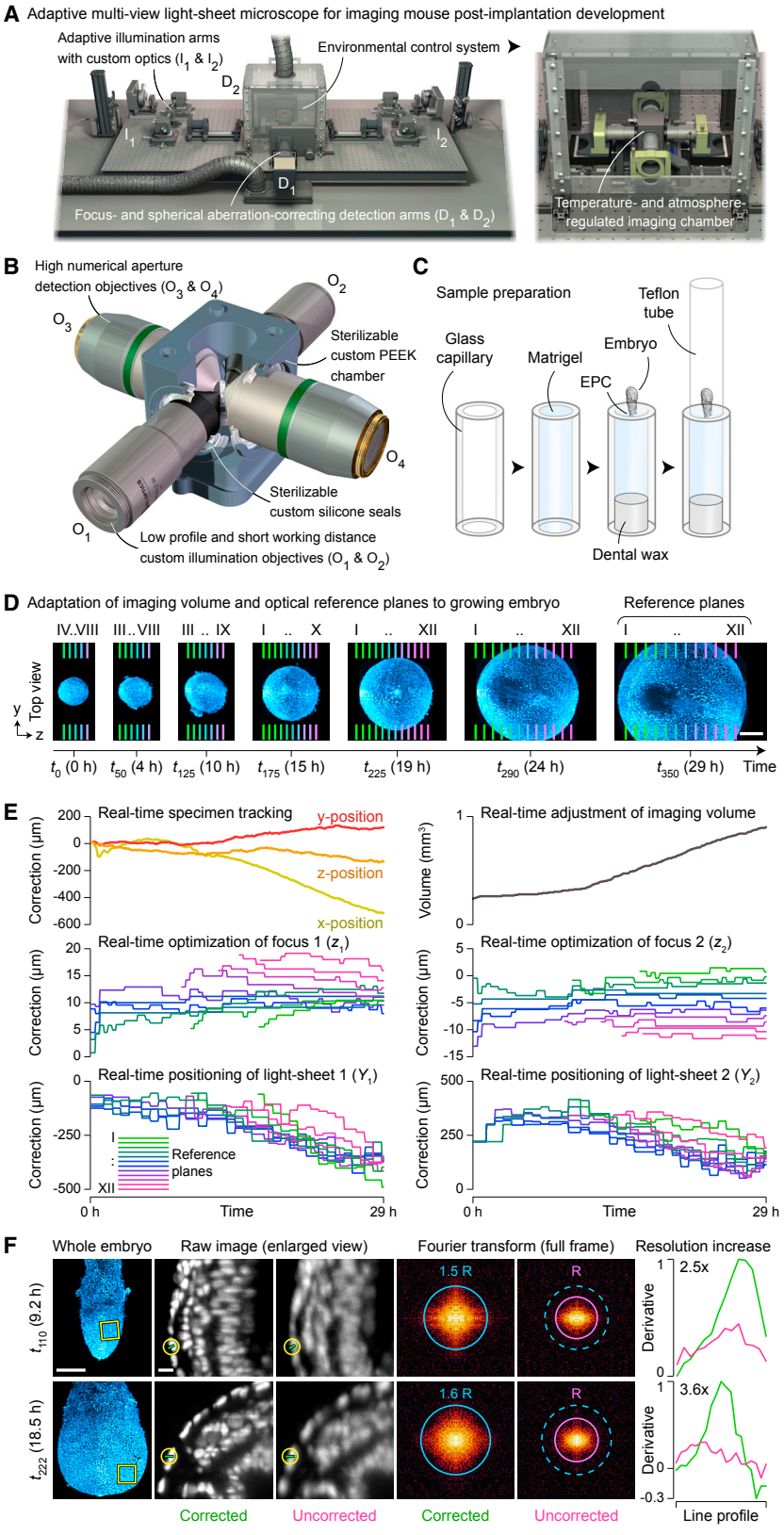


Figure 1. Light-Sheet Microscope for Adaptive Imaging of Mouse Embryo Development

(A) Adaptive multi-view light-sheet microscope for live imaging of mouse embryos, including Maus Haus environmental control system.

(B) Sample chamber and objectives. Custom illumination objectives (O_1, O_2) provide short working distance and spatial compatibility with high numerical aperture detection objectives (O_3, O_4).

(C) Mounting of post-implantation embryos. A glass capillary with Matrigel is sealed on one end with dental wax. The embryo's EPC is embedded in the opposite end. This assembly is placed either directly into the chamber for open culture or surrounded by a 25-µm-thin FEP cylinder.

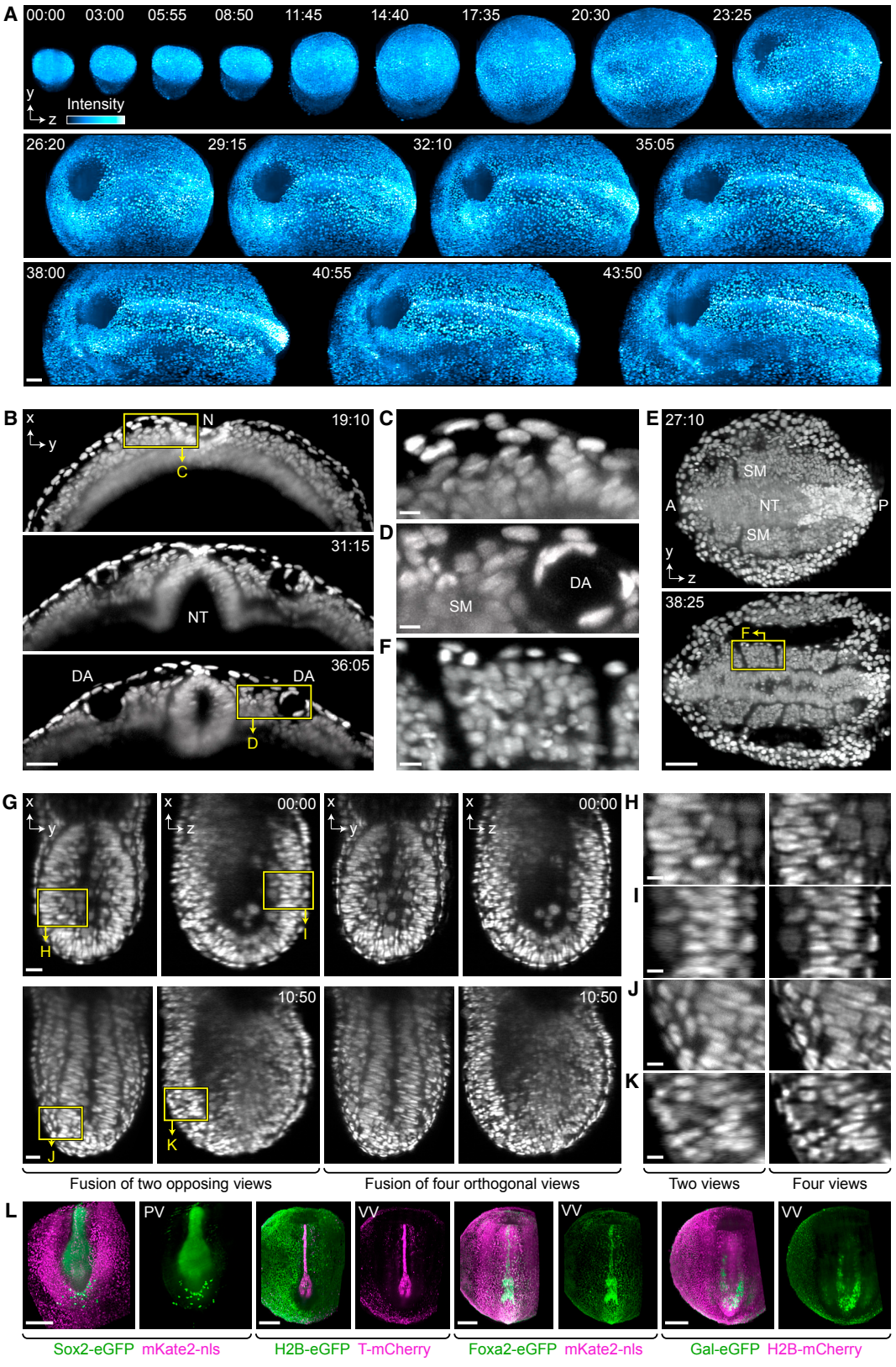
(D) Automated expansion of imaging volume and positioning of reference planes for adaptive imaging with optimal image quality during embryo development and growth.

(E) Adaptive correction of embryo position and imaging volume, focus optimization for light sheets and detection objectives (z_1, z_2), and optimization of light-sheet waist positions (Y_1, Y_2) for all dynamic reference planes in one imaging experiment.

(F) Left: comparison of image quality with ("Corrected") or without ("Uncorrected") adaptive corrections. Right: Fourier transforms and quantifications of resolution improvements (for regions marked by green lines, as described in Royer et al. [2016]).

Scale bars (µm): 200 (D; F, embryo), 15 (F, enlarged view).

See also Video S1 and Figures S1 and S2.



(legend on next page)

volumetric imaging data acquired by the microscope, stabilize the 3D position of the embryo in the center of the field of view, and continuously adjust the size of the imaging volume to match the growth rate of the embryo. Using the dynamic geometrical sample model, the adaptive imaging framework then divides the specimen volume into a set of reference regions, for which it determines the locally optimal light-sheet geometry and geometrically matches detection focal planes to illumination planes (Figure 1E; Video S1B). Thereby, the average defocus-induced mismatch between light sheets and detection focal planes is reduced to 0.06 ± 0.04 μm (Table S1), spatial resolution is increased 3.3-fold, and signal strength is increased 2.1-fold (Figure 1F; Video S1C). This continuous, automated adaptation also allows the microscope to run unattended for days and greatly reduces the data rate.

To minimize light exposure of the specimen, all measurements of the adaptive imaging framework are performed with low laser power and structured illumination. Enabling full system optimization with all degrees of freedom increases the number of images recorded per experiment by 5.2% and the energy load on the sample by 2.3% (n = 11 time-lapse experiments), which is negligible with respect to specimen viability. The total energy load in our microscope is so low that fluorescent protein production outperforms photobleaching in all experiments reported in this study, leading to a doubling of fluorescence levels every 10 hr on average.

Multi-day High-Resolution Imaging of Post-Implantation Mouse Development

These advancements in microscope design, algorithmic strategies for real-time adaptive imaging, and embryo culture allowed us to capture the development of the mouse embryo in high resolution from gastrulation to organogenesis *in toto* over the course of nearly 48 hr (from E6.5 d.p.c. to E8.5 d.p.c.; Video S2A; Figures 2A and S1A–S1C). We acquired opposing views of the embryo for up to two color channels in 5 min intervals, visualizing and tracking individual cells across the embryo at varying depths and observing morphodynamic processes and tissue formation as they occur (Figures 2B–2F). Notably, although we achieved single-cell resolution in endoderm and mesodermal tissues, the ectoderm presents a particular challenge due to its dense and highly scattering structure. More ventral ectodermal structures and those closest to the detection objectives could generally be resolved at the single-cell level; however, the dorsal-most and internal structures such as the surface ectoderm and medial head folds are not easily accessible (Figure 2B). Careful selection of fluorescent reporters proved essential to achieve cellular resolution in these optically complex areas and in deeper tissues such as the heart. Selective lineage labeling strategies can be used to generate a compelling overview of heart development (Video S3A) but lack the high resolution at later stages to clearly distinguish individual cells. While far-red reporters such as mKate2 provide a slight advantage in discerning individual cells and structures (Video S2B), bright near-infrared fluorescent reporter lines are needed to achieve greater depth and contrast. To demonstrate these advantages, we utilized a Histone2B-miRFP703 reporter line (Gu et al., 2018) to image the developing heart in greater cellular detail than has been previously possible (Video S3B). Using this line, we were able to image up to 600 μm deep into the anterior region of the embryo, encapsulating nearly the entire linear heart tube at single-cell resolution, as well as regions of the head-fold and foregut pocket that were previously inaccessible.

Spatial resolution, particularly in crowded, low-contrast regions such as the disorganized, rapidly mixing primitive streak (Figures S3A and S3B), can be further improved by orthogonal four-view imaging (Video S2C; Figures 2G–2K; STAR Methods). Alternatively, mosaic reporter strategies provide an excellent opportunity to track and visualize individual cell behaviors in a crowded and dynamic environment. The reconstruction of lineages had previously been accomplished by manual single-cell labeling (Lawson and Pedersen, 1987; Tam et al., 1997); however, this approach lacks the spatial and temporal resolution provided by live imaging and adds additional manipulations to the embryo. By using a mosaic Cre/loxP reporter strategy, we created detailed lineage trees for single cells as they migrate through the primitive streak (Figures S3C and S3D). Cells can be tracked starting from their positions in the epiblast and as they mix, divide, and migrate through the streak.

Although these recordings appear to represent "normal" and expected embryonic development over the stages imaged, the post-implantation embryo cannot be returned to the uterus to verify competency. We thus sought to ensure that imaging conditions did not perturb normal developmental processes by utilizing a range of existing fluorescent reporters known to spatially recapitulate endogenous gene expression in various tissues (Figures S3E–S3H and 2L; Video S3C and S3D). These experiments show that embryos not only develop normally with respect

Figure 2. Long-Term Live Imaging of Post-Implantation Mouse Development at the Cellular Level

(A) Selected projections from H2B-eGFP channel of a CAGTAG1 expressing mouse embryo over 44 h of continuous imaging, from early streak (E6.5) to somite stages (E8.5).

(B–D) (B) Projections of 45-μm-thick cross-section through mid-plane of embryo in (A), demonstrating ability to resolve individual cells across germ layers. Cross-section and zoomed-in segment (C) highlight node (N), mesoderm, and endoderm regions. (D) shows cross-section of angioblasts assembling dorsal aortae (DA).

(E and F) (E) Projections of 45-μm thick cross-sections of neural tube (NT) and somitic mesoderm, with surrounding regions in lateral plate and endoderm. (F) Zoomed-in region showing single somite.

(G) En face (XY) and lateral (XZ) slices of two-view (left) and four-view (right) image data of streak-stage embryo expressing mKate2-nls.

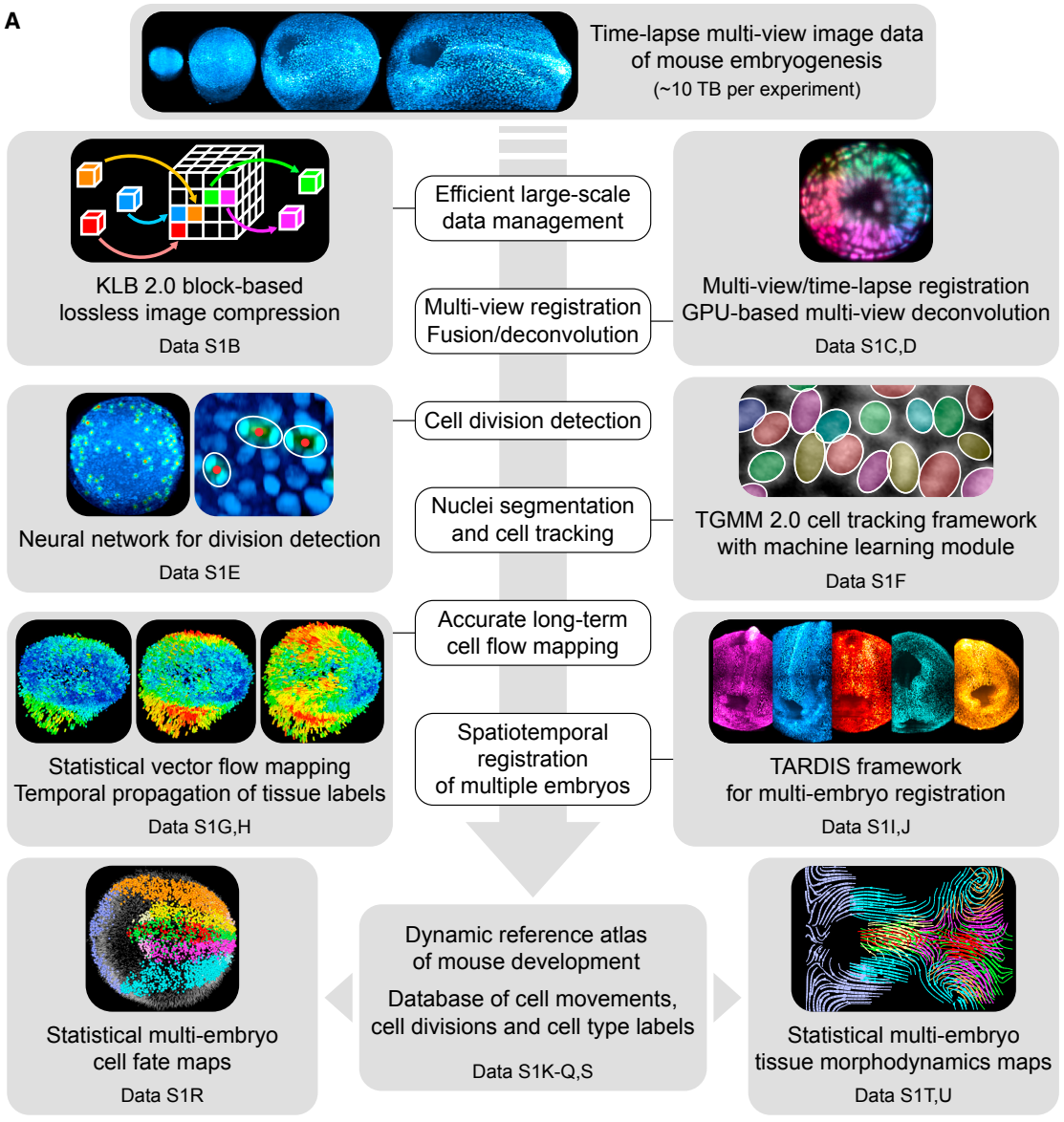
(H–K) Enlarged views of regions marked in (G), showing contrast and resolution improvement by four-view deconvolution (right) versus two-view content-based fusion (left).

(L) Live-reporter expression in developing embryos for Sox2-eGFP, T-mCherry, Foxa2-eGFP, Gal-eGFP, displaying expected expression and localization after long-term imaging.

Time (hh:mm). SM, somite. NT, neural tube. A, anterior. P, posterior. PV, posterior view. VV, ventral view.

Scale bars (μm): 200 (L), 100 (A, E), 50 (B), 25 (G), 15 (F), 10 (C, D, H–K).

See also Videos S2 and S3 and Figures S1 and S3.



(legend on next page)

to their morphology, but also properly express patterning genes over the course of development. The viability of these dual reporters greatly expands the avenues of study for early mouse development.

To conclude this section, we present an example of the versatility of this technique to image specific cell types or populations over long timescales. We used a Sox2-eGFP reporter line (Video S3C) to visualize the migration of primordial germ cells (PGCs) (Campolo et al., 2013). Primordial germ cells first arise at around E6.5 d.p.c. in the proximal posterior bordering the extra-embryonic and embryonic regions (Saitou and Yamaji, 2012) and later migrate to the genital ridge (Molyneaux et al., 2001). The use of the Sox2-eGFP reporter allowed us to clearly distinguish PGCs as they arose from a small cluster in the proximal posterior, migrated into the endoderm layer, and began to actively wander about, sending out small membrane protrusions (Video S3C) until they were drawn inward by the formation of the hindgut portal. During amnion inflation, a few of the more proximal PGCs were pulled along with the allantois toward the extra-embryonic region (ExE). These cells never rejoined their embryonic counterparts during imaging and are presumably the origin of PGCs found in the extra-embryonic region (Anderson et al., 2000; Ginsburg et al., 1990). This segregation of PGCs into the extra-embryonic region appears to be an unintended consequence of the rapidly expanding amnion, as embryos will exhibit multiple different behaviors regarding the more proximal PGCs in this region. In every embryo examined (n = 4), we observed two or three PGCs to act in one of four ways: (1) join the ExE; (2) be drawn along with the allantois proximally but then turn around and rejoin the embryonic PGCs; (3) one rejoins while the other remains in the ExE; or (4) yet another one rejoins, while the other undergoes apoptosis. It is unclear why, with such a small initial population size, the fate of these cells would seemingly be left to chance. The power of the adaptive imaging methodology presented here enabled us to visualize this population over long timescales, with a richness and detail that has previously been unattainable.

Automated Cell Tracking Across the Post-Implantation Mouse Embryo

Post-implantation mouse development presents a unique challenge not only for live imaging, but also for the analysis of the resulting image datasets. To systematically extract quantitative information on cellular dynamics and—ultimately—biological insights from the image data, we developed a modular computational framework (Figure 3A). Our tools facilitate efficient and accurate cell segmentation, long-term cell tracking, detection of cell divisions, reconstruction of high-resolution dynamic fate maps, high-resolution mapping of tissue morphodynamics, and registration and quantitative comparison of cellular dynamics across multiple embryos (STAR Methods; Videos S4, S5, S6, and S7; Data S1B–S1U). We provide a comprehensive guide that explains the practical use of this framework and includes example data (STAR Methods). In the following sections, we briefly describe each module and the respective biological analyses they enable, starting here with automated, whole-embryo cell tracking.

A 48 hr recording of gastrulation and early organogenesis in the mouse embryo captures the dynamic behavior of tens of thousands of cells over hundreds of time points. Conventional cell segmentation and tracking algorithms are not equipped to handle the size and complexity of these datasets, and manual annotation by a human would ostensibly take years for a single dataset. We thus developed an improved version of our Bayesian cell tracking framework TGMM (Tracking with Gaussian Mixture Models), which was originally designed for automated cell tracking in developing zebrafish and *Drosophila* embryos. We successfully tested TGMM on short-term recordings of early streak-stage mouse embryos comprising less than 1,000 cells (Amat et al., 2014) but found that it failed on our long-term recordings of the up to 200-fold larger embryo volumes and 20-fold higher cell counts encountered in the present study. We thus developed TGMM 2.0, which employs a machine learning approach to division detection utilizing both lineage-based and image-based features (STAR Methods; Data S1F).

Over the full 48 hr imaging period and across all tissues in the mouse embryo, TGMM 2.0 achieved an average linkage accuracy of 93.0%–94.8% and an average segmentation accuracy of 96.7%–97.6% (n = 8,982–30,962 cell position and linkage annotations), depending on the fluorescent marker strategy (Table S1). We note that while we designed TGMM 2.0 as a tool for cell tracking and recommend its use for this purpose, we discourage its use as a cell division detector (except in smaller model systems, such as zebrafish and *Drosophila*). Even with the improvements introduced in TGMM 2.0, cell division detection performance itself is still limited and cannot easily be improved within the scope of the current algorithmic design. We thus set out to develop a new computational tool that complements TGMM and offers a robust and automated way of identifying cell divisions in the raw image data of developing embryos.

Efficient Detection of Cell Divisions Using Deep Learning

Cell proliferation rates and timing are of critical importance to cell differentiation and tissue formation, but little is known about tissue- and embryo-wide division patterns or how they might evolve over time. Although TGMM 2.0 incorporates a classifier for identifying dividing cells and linking a parent cell with its daughters, this approach faces the same challenges as those

Figure 3. Data Processing Framework and Neural Network for Cell Division Detection

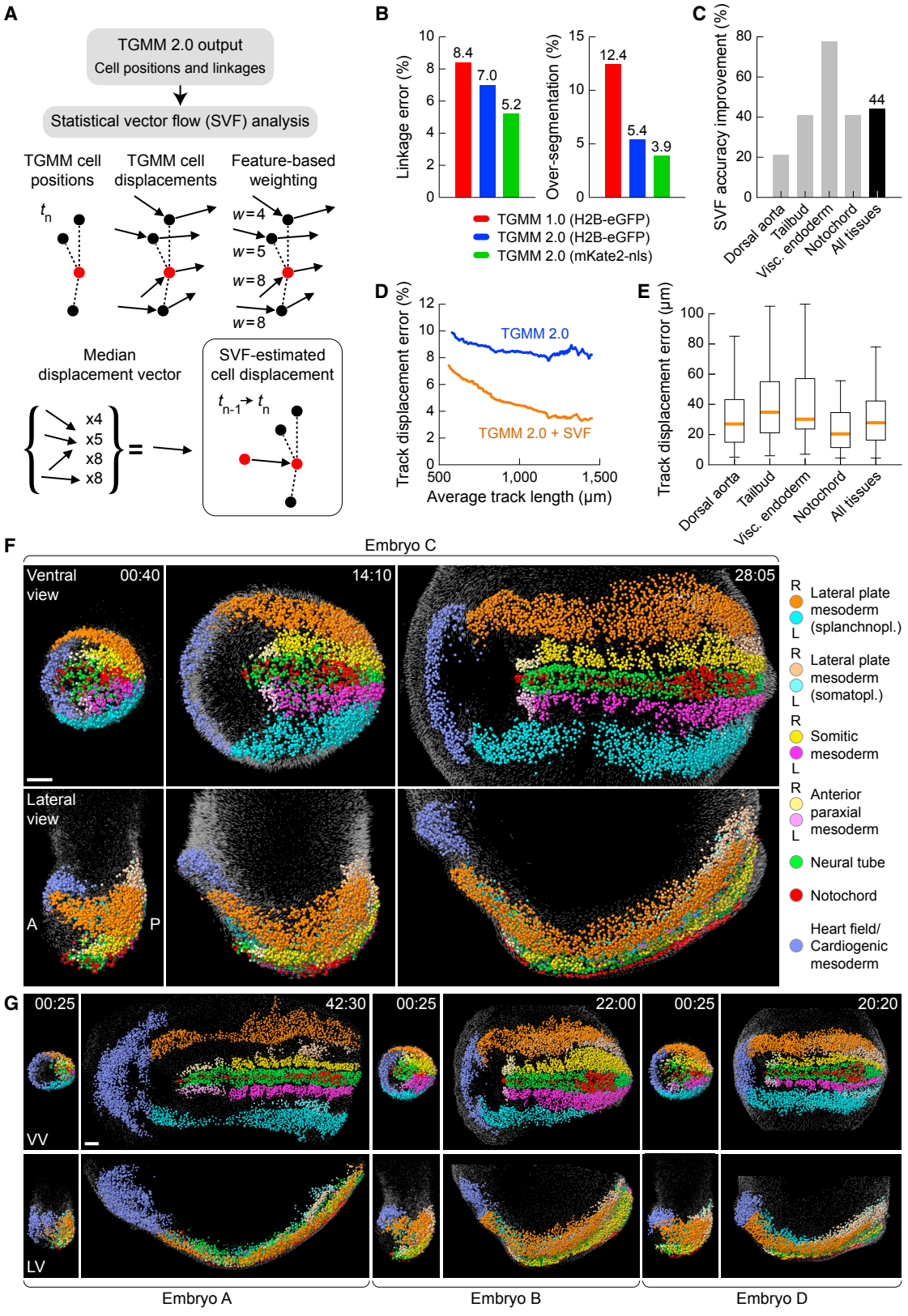
(A) Overview of image processing and data analysis modules.

(B) Cell division prediction with a convolutional neural network (CNN). Top: projection of example image region. Middle: predicted locations of dividing cells (intensity reflects prediction confidence). Bottom: images overlaid with predictions, showing correct identification of two neighboring dividing cells.

(C) Precision-recall curves of CNN division detector for two embryos at three time points (training on embryo A only). TGMM 2.0 machine learning classifier shown for comparison.

Scale bar (μm): 10 (B).

See also Video S4A.



(legend on next page)

confronting the tracking algorithm itself; rapid cell movements and heterogenous nuclear shapes and sizes complicate image analysis. Fortunately, the use of histone reporters for tracking cell nuclei provides an additional benefit: divisions are readily identifiable by the condensation and increase in brightness of chromatin as the metaphase plate appears, followed by the appearance of two smaller, bright daughters. This stereotyped event can be easily identified, even in regions of poor contrast and between different cell types. We thus sought to leverage the strength of deep learning and train a convolutional neural network (CNN) to automatically and accurately identify cellular divisions in 3D and time (STAR Methods; Data S1E).

Our network correctly locates and identifies the presence of a single metaphase plate even in a highly crowded cellular environment or where multiple divisions are occurring in the same spatiotemporal window (Figure 3B). We determined the precision and recall of this detector to be 0.81 and 0.81, respectively; i.e., 81% of all detections are true cell divisions, and 81% of all true divisions in the embryo are retrieved by the CNN (*f*-score = 0.81; n = 381 annotated divisions in three volumes), which corresponds to a detection accuracy of 99.8%. This represents a 3.7-fold improvement in precision and a 24% increase in recall compared to TGMM 2.0 (Table S1). We confirmed that the CNN can be effectively applied to new data without the need for retraining: application to data from a second embryo produced an *f*-score of 0.80 (n = 479 annotated divisions in three volumes), indicating that the network generalizes well to new experiments (Figure 3C). Notably, the CNN was able to retrieve on average more than twice the number of cell divisions found by a human annotator. A time-lapse reconstruction of cell divisions across an entire developing embryo is shown in Video S4A and included in Data S1P.

Reconstructing High-Resolution Dynamic Cell Fate Maps Across the Mouse Embryo

Because nuclei of even the same lineage can experience huge variations in size, shape, and intensity, we performed a detailed parameter sweep to determine the regime of optimal TGMM performance in post-implantation development (Figure S4). Even still, the photosensitivity of the mouse embryo limits temporal resolution, and rapid cell movements in the mouse embryo combined with their complex, fluctuating shapes introduce temporal gaps and/or linkage errors in cell tracks over long timescales. While the shorter, fragmented tracks generated by TGMM faithfully recapitulate the movements of individual cells, TGMM alone is insufficient to reconstruct cell fates over an entire 48 hr experiment. We thus took advantage of the high-quality data TGMM provides for local spatiotemporal windows and enhanced the accuracy of long-term cell tracking by augmenting individual cell tracks through a statistical assessment of cell behavior in a cell's local neighborhood (Figure 4A). The integration of TGMM 2.0 with this statistical vector flow (SVF) analysis not only improves spatial accuracy of cell tracks, but also reconstructs continuous cell tracks (rather than sets of short, disconnected tracks) for the full duration of the experiment (Video S4B; STAR Methods; Data S1G and S1H).

Compared to TGMM 2.0 alone (Figure 4B), combined use of TGMM and SVF has a positive effect on cell tracking performance across all tissues (Table S1), including difficult cases such as the thin notochord, and results in a 44% improvement in accuracy across all cell tracks (Figure 4C; n = 285 annotated cell tracks). This number reflects the decrease in spatial mismatch between automatically versus manually reconstructed cell tracks. This performance boost increases with track length and tracking period: for tracks with a length of 1,500 μm, SVF improves accuracy by as much as 140%, reducing track displacement errors (versus ground truth) to 3% (Figure 4D). Overall, we determined that our combined TGMM and SVF approach reconstructs individual cell tracks with an average error of 34.2 μm (two cell diameters) over 48 hr and across all tissues (Figure 4E).

With the ability to accurately reconstruct cell tracks from beginning to end we systematically queried the origin of cell fates by doing the following: (1) manually segmenting individual tissues or regions of interest in the image data at the last time point of the experiment, (2) then locating the corresponding SVF objects within these manual tissue labels, and (3) following the labeled cell tracks backward in time to the beginning of the experiment. This approach provides a complete reconstruction of the trajectories of cells as they migrate out of the primitive streak and assemble into their corresponding tissues. As a first

Figure 4. Long-Term Cell Tracking and Reconstruction of Cell Fates in the Mouse Embryo

(A) SVF workflow: SVF-based cell displacement estimate is computed from TGMM-based cell displacements using feature-based weighting. The degree of neighborhood conservation is considered to calculate median displacement vectors for each cell.

(B) Linkage errors and over-segmentation in TGMM 1.0 versus 2.0 (n = 8,982/30,962 cell position and linkage annotations for H2B-eGFP and mKate2-nls data, respectively).

(C) SVF post-processing improves TGMM cell tracking accuracy on average by 44% across all cell types (n = 285 cell tracks spanning all stages, H2B-eGFP reporter).

(D) Average track displacement error (versus ground truth) as a function of track length when using TGMM 2.0 + SVF versus TGMM 2.0 alone. Statistics as in C.

(E) Track displacement error for different cell types using TGMM 2.0 + SVF (orange line: mean, box: $25^{th}/75^{th}$ percentiles, whiskers: extremum data within 1.5 × interquartile range). Average error across all cell types is 34.2 μm over full imaging experiment. Statistics as in C.

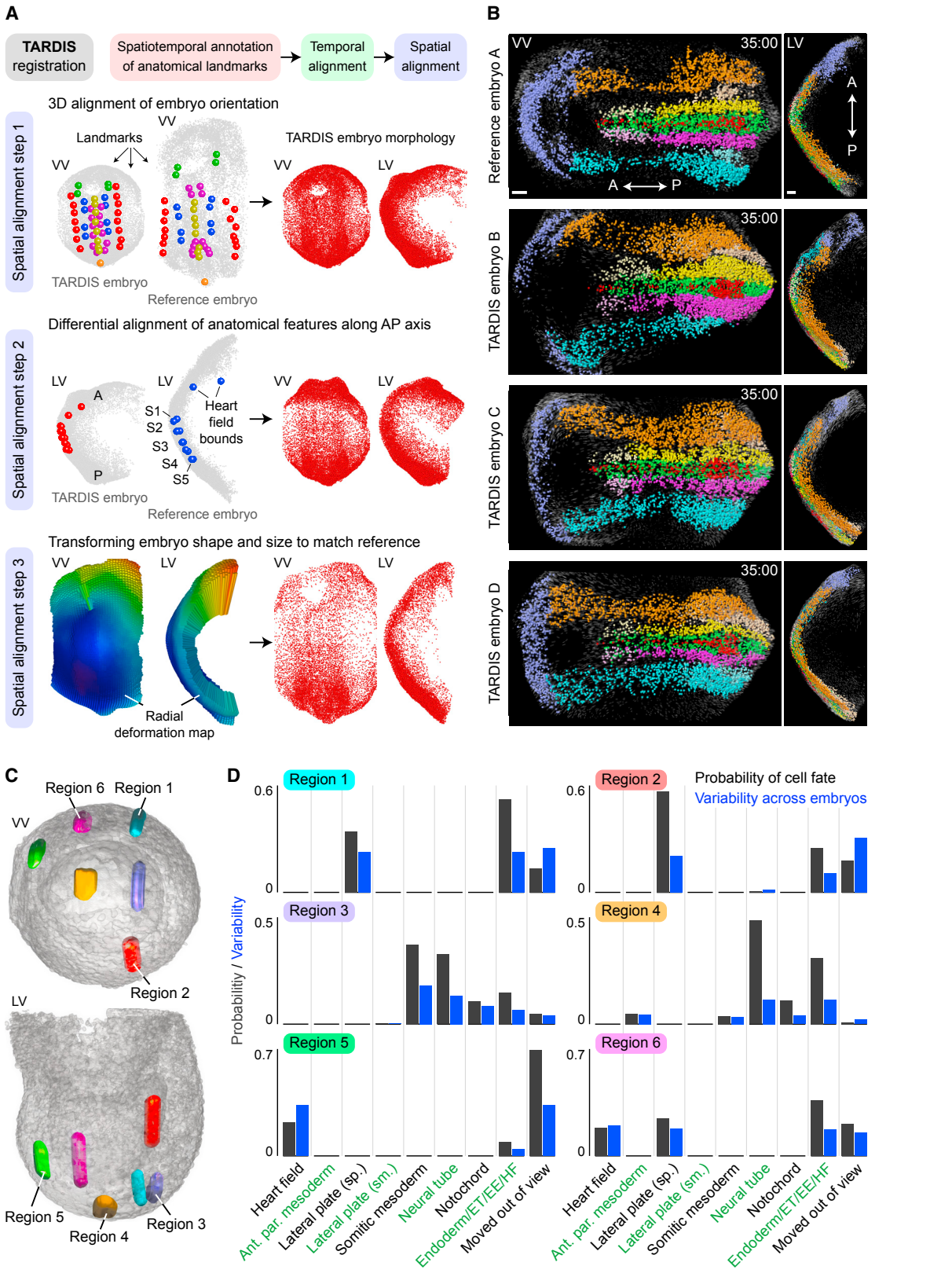
(F) Reconstruction of an embryo shown at three time points of an experiment from mid/late-streak stage to early somite stage, using TGMM 2.0 + SVF. The dynamic fate map was created by labeling tissues in the image data at the last time point, transferring labels to SVF objects (spheres) and propagating labels backward in time.

(G) Cell-fate reconstructions for three additional embryos ("A," "B," "D"). Similar tissue locations and patterning are seen across embryos and are consistent across developmental stages.

Time (hh:mm). VV = ventral view, LV = lateral view.

Scale bars (μm): 200 (F, G).

See also Videos S4B and S5 and Figure S4.



(legend on next page)

demonstration, we labeled a set of tissues based on their easily identifiable anatomical features, but we note that our method is not limited to this selection. The heart field, lateral plate mesoderm (split into splanchnopleure and somatopleure), somitic mesoderm (including condensed somites and pre-somitic mesoderm), anterior paraxial mesoderm (mesoderm anterior of first somite), endoderm, neural tube, and notochord were manually labeled and dynamically reconstructed in four embryos (Video S5A; Figures 4F and 4G). By comparing results across embryos, we found a strong similarity between origins and movements of individual tissue types, regardless of size and shape of the embryo (Video S6A). A time-lapse visualization of a dynamic fate map including the endoderm is shown in Video S5B and an interactive 4D visualization is provided as Data S1O.

Notably, our spatiotemporal fate maps are also in agreement with the individual snapshots that were laboriously generated by others prior to the availability of live imaging techniques (Lawson and Pedersen, 1992; Takaoka et al., 2011; Tam and Beddington, 1992; Tam et al., 1997) and confirm that there is little mixing of the different mesodermal layers once cells exit the primitive streak. This suggests that when and where a cell exits the streak is of critical importance to determining its final fate, an order that is strongly maintained by its local environment. In contrast to early static maps, however, and as shown in the next sections, our dynamic, high-resolution reconstructions allow us to visualize the interactions of multiple cell types over time, comprehensively follow cell behavior in a single embryo, determine the degree of variability between embryos in establishing cell fates, and require no additional physical manipulation or perturbation of the embryo.

Importantly, this technique can be applied to any tissue, cell type, time period, or region of interest within the embryo. In our next example, we used reporters for Brachyury and Foxa2 to further demonstrate the accuracy of SVF and trace the origins of cells that give rise to the node and notochord. We visualized and quantified single-cell behaviors in the convergent extension of the anterior portion of the notochord and subsequent elongation as the posterior axis expands (Video S3D). Using dual-expressing T-mCherry and H2B-eGFP embryos, we tracked cells in the ubiquitous nuclear channel and masked the last time point with the location of T-mCherry positive cells in the node and notochord to seed SVF propagation (Video S5C). Thereby, we verified by direct comparison with the image data that cells tracked over a period of more than 24 h with SVF precisely matched the location and dynamics of the T-mCherry live reporter (Figures S5A and S5B): 94% of all cell tracks matched the dynamic expression pattern and the remaining 6% exhibited an average mismatch of only 5.0 ± 4.4 μm.

Building a Dynamic Atlas of Development by Spatiotemporal Registration of Multiple Embryos

Although we can accurately track and characterize cell fates and tissue dynamics in individual embryos, in order to employ meaningful statistical analyses or measure variability of developmental processes, it is necessary to jointly analyze data from multiple embryos. However, the wide variation in size, shape, and even rate of tissue formation across normal mouse development makes direct comparisons between embryos even of the same developmental stage problematic. We thus developed the registration method TARDIS (time and relative dimension in space) that combines manually annotated spatiotemporal landmarks and information on local cell distributions obtained from TGMM to map multiple embryos in space and time onto a single reference embryo (Figure 5A; Video S6C; STAR Methods; Data S1I).

Using TARDIS, we registered in space and time four different embryos, which encompass a range of developmental timescales and size and shape variations, with an average registration error of 41.5 μm (Video S6D; Figures 5B, S5C, and S5D; unregistered embryos are shown in Figures 4F and 4G). Despite these dramatic differences in size and shape, the resulting TARDIS embryos preserve the motion and morphology of different tissues as compared to their original form. We note that our quantifications of cell dynamics, such as movement speed, movement direction, and cell density, are performed in the original, untransformed embryos to ensure that these quantifications are not biased by the registration procedure itself. Equipped with these registered datasets, we set out to generate a reference atlas of mouse embryonic development (Data S1K–S1O). The purpose of this atlas is not only to assess quantitatively how conserved or variable cellular dynamics are across different individuals, but also to create a statistically average embryo that can be used as a point of comparison between normal embryos (e.g., to query a region of interest before cells in that region are fully specified) or for the quantitative analysis of mutant phenotypes or otherwise perturbed embryos.

A Statistical Fate Map of the Post-Implantation Mouse Embryo

TARDIS makes it possible to assess cell behavior in the developing mouse embryo in a statistically meaningful way. As the first application of this approach, we constructed a whole-embryo statistical fate map (STAR Methods; Data S1R) based on the

Figure 5. Spatiotemporal Registration of Mouse Embryo Development Using TARDIS

(A) TARDIS overview: embryos are aligned in time using manual annotations, then aligned in space by rigid registration to a reference embryo using spatial landmarks (step 1), differential alignment of anatomical features along anterior-posterior axis (step 2), and transformation of their shape and size to match the reference embryo (step 3). Left: examples of landmarks and transformation maps are shown. Right: resulting embryo morphology.

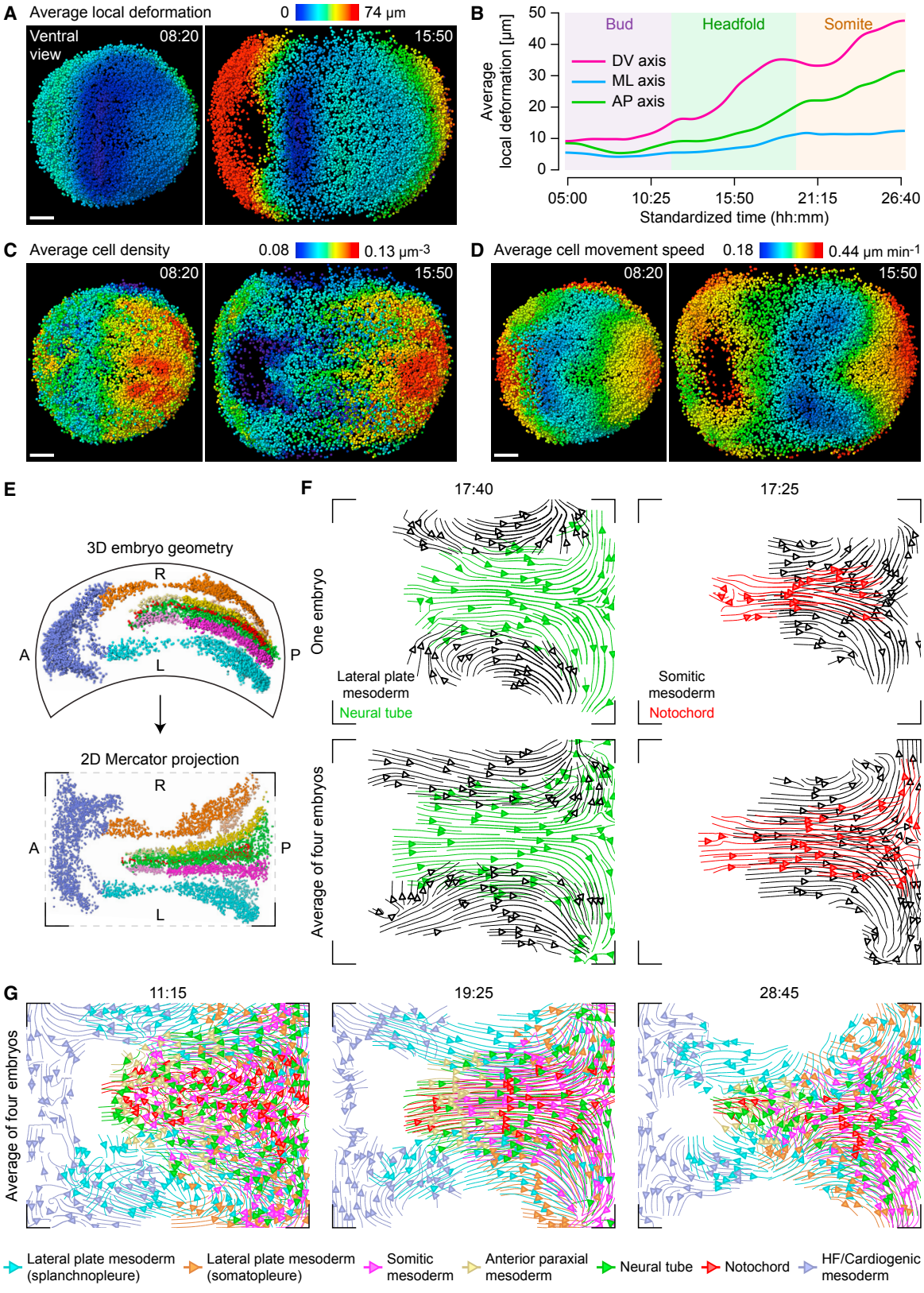
(B) TARDIS registered embryos A–D shown side-by-side. Although embryos can vary dramatically in size and shape prior to registration (Figures 4F and 4G), they are closely aligned post-registration while maintaining overall tissue morphology and movements. Colors as in Figure 4F.

(C and D) Statistical cell fate predictions (D) for six example regions in an embryo selected at an early time point (C) and computed from combined dynamic fate maps of four TARDIS-registered embryos. Dark gray bars represent average probability for a region to become a specific tissue. Blue bars represent variability between embryos (standard deviation of probabilities). Category "Moved out of view" represents cells that moved outside imaging volume during recording.

VV, ventral view. LV, lateral view. A, anterior. P, posterior. ET, endothelial. EE, extra-embryonic. HF, head fold.

Scale bar (μm): 200 (B).

See also Video S6 and Figure S5.



(legend on next page)

tissue annotations we generated for four individual embryos by jointly analyzing the data of the respective TARDIS-registered embryos. We systematically computed the average percentages of local cell fate contributions across all TARDIS embryos for all labeled tissues (Video S6E). These combined statistical results closely correspond with and support what has been observed for past tissue or single-cell labeling of individual embryos (Lawson, 1999; Lawson and Pedersen, 1992; Tam and Beddington, 1992). However, while past observations relied on small cell populations or individually labeled cells, our method enables us not only to visualize cell fate dynamics and establishment of all cell populations in an embryo simultaneously, but also to make direct comparisons between multiple embryos.

Although it is usually possible to determine the identity of a cell at the end of an experiment, this is not helpful if one wishes to know in real time where a cell might end up before arriving at that location. Labeling, isolating, optogenetically manipulating, or otherwise perturbing cell fates requires *a priori* knowledge of the probability of cells to assume a particular fate. Using the TARDIS atlas, we queried the probability that certain cell populations in a mid- to late-streak stage embryo would assume the fate of one of the previously labeled tissue groups by manually selecting regions of interest in the query embryo and visualizing the final predicted contribution to different tissues 24 hours later (Figures 5C and 5D). We included difficult "edge" cases, such as the endoderm (region 1), as well as locations in the mesoderm (region 2), ectoderm (region 4), and combinations of both (region 3) (Figures 5C and 5D). Correct statistical classification of the endoderm region is exceptionally difficult, as it requires both accurate long-term tracking of this single-cell-layer thick tissue and precise registration of all TARDIS embryos with an accuracy approaching a single-cell diameter. Nevertheless, even for this challenging setting, the majority of TARDIS predictions were indeed in the endoderm-containing category, with a smaller contribution from the underlying lateral plate mesoderm. Additionally, we tracked a set of individual cells through the developing embryo and showed how the TARDIS database can be used to measure, as a function of time, the cell fate composition in the local neighborhoods of these tracked cells as well as the variability of cell fate composition across embryos (Video S5D).

While TARDIS provides a useful tool to enable the mapping of multiple embryos onto one reference scaffold, we also designed a complementary method to combine the information from multiple embryos into a single, average embryo (STAR Methods). Related concepts have been successfully employed for later developmental stages, using fixed embryos imaged with micro-CT (Wong et al., 2014). By contrast, our approach relies on dynamic cellular information obtained from fluorescence imaging, using live embryos at earlier stages. This approach creates a statistically "average" embryo (Data S1S, Video S7), avoiding the use of a reference embryo as a physical scaffold for data visualization and analysis, and allows for the direct measurement and visualization of average embryo shape and development as well as of the variability of shape and developmental parameters across individuals (Figures S3E–S3G; Video S7A–S7C). We built this average embryo as the mean of our four reconstructed embryos after registering all embryos to their global center of mass.

As an example of the different metrics that can be quantified from the average embryo, we examined the average local deformation, cell density, and cell movement speed across the development of the average embryo (Figures 6A–6D and S5E–S5G; Video S7A–S7C). The average local deformation (Figures 6A and 6B) reports to which extent the contributing embryos had to be deformed in a given area to match the average embryo shape, i.e., it quantifies variability in local shape of individuals. Across the DV, ML, and AP axes, average variability in shape is less than 50 μm, and local deformations are generally significantly smaller, particularly at early developmental stages or along the ML and AP axes. As with the TARDIS registration, individual points or tissue regions can be labeled in the average embryo and propagated in time (Data S1J, STAR Methods) to determine their developmental origin, fate, or the variability in cell neighborhood over time and across embryos (Video S7D). The unique strength of performing such analyses in the average embryo is that the results represent the statistically average cell behaviors across individuals.

It is important to note here that the labeling scheme for the query regions and predictions for both TARDIS and the average embryo can be entirely arbitrary. For example, labeling criteria could include anatomical features, a second fluorescent reporter, or the expression of genes of interest. To demonstrate the latter strategy, we used the T-mCherry reporter and masking scheme for SVF described previously to map a single time point of an H2B-eGFP and T-mCherry expressing embryo onto the average embryo. Using a single time point from the T-mCherry expressing embryo to label corresponding T-positive locations

Figure 6. Quantifying Stereotypy and Variability of Local Cell Dynamics Across Embryos

(A and B) Visualization and quantification of differences in local embryo shape across four rigidly aligned embryos. DV, dorsoventral, ML, mediolateral, AP, anteroposterior.

(C) Average local cell densities and (D) average local cell movement speeds at two time points in the average embryo.

(E) 2D flattening of a 3D embryo using the Mercator projection. The 3D embryo geometry is represented using a spherical coordinate system with radial dimension ρ and angular dimensions θ and ϕ (corresponding to the horizontal and vertical dimensions of the Mercator map). A, anterior. P, posterior. L/R, left/right side of embryo.

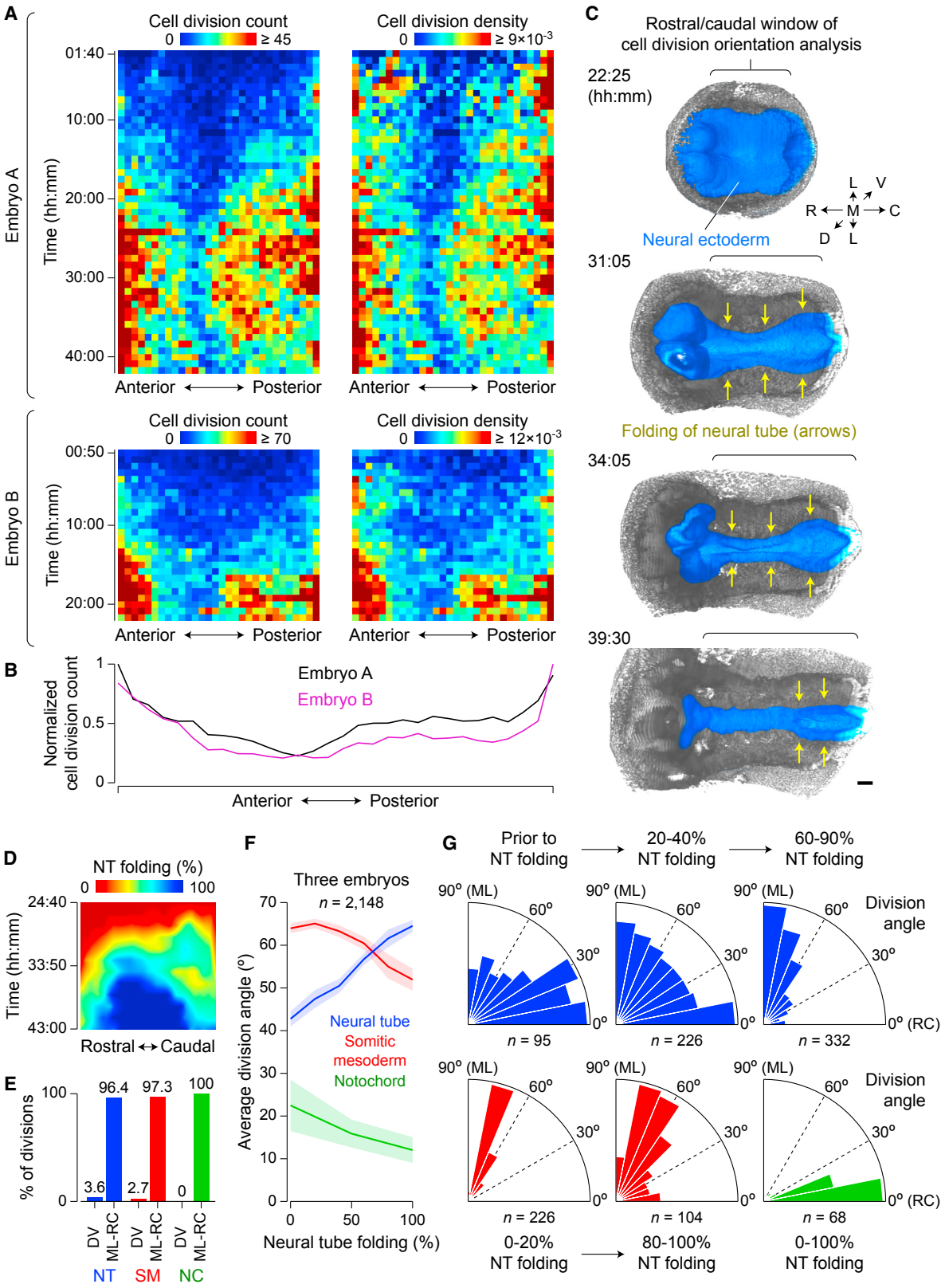
(F) Overlay of cell movements in lateral plate mesoderm and neural tube (left) and in somitic mesoderm and notochord (right), for a single embryo (top) and averaged across four embryos at the early head fold stage. 3D tissue movements were computed from respective cell movements reconstructed with TGMM 2.0 + SVF over 100 min and visualized in 2D using Mercator projection.

(G) Overlay of movements for all labeled tissues combined into one morphodynamic fate map, shown at mid-bud, early head fold, and early somite stages. Top: map for single embryo, bottom: average map for four embryos.

Time (hh:mm).

Scale bars (μm): 100 (A, C, D).

See also Videos S4C, S4D, and S7 and Figure S5.



(legend on next page)

in the average embryo, we tracked the progression and location of T-positive cells in the average embryo backward and forward in time (Video S7D). These results are in close agreement with the expected expression patterns and dynamics of notochord development visualized in live embryos (Video S3D). To facilitate the general use of our average embryo database for reconstructing the developmental history and fate of arbitrary cell populations defined by image data obtained from other imaging experiments (using live or fixed embryos), we provide software tools for mapping snapshot image data of gene expression patterns and other types of label masks onto the average embryo database (Data S1J).

Mapping Tissue Morphodynamics

Using TARDIS, we combined quantitative measurements of the behavior of individual cells and tissues across individuals and made comparisons where they might otherwise be prevented by differences in embryo geometry or developmental timing. Complementing our statistical analysis of cell fate described above, we developed a similar approach for the quantitative investigation of cell movements, specifically the direction and speed of cell populations during the formation of tissues (STAR Methods; Data S1T and S1U). The construction of such maps of tissue morphodynamics allows us not only to visualize the flow of different tissues in comparison to their neighbors, but to measure changes in velocity and tissue size over the course of development as well as variability in tissue movements across individuals (Figures 6E–6G; Video S4C and S4D).

A Spatiotemporal Switch in the Orientation of Cell Divisions during Neural Tube Morphogenesis

In this final section, we combined our dynamic fate maps created with TGMM 2.0 and SVF, our maps of tissue morphodynamics and our CNN-based reconstruction of cell divisions across the developing embryo into a single database (STAR Methods). With this resource, we can jointly analyze the spatiotemporal distribution of cell divisions, local cell movements, and cell fates and large-scale morphogenetic changes over the course of development. A visualization of the combined data in this database (Video S6B) showed that the vast majority of cell divisions in later stages of development occur in the anterior- and posterior-most regions of the embryo (Figure 7A). Comparative analysis of a second embryo confirmed that divisions generally cluster at the leading edge of growing tissues (Figure 7A) and revealed similarities not only at the ends but along the entire AP axis (Figure 7B).

The high spatiotemporal resolution of our image data enabled the analysis of cell divisions not just by their location in the embryo but also with respect to their orientation. For example, the striking changes in tissue geometry during neural tube formation led us to ask whether the rate and orientation of divisions in the neural tube changed as a function of elongation and closure. The first site of neural tube closure can be visualized at the hindbrain/cervical boundary, after which the neural tube progressively closes in a zipper-like fashion in both rostral and caudal directions (Video S2D). An examination of division orientation using fixed tissues in chick and mouse (Sausedo et al., 1997) determined that a large percentage of cells divided along the RC axis, with some dividing along the ML axis during bending, but was unable to correlate this with the progression of neural tube closure. The image data and computational reconstructions generated here allowed us to examine this dynamic, mechanical process in detail (Figures 7C–7F). While the neural ectoderm is the most challenging part of the embryo to capture at a single-cell level, the vast majority of divisions occur on the luminal surface, making these events well separated and easily identifiable. Additionally, we quantified the orientation of cell divisions in neighboring tissues to determine if division orientation was a shared function of axis elongation or unique to the developmental stage of the tissue in question.

Prior to neural tube closure, when the primary tissue-wide morphogenetic process in the neural tube is tissue elongation, cells in the neural tube preferentially orient their divisions along the RC axis (Figure 7G). Division orientation subsequently becomes more uniformly distributed as folding progresses. Then, as the neural tube reaches the final stages of closure, cells exhibit a strong tendency to divide along the ML axis. As elongation proceeds and the neural plate narrows, it begins to bend dorsally, elevating the neural folds toward the midline, and divisions reorient to divide along the ML axis. We furthermore observed what appears to be significant 90° spindle rotation prior to division in the neural tube during these stages; however, higher temporal resolution will be necessary to quantify this behavior.

Finally, we examined neighboring tissues to determine if they too showed a preference for division orientation or a change in division angle preference as a function of their development (Figure 7G). In the somitic mesoderm, newly condensed or condensing somites appear to display a preference for ML divisions. However, this preference is lost over time as somites mature. By contrast, notochord cells, while largely quiescent, divide almost exclusively along the rostral-caudal axis. While the embryo-wide process of axis elongation almost certainly

Figure 7. A Temporal Switch in Cell Division Orientation During Neural Tube Closure

(A) Spatiotemporal distributions (left) and densities (right) of cell divisions in two embryos. As a function of time, embryos were subdivided along anterior-posterior axis into 30 sectors with equal cell counts (horizontal axis).

(B) Time-integrated, normalized cell division distributions for embryos in (A).

(C) Surface rendering of neural tube (NT) development as seen from inside of embryo.

(D) Neural tube folding map obtained by manual annotations of local NT folding angles (n = 564). 100% folding corresponds to NT closure at that location.

(E) Percentages of divisions that occur primarily along dorso-ventral axis (DV) versus within mediolateral-rostrocaudal plane (ML-RC) for three tissues.

(F) Division angle versus NT folding progression for three tissues, based on division data from three embryos.

(G) Distribution of division angles as a function of NT folding.

R, rostral. C, caudal. D, dorsal. V, ventral. M, medial. L, lateral. SM, somitic mesoderm. NC, notochord.

Scale bar (μm): 100 (C).

See also Videos S2D, S4A, and S6B.

imposes some degree of physical force on these tissues, our observations here imply that cell division orientation, at least, is a more local feature distinct to each tissue in space and time.

DISCUSSION

A Framework for Studying Whole-Embryo Development at the Single-Cell Level

The light-sheet microscope presented and applied here enabled, for the first time, imaging of mouse development *in toto* and with cellular resolution from gastrulation to early organogenesis. The resulting image data allowed us to generate the first dynamic fate maps, systematically visualize tissue formation and developmental processes at an unprecedented level of detail, and quantitatively examine dynamic processes such as the embryo-wide patterns of cell division and cell behavior during neural tube formation. In the future, further advances of this technique and the use of new, brighter, and near-infrared fluorescent reporters will aid in our ability to resolve with even greater detail those anatomical regions with the most challenging optical properties, such as the deepest layers of the ectoderm. Such advances may also enable the imaging of later stages in development, large explants or small organ cultures, where lack of transparency becomes the major limitation for light microscopy.

With the ability to systematically image the dynamic behavior of individual cells throughout the embryo came a need to develop a suite of computational methods capable of processing and analyzing the massive and highly complex image data. The result of this synergy between advances in imaging and computational image analysis is the generation of a statistical atlas of mouse development, which systematically maps cell movements, cell proliferation, cell fates, and tissue-level morphodynamic processes across the embryo and furthermore quantifies variability of these developmental parameters across individuals. These datasets and tools provide a wealth of information that can be used to address a wide range of questions on post-implantation development. For example, the ability to predict the final positions and fates of cells will allow for sampling individual cell populations at different points in time and profiling or altering their various transcriptional and behavioral states as they develop. Additionally, our dynamic atlas can be used as a platform for high-resolution mutant phenotyping by providing a systematic, quantitative comparison of "normal" development to mutants or altered states.

Dissemination and application of microscopy and image analysis methods

The ability to quantitatively examine the morphodynamic events patterning a developing embryo is of substantial value to any field. Importantly, the imaging system, computational tools, and resources we provide here are not limited to the study of post-implantation mouse development, or even to only mouse development, and may be broadly adopted for a wide range of developmental organisms and timescales, including tissue and stem cell systems, organoids or 3D culture systems. The software tools developed in this work for data management, multi-view image processing, and TGMM-based cell segmentation and tracking are applicable to a wide spectrum of live imaging data and are now also used routinely in fruit flies, zebrafish, crustaceans, and organ explants. In future work, we furthermore envision an extension of the CNN-based cell division detector, SVF cell tracking method, and TARDIS framework for multi-embryo registration to other model systems.

Considering this wide application space, we have designed all tools and resources accordingly so that they may serve a wide spectrum of groups and users. The documented code of all computational tools (Data S1B–S1J), a comprehensive practical guide to using our software including example image datasets (STAR Methods), the blueprints of the microscope (Data S1A), all databases of mouse embryonic development generated in this study (Data S1K–S1U), and interactive 4D visualizations of our cellular-resolution dynamic atlas (Data S1O) are available as a public repository. Future work will focus on continued improvements to this computational toolkit to enable new analyses of complex image data, as well as developing a new cell lineaging framework based in its entirely on a deep learning paradigm.

Access to the custom—and often highly experimental—microscopes that enable such advancements has been notoriously difficult. However, it is of critical importance that such instruments are made available to the community immediately and independently of any need for synergy or collaborative coordination. As such, we have partnered with the Janelia Advanced Imaging Center and built a copy of our adaptive light-sheet microscope with all capabilities and functions described in this work. This instrument is dedicated solely to the free use by the general scientific community and maintained by a team of experts that guide users through sample preparation, imaging, and image analysis (https://www.aicjanelia.org/).

The steady march of advancing imaging technologies and new computational methods makes this a truly exciting time for the field of mouse development, which has largely been left behind by new imaging strategies in favor of more 'amenable' organisms like D*rosophila* or zebrafish. The comprehensive methodological framework we have presented here and the generation of a dynamic atlas of early mouse development will hopefully serve as the first step in the next phase of our understanding of mammalian development as a whole.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Transgenic mice and reporters
  - Sample preparation, embryo culture and imaging of mouse embryos
- METHOD DETAILS
  - Adaptive light-sheet microscope for imaging post-implantation mouse development
  - Large-scale image data management
  - Robust affine 3D registration of multi-view and time-lapse image data
  - TGMM 2.0 framework for cell nuclei segmentation and tracking

- Long-term tracking of cell movements with Statistical Vector Flow (SVF)
- Convolutional neural network for automated detection of cell divisions
- TARDIS algorithm for spatiotemporal registration of multiple embryos
- Constructing an average embryo from multiple reconstructed embryos
- Probabilistic and statistical cell fate maps
- Computation of tissue morphodynamics maps
- A practical guide to the framework for analyzing image data of mouse development

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, seven videos, two data files, and one methods file and can be found with this article online at https://doi.org/10.1016/j.cell.2018.09.031.

A video abstract is available at https://doi.org/10.1016/j.cell.2018.09.031#mmc12.

ACKNOWLEDGMENTS

We thank N. Clack and A. Taylor for help with developing MVD and TGMM 2.0 software; B. Coop and Janelia jET for mechanical designs and custom parts; M. Coleman for LabVIEW software; Janelia AIC for providing free microscope access; J. Kuhl for illustrations in Figures S1A–S1C; C. Wang for help with custom objective designs; B. Höckendorf for help with developing KLB; S. Preibisch for helpful discussions about image registration; J. Liu for Brachyury-H2B-mCherry mice; B. Gu, E. Posfai, and J. Rossant for H2B-miRFP703 mice; S. Srinivas, T. Watanabe, W. Lemon and A. Hantman for help with exploring mouse live imaging assays; J. Green and S. Srinivas for helpful comments on the manuscript; and Janelia Vivarium Team for animal care. All aspects of this work were funded by the Howard Hughes Medical Institute. G.M. acknowledges additional funding from the French National Research Agency (Project Dig-Em: ANR-14-CE11-0013) and L.A.R. from the Chan Zuckerberg Biohub.

AUTHOR CONTRIBUTIONS

Conceptualization: P.J.K., K.M., and L.G.; Methodology: K.M., P.J.K., L.G., F.A., K.B., A.B., L.A.R., S.C.T., and G.M.; Software: L.G., P.J.K., F.A., A.B., L.A.R., K.B., and G.M.; Formal Analysis: L.G., P.J.K., F.A., and K.M.; Investigation: K.M., L.G., F.A., and A.B.; Data Curation: K.M.; Writing: P.J.K., K.M., and L.G.; Visualization: P.J.K., K.M., and L.G.; Supervision: P.J.K., S.C.T., and K.B.; Project Administration: P.J.K., K.M., S.C.T., and K.B.; Funding: P.J.K.

DECLARATION OF INTERESTS

P.J.K., L.A.R., and R. Chhetri filed provisional US patent application 62,354,384 for adaptive light-sheet microscopy on June 24, 2016. P.J.K. holds US patent 9,404,869 for simultaneous multi-view light-sheet microscopy issued on August 2, 2016.

Received: December 1, 2017
Revised: May 12, 2018
Accepted: September 13, 2018
Published: October 11, 2018

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevedo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. arXiv, 1603.04467. https://arxiv.org/abs/1603.04467.

Abe, T., Kiyonari, H., Shioi, G., Inoue, K., Nakao, K., Aizawa, S., and Fujimori, T. (2011). Establishment of conditional reporter mouse lines at ROSA26 locus for live cell imaging. Genesis *49*, 579–590.

Amat, F., and Keller, P.J. (2013). 3D Haar-Like Elliptical Features for Object Classification in Microscopy. ISBI, 1194–1197.

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., Höckendorf, 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.

Anderson, R., Copeland, T.K., Schöler, H., Heasman, J., and Wylie, C. (2000). The onset of germ cell migration in the mouse embryo. Mech. Dev. *91*, 61–68.

Astola, J., Haavisto, P., and Neuvo, Y. (1990). Vector Median Filters. Proc. IEEE *78*, 678–689.

Attali, D., and Boissonnat, J.D. (2003). Complexity of the Delaunay triangulation of points on polyhedral surfaces. Discrete. Comput. Geom. *30*, 437–452.

Besl, P.J., and Mckay, N.D. (1992). A Method for Registration of 3-D Shapes. PAMI *14*, 239–256.

Campolo, F., Gori, M., Favaro, R., Nicolis, S., Pellegrini, M., Botti, F., Rossi, P., Jannini, E.A., and Dolci, S. (2013). Essential role of Sox2 for the establishment and maintenance of the germ cell line. Stem Cells *31*, 1408–1421.

Chhetri, R.K., Amat, F., Wan, Y., Höckendorf, B., Lemon, W.C., and Keller, P.J. (2015). Whole-animal functional and developmental imaging with isotropic spatial resolution. Nat. Methods *12*, 1171–1178.

Gabriel, K.R., and Sokal, R.R. (1969). A new statistical approach to geographic variation analysis. Syst. Zool. *18*, 259–278.

Ginsburg, M., Snow, M.H., and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. Development *110*, 521–528.

Gu, B., Posfai, E., and Rossant, J. (2018). Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. Nat. Biotechnol. *36*, 632–637.

Guignard, L., Godin, C., Fiuza, U.M., Hufnagel, L., Lemaire, P., and Malandain, G. (2014). Spatio-temporal registration of embryo images. ISBI, 778–781.

Guimond, A., Meunier, J., and Thirion, J.P. (2000). Average brain models: A convergence study. Comput. Vis. Image Underst. *77*, 192–210.

Herion, N.J., Salbaum, J.M., and Kappen, C. (2014). Traffic jam in the primitive streak: the role of defective mesoderm migration in birth defects. Birth Defects Res. A Clin. Mol. Teratol. *100*, 608–622.

Huh, S., Ker, D.F.E., Bise, R., Chen, M., and Kanade, T. (2011). Automated Mitosis Detection of Stem Cell Populations in Phase-Contrast Microscopy Images. IEEE Trans. Med. Imaging *30*, 586–596.

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E.H.K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science *305*, 1007–1009.

Ichikawa, T., Nakazato, K., Keller, P.J., Kajiura-Kobayashi, H., Stelzer, E.H., Mochizuki, A., and Nonaka, S. (2014). Live imaging and quantitative analysis of gastrulation in mouse embryos using light-sheet microscopy and 3D tracking tools. Nat. Protoc. *9*, 575–585.

Imuta, Y., Kiyonari, H., Jang, C.W., Behringer, R.R., and Sasaki, H. (2013). Generation of knock-in mice that express nuclear enhanced green fluorescent protein and tamoxifen-inducible Cre recombinase in the notochord from Foxa2 and T loci. Genesis *51*, 210–218.

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. ICML, 448–456.

Jones, E.A., Crotty, D., Kulesa, P.M., Waters, C.W., Baron, M.H., Fraser, S.E., and Dickinson, M.E. (2002). Dynamic in vivo imaging of postimplantation mammalian embryos using whole embryo culture. Genesis *34*, 228–235.

Kabra, M., Robie, A.A., Rivera-Alba, M., Branson, S., and Branson, K. (2013). JAABA: interactive machine learning for automatic annotation of animal behavior. Nat. Methods *10*, 64–67.

Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H.K. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322*, 1065–1069.

Kingma, D., and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv, 1412.6980. https://arxiv.org/abs/1412.6980.

Kuhn, H.W. (1955). The Hungarian method for the assignment problem. Nav. Res. Log. *2*, 83–97.

Lawson, K.A. (1999). Fate mapping the mouse embryo. Int. J. Dev. Biol. *43*, 773–775.

Lawson, K.A., and Pedersen, R.A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. Development *101*, 627–652.

Lawson, K.A., and Pedersen, R.A. (1992). Clonal analysis of cell fate during gastrulation and early neurulation in the mouse. Ciba Found. Symp. *165*, 3–26.

Liu, Y.K. (2013). Noise reduction by vector median filtering. Geophysics *78*, V79–V86.

Lolas, M., Valenzuela, P.D., Tjian, R., and Liu, Z. (2014). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. Proc. Natl. Acad. Sci. USA *111*, 4478–4483.

Matula, D.W., and Sokal, R.R. (1980). Properties of Gabriel graphs relevant to geographic variation research and the clustering of points in the plane. Geogr. Anal. *12*, 205–222.

Molyneaux, K.A., Stallock, J., Schaible, K., and Wylie, C. (2001). Time-lapse analysis of living mouse germ cell migration. Dev. Biol. *240*, 488–498.

Ourselin, S., Roche, A., Prima, S., and Ayache, N. (2000). Block matching: A general framework to improve robustness of rigid registration of medical images. MICCAI *1935*, 557–566.

Piliszek, A., Kwon, G.S., and Hadjantonakis, A.K. (2011). Ex utero culture and live imaging of mouse embryos. Methods Mol. Biol. *770*, 243–257.

Rivera-Pérez, J.A., and Hadjantonakis, A.K. (2014). The dynamics of morphogenesis in the early mouse embryo. Cold Spring Harb. Perspect. Biol. *7*. https://doi.org/10.1101/cshperspect.a015867.

Royer, L.A., Lemon, W.C., Chhetri, R.K., Wan, Y., Coleman, M., Myers, E.W., and Keller, P.J. (2016). Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms. Nat. Biotechnol. *34*, 1267–1278.

Saitou, M., and Yamaji, M. (2012). Primordial germ cells in mice. Cold Spring Harb. Perspect. Biol. *4*. https://doi.org/10.1101/cshperspect.a008375.

Sausedo, R.A., Smith, J.L., and Schoenwolf, G.C. (1997). Role of nonrandomly oriented cell division in shaping and bending of the neural plate. J. Comp. Neurol. *381*, 473–488.

Shoemake, K. (1985). Animating rotation with quaternion curves. SIGGRAPH *19*, 245–254.

Strnad, P., Gunther, S., Reichmann, J., Krzic, U., Balazs, B., de Medeiros, G., Norlin, N., Hiiragi, T., Hufnagel, L., and Ellenberg, J. (2016). Inverted light-sheet microscope for imaging mouse pre-implantation development. Nat. Methods *13*, 139–142.

Susaki, E.A., Tainaka, K., Perrin, D., Yukinaga, H., Kuno, A., and Ueda, H.R. (2015). Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. Nat. Protoc. *10*, 1709–1727.

Takaoka, K., Yamamoto, M., and Hamada, H. (2011). Origin and role of distal visceral endoderm, a group of cells that determines anterior-posterior polarity of the mouse embryo. Nat. Cell Biol. *13*, 743–752.

Tam, P.P., and Beddington, R.S. (1992). Establishment and organization of germ layers in the gastrulating mouse embryo. Ciba Found. Symp. *165*, 27–41, discussion 42–49.

Tam, P.P., Parameswaran, M., Kinder, S.J., and Weinberger, R.P. (1997). The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development *124*, 1631–1642.

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.

Udan, R.S., Piazza, V.G., Hsu, C.W., Hadjantonakis, A.K., and Dickinson, M.E. (2014). Quantitative imaging of cell dynamics in mouse embryos using light-sheet microscopy. Development *141*, 4406–4414.

Wolff, C., Tinevez, J.Y., Pietzsch, T., Stamataki, E., Harich, B., Guignard, L., Preibisch, S., Shorte, S., Keller, P.J., Tomancak, P., et al. (2018). Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb. Elife *7*, e34410.

Wong, M.D., Maezawa, Y., Lerch, J.P., and Henkelman, R.M. (2014). Automated pipeline for anatomical phenotyping of mouse embryos using micro-CT. Development *141*, 2533–2541.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Fluorobrite DMEM	ThermoFischer Scientific	Cat# A1896701
Matrigel GFR, phenol-red free	Corning	Cat# 356231
Dental wax	Darby Dental Supply, LLC	Cat# 0928094
Penicillin-Streptomycin	ThermoFischer Scientific	Cat# 15140122
GlutaMax	ThermoFischer Scientific	Cat# 35050061
MEM Non-Essential Amino Acids	ThermoFischer Scientific	Cat# 11140076
Fetal Bovine Serum	ThermoFischer Scientific	Cat# 10082139
Glass capillary tubes	Hilgenberg	Cat# 1470253
Teflon FEP Tubes	Zeus Inc.	Custom order
Deposited Data		
Technical drawings of the adaptive light-sheet microscope	This paper	Data S1A https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo A: database of cell tracks with tissue annotations	This paper	Data S1K https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo B: database of cell tracks with tissue annotations	This paper	Data S1L https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo C: database of cell tracks with tissue annotations	This paper	Data S1M https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo D: database of cell tracks with tissue annotations	This paper	Data S1N https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo A: interactive visualization of annotated cell tracks	This paper	Data S1O https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo A: database of cell divisions	This paper	Data S1P https://doi.org/10.17632/fky6dkhmb3.1
Average embryo: landmark annotations in the average embryo	This paper	Data S1Q https://doi.org/10.17632/fky6dkhmb3.1
TARDIS statistical fate map: statistical integration of four embryos	This paper	Data S1R https://doi.org/10.17632/fky6dkhmb3.1
Average embryo: statistical integration of four embryos	This paper	Data S1S https://doi.org/10.17632/fky6dkhmb3.1
TARDIS embryo A: database of tissue morphodynamics	This paper	Data S1T https://doi.org/10.17632/fky6dkhmb3.1
TARDIS database of tissue morphodynamics: statistical integration of four embryos	This paper	Data S1U https://doi.org/10.17632/fky6dkhmb3.1
4D microscopy data of embryos A-D imaged with adaptive light-sheet microscopy	This paper	https://idr.openmicroscopy.org/webclient/?show=project-502
Experimental Models: Organisms/Strains		
Mouse: H2B-eGFP	RIKEN CLST	Abe et al., 2011
Mouse: H2B-mCherry	RIKEN CLST	Abe et al., 2011
Mouse: R26-CAG-nuc-3xmKate2-nls	RIKEN BRC	Susaki et al., 2015
Mouse: CAG-tdTomato-2A-H2B-EGFP	Srinivas laboratory	Trichas et al., 2008
Mouse: Brachyury-IRES-H2B-mCherry	Liu laboratory	Lolas et al., 2014
Mouse: Sox2-eGFP	The Jackson Laboratory	JAX: 017592
Mouse: Foxa2-eGFP-CreERT2	RIKEN CLST	Imuta et al., 2013
Mouse: Gal-eGFP	Mutant Mouse Resource & Research Centers	MMRRC: 016342-UCD
Mouse: H2B-miRFP703	Rossant laboratory	Gu et al., 2018
Mouse: ROSA mT/mG	The Jackson Laboratory	JAX: 007576
Mouse: Nkx2.5-Cre	The Jackson Laboratory	JAX: 024637

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
KLB 2.0 block-based lossless image compression file format	This paper	Data S1B https://doi.org/10.17632/fky6dkhmb3.1 https://bitbucket.org/fernandoamat/keller-lab-block-filetype/downloads/
Block-matching multi-view/time-lapse image registration pipeline	This paper	Data S1C https://doi.org/10.17632/fky6dkhmb3.1 https://gitlab.inria.fr/greg/Klab-BlockMatching https://github.com/leoguignard/Time-registration
GPU-based multi-view image deconvolution software	This paper	Data S1D https://doi.org/10.17632/fky6dkhmb3.1 https://bitbucket.org/fernandoamat/multiview-deconvolution
Convolutional neural network for division detection	This paper	Data S1E https://doi.org/10.17632/fky6dkhmb3.1 https://github.com/rueberger/division_detection https://github.com/kristinbranson/DivisionDetectionPostProcessing
TGMM 2.0 cell tracking framework with machine-learning module	This paper	Data S1F https://doi.org/10.17632/fky6dkhmb3.1 https://bitbucket.org/fernandoamat/tgmm-paper/overview
Computational framework for Statistical Vector Flow (SVF) analysis	This paper	Data S1G https://doi.org/10.17632/fky6dkhmb3.1 https://github.com/leoguignard/SVF
Converter for interactive 4D visualization of Statistical Vector Flow (SVF) databases	This paper	Data S1H https://doi.org/10.17632/fky6dkhmb3.1 https://github.com/leoguignard/SVF2MaMuT
TARDIS framework for spatiotemporal registration of multiple embryos	This paper	Data S1I https://doi.org/10.17632/fky6dkhmb3.1 https://github.com/leoguignard/TARDIS
Average embryo database query using external label masks	This paper	Data S1J https://doi.org/10.17632/fky6dkhmb3.1 https://github.com/leoguignard/I2AE
Preconfigured modules and example data for the practical software guide (full version)	This paper	https://doi.org/10.25378/janelia.6252800 https://figshare.com/s/48d3843efc538ac8f043
Preconfigured modules and example data for the practical software guide (compact version)	This paper	https://doi.org/10.25378/janelia.6252938 https://figshare.com/s/37776627713bf4408a72

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Philipp J. Keller (kellerp@janelia.hhmi.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Transgenic mice and reporters

Mouse lines used in this study were obtained from the following sources: H2B-eGFP and H2B-mCherry (RIKEN, Abe et al., 2011), R26-CAG-nuc-3xmKate2-nls (RIKEN, Susaki et al., 2015), CAG-tdTomato-2A-H2B-EGFP (CAGTAG1) mice (Trichas et al., 2008), Brachyury-IRES-H2B-mCherry (Lolas et al., 2014), Sox2-eGFP (JAX #017592), Foxa2-eGFP (RIKEN, Imuta et al., 2013), Gal-eGFP (MMRRC #016342-UCD), ROSA mT/mG (JAX #007576), Nkx2.5-Cre (JAX #024637), and H2B-miRFP703 (Gu et al., 2018). Mice were maintained on a mixed background of 129S, C57BL/6J and in the case of mKate2-nls CD-1 as well. Time pregnant females were obtained from natural matings by crossing stud males of breeding age with CD-1 females 2-3 months of age or a second reporter mouse of breeding age. Presence of a copulatory plug was denoted as day 0.5 d.p.c.

Sample preparation, embryo culture and imaging of mouse embryos

On the day of the experiment (typically E6.5 d.p.c. or E7.5 d.p.c.) time-pregnant females were sacrificed by cervical dislocation, embryos were dissected in media containing 5% FBS in Fluorobrite DMEM (ThermoFisher, Cat. no. A1896701 and 10082139) on a 37°C heated Tokai Hit stage (TPi-SZX2AX). The Reichert's membrane was removed with #5 forceps, taking care to leave the ectoplacental cone (EPC) intact. 3 mm glass capillary tubes (Hilgenberg, #1470253) were filled with Matrigel (GFR phenol-red free, Corning Cat. No. 356231) and allowed to solidify at 37°C. One end of the Matrigel-filled capillary was then sealed with dental wax (Darby Dental Supply, LLC, #0928094). Embryos selected for imaging were individually placed in capillaries by carefully embedding the EPC in the Matrigel column. Custom Teflon FEP tubing (Zeus) with an O.D. matching the I.D. of the glass capillary was then inserted into the capillary, dissection media was withdrawn with a fine-tip pipette and replaced with imaging media containing 40%-50% rat serum in

Fluorobrite DMEM (ThermoFisher, Cat. A1896701) supplemented with Pen-Strep (ThermoFischer, Cat. 15140122), GlutaMax (ThermoFischer, Cat. 35050061), and MEM non-essential amino acids (ThermoFischer, Cat. 11140076). Rat serum was purchased from Envigo (Whole Embryo Culture serum) or obtained from male Long-Evans rats (Charles River, Strain 006) 4-6 months of age.

Mounted embryos were then placed directly into the environmentally controlled "Maus Haus" (see full description in section "*Adaptive light-sheet microscope for imaging post-implantation mouse development*" below). 5% CO2 and 5% O2 were supplied by a customized Okolab Bold line system which also maintained the Haus temperature at a constant 37°C for the duration of the experiment. Embryos were imaged during developmental stages E6.0 d.p.c. to E8.5 d.p.c. every 5 min with a z-step size of 2.031 μm, a laser power of ~150-250 μW, and a laser sweep time of 20 ms per image.

Embryos were staged according to the criteria outlined by Downs and Davies, Kaufman's original Atlas of Mouse Development, the revised staging system by Lawson and Wilson, and the online eMouseAtlas (eMouseAtlas Community Resource).

METHOD DETAILS

Adaptive light-sheet microscope for imaging post-implantation mouse development

Design principles of the adaptive light-sheet microscope and mouse imaging assay

In this section, we elaborate on the technical advances and design principles introduced in the light-sheet microscope described in the Results section "*Adaptive multi-view light-sheet microscope for imaging mouse development*".

For optimal mouse embryo culture conditions during imaging, we fabricated a custom incubation enclosure and sample chamber that, coupled with an Okolab Bold line environmental system, provides highly stable temperature and atmospheric control (Figure 1A; Figure S1A). Embryos are cultured and imaged within this enclosure inside a sample chamber fabricated from heat or chemical-sterilization resistant materials (PEEK or Ultem). To mitigate the impact of light scattering and optical aberrations on image quality while facilitating rapid imaging of the entire embryo, the specimen is simultaneously illuminated with scanned light sheets (Keller et al., 2008) from two opposing directions while images are acquired with confocal slit detection, using two opposing detection objectives. Dipping objectives are used for both illumination and detection, and both sample post and objectives are enclosed by molded silicone seals (Figure 1B). We opted for the use of short working distance dipping optics to avoid unnecessary optical interfaces and minimize the optical path length both for light-sheet illumination and fluorescence detection, which in turn minimizes the effect of light scattering in the serum even when using an open embryo culture. The presence of four dipping objectives in one space severely restricts the available geometry for objective pairs (Figure 1B) and a system design based exclusively on commercial objectives would either prohibit the use of high numerical aperture detection objectives or require the use of long working distance illumination objectives. This restriction mandated the design of custom illumination objectives designed with a very small nose angle (33°), which reduce working distance to 4 mm (leaving enough space for embryo growth and sample access) and are fully compatible with detection objectives up to a numerical aperture of 1.0 (Figure S1E-S1H). Compared to the highest-quality, commercial illumination objectives with a long working distance, our low-profile custom design reduces light loss by a factor of 150 in a typical culturing medium containing 50% rat serum (Figure S2A). This performance feature and related improvements in image quality are the result of substantially reduced light scattering/absorption and aberrations due to the short illumination light path inside the medium.

Traditionally, mouse embryos have been grown in roller culture to ensure sufficient gas exchange and proper development (Piliszek et al., 2011); however, such conditions are obviously incompatible with imaging, and static culture methods have been developed that support normal post-implantation growth (Jones et al., 2002). We further adapted these static culture conditions for our system and ensure sufficient gas exchange by replacing the culture chamber atmosphere at a rate of at least 15 times per minute. The sample positioning stage is placed underneath the sample chamber, such that samples are mounted in a vertical position. While embryos like *Drosophila* and zebrafish may be mounted and held in place with agarose cylinders, the dramatic growth and sensitivity of the mouse embryo prohibits any kind of mechanical constraints. We use hollow glass capillaries filled with Matrigel as a supportive base for the embryo by gently embedding the sticky ectoplacental cone into the Matrigel, leaving the extra-embryonic and embryonic regions of the embryo above freely floating in media (Figure 1C). Embryos can then be placed either directly into low-volume chambers or enclosed in an ultra-thin Teflon FEP tube (25 μm wall thickness) with culture media and mounted in chambers filled with water (see sections "*Transgenic mice and reporters*" and "*Sample preparation, embryo culture and imaging of mouse embryos*" above).

Our adaptive imaging approach comprises several crucial advances over our earlier AutoPilot system (Royer et al., 2016). This latter method employed a static model of the specimen that was provided by the user and assumed a fixed, pre-defined specimen size and geometry. The user manually selected several reference locations in the specimen, at which the system monitored image quality, and the microscope then observed and reacted to optical changes only in these locations. We found this basic approach to be fundamentally ill-suited to imaging the developing post-implantation mouse embryo, which undergoes dramatic changes in size and shape and exhibits continuous movements throughout the time-lapse experiment. Both in non-adaptive light-sheet microscopy and in our original AutoPilot framework, the magnitude of detection defocus errors is on average at least as large as the depth of the focal volume of the detection objective itself (1.93 μm confocal parameter for Nikon 16x/0.8 and 1.14 μm for Zeiss 20x/1.0 objectives).

The advanced adaptive imaging framework described here facilitates a dynamic adaptation of the microscope configuration to local optical conditions, which serves the purpose of first optimizing and subsequently maintaining high spatial resolution across the embryo and over the time course of the experiment. As the embryo moves and/or grows, the adaptive imaging framework corrects the positions of the reference regions accordingly and adjusts their number and relative spacing to maintain adequate spatial sampling of

the constantly changing optical conditions across the developing mouse embryo (Figure 1D; Video S1B). To compensate for the optical heterogeneity and dramatic growth of the embryo, focal plane and light-sheet waist locations typically need to be adjusted by approximately 5 μm and 200 μm, respectively, over the course of 24 hr. Post-optimization, focal plane offsets are typically spread over a 5 μm range across the embryo in early developmental stages and over a 15 μm range during late developmental stages. Thus, adaptive imaging capabilities are critically needed to maintain high spatial resolution and the ability to track individual cells throughout development. The spatial sampling and temporal frequency of our aberration corrections are set to ensure that less than 2% of all defocus errors remain uncorrected. Thereby, the average defocus-induced mismatch between light sheets and detection focal planes is reduced to only 0.06 ± 0.04 μm, which is negligible compared to the confocal parameter of the detection optics and the size of the cellular structures imaged in the mouse embryo (Table S1; Figures S2B and S2C). We evaluated the quantitative impact of the adaptive imaging framework on image quality in developing mouse embryos and found that, on average, spatial resolution is increased by a factor of 3.3, signal strength by a factor of 2.1 and the cut-off radius in frequency space by 60% (measured across $n = 5$ different locations and developmental stages; Figure 1F). A continuous side-by-side comparison of adaptive vs. non-adaptive imaging of post-implantation mouse development over the course of a time-lapse imaging experiment is shown in Video S1C.

Complementing the core adaptive imaging capabilities described above, we also introduced an optional degree of freedom in the AutoPilot framework that enables compensation for spherical aberrations in fluorescence imaging. By jointly controlling the positions of objectives and tube lenses in the detection arms with motorized stages, this degree of freedom can adapt the microscope to spatially varying spherical aberrations. In general, such variability is expected as a function of imaging depth in the embryo if the average refractive indices in the embryo and in the culturing medium differ. We observed this effect experimentally and confirmed that adaptive system correction leads to a corresponding improvement in image quality (Figures S2D and S2E); however, we also found that the relative impact of these corrections is considerably smaller (by a factor of ~6) than that of the illumination and detection focus corrections discussed above. We concluded from these data that it is preferable to avoid the additional imaging time and light exposure of the specimen required for continuous spherical aberration corrections.

We note that the design of the adaptive light-sheet microscope supports simultaneous acquisition of opposing views as well as more complex multi-view imaging strategies, such as orthogonal four-view imaging, which can be achieved by sequentially acquiring two sets of opposing views and physically rotating the embryo by 90 degrees in between. Due to the rotation of the point-spread function relative to the embryo, such orthogonal sets of views contain complementary frequency content that can be combined by multi-view registration and multi-view deconvolution. However, we found that the acquisition of two opposing views alone already offers good coverage of the mouse embryo at the cellular level, and we thus recommend four-view imaging only for short-term imaging of very crowded regions (e.g. migration of cells out of the primitive streak) or for investigations at the sub-cellular level, considering that this approach necessitates a two-fold increase in light exposure of the embryo.

Microscope layout and main system components

The light-sheet microscope we designed for live imaging of post-implantation mouse development (Figure S1A, Data S1A) consists of four main components: 1) Two bi-lateral scanned light-sheet illumination arms, 2) two wide-field fluorescent detection arms, 3) the environmental controls and Maus Haus incubation chamber and 4) the microscope control infrastructure. The illumination arms comprise an Omicron SOLE-6 multi-laser system with 488 nm, 515 nm, 561 nm, 594 nm, 642 nm, and 685 nm wavelengths with two exit ports and one QiOptiq kineFlex single-mode fiber for each illumination arm, each of which are connected to fiber-to-free-space collimators. The light path then travels through a filter wheel (96A351, driven by a MAC 6000 controller, Ludl) and shutter (Uniblitz LS6 with VMM-D3 three channel shutter driver), two relay lenses (49361-INK, Edmund Optics), one on either side of a dual-axis XY galvanometer scanning system (6125LH and RH from Cambridge Technology with dual-axis driver), before entering a second, identical XY galvanometer scanning system and then through an f-theta lens (S4LFT4375, Sill Optics), illumination tube lens (49361-INK, Edmund Optics) and a custom designed water-dipping 6.4× illumination objective (manufactured by Special Optics, model no. 54-12.5-31, see Figure S1E-S1H). Illumination objectives are mounted on 800 μm travel PIHera linear piezo stages (P-628.1CD and E-665 LVPZT-Amplifier/Servo controller, Physik Instrumente) coupled to custom made adjustable bases. Dipping objectives enter a custom fabricated heat and chemical sterilization resistant sample chamber made from black PEEK or Ultem and are held in place by custom molded silicone seals (Albright Technologies) to prevent leaks. Additionally, the sample chamber was designed with moats and overflow drains to protect sensitive electronics in the case of leaks or spills. The sample is held in place underneath the objectives by a custom stainless-steel sample holder, which is connected to a post on the stage assembly by a strong magnet. To ensure proper rotation this holder-to-post connection is guided by a combination of ball-bearings on the holder set into grooves notched out of the post. The stage assembly itself is comprised of three linear translation stages for XYZ positioning, and one rotation stage which is directly coupled to the sample post (M-116.DG and M111.2.DG with PI C-809.40 motion controller, Physik Instrumente). The fluorescent detection arms are comprised of water dipping objectives (either Nikon CFI LWD 16×/0.8W or Carl Zeiss Plan-Apochromat 20×/1.0W) mounted on 250 μm travel PIHera linear piezo stages (P-622.1CD and E-665 LVPZT-Amplifier/Servo controller, Physik Instrumente) coupled to custom adjustable bases. Following the light path, each detection arm furthermore consists of a detection filter wheel with internal shutter (96A354 with MAC 6000 modular controller, Ludl), a detection tube lens (either Nikon ITL200 or Carl Zeiss 425308-000-000) mounted on a 1500 μm linear travel piezo stage (PI-629.1CD with E-665 LVPZT-Amplifier/Servo controller, Physik Instrumente), and an sCMOS camera (Orca Flash 4.0 v2, Hamamatsu).

The Maus Haus environmental system and incubation chamber comprises an Okolab Bold Line heater and atmospheric control providing a mixture of air, CO_2, N_2, or O_2 depending on the desired concentrations. The Maus Haus incubator was custom fabricated

from laser-cut acrylic panels that were designed for ease of access, sterility, temperature stability and a baffled air-exchange system to minimize vibrations. This enclosure covers the sample chamber, dipping and illumination objectives, and four-axis stage assembly. Typical light-sheet drift from ambient (21°C) to imaging temperatures (37°C) is generally around 50 μm, however the system is highly stable once it is aligned to imaging temperatures, even with repeated heating and cooling cycles.

In total, the microscope presented here uses 77 different types of custom designed mechanical components, which are combined into 28 types of multi-part assemblies (such as the Maus Haus incubator, sample chamber, sample positioning system, etc.). To enable the replication of this instrument, we provide a complete set of technical drawings for all parts and assemblies, together with a detailed parts list, as supplemental data (Data S1A).

Mouse light-sheet microscope electronics and control infrastructure

Image acquisition and the control of all microscope components is facilitated via a distributed system consisting of a computer workstation (for data acquisition and user interface) and a National Instruments PXI-8110 embedded controller (for real-time control of microscope components) inside of a PXI-1042Q chassis containing a PXI-7350 motion controller for the stage assembly, a PXI-8432/2 RS232 interface connected to the filter wheel controllers, and two PXI-6733 input/output modules integrated with two BNC-2110 connector blocks, which output analog and digital waveforms controlling individual microscope components and monitor camera states. The microscope control computer itself is directly connected to the two Orca Flash 4.0 v2 cameras by way of FireBird CameraLink frame grabbers (Active Silicon) and is based on a Colfax SX6750 workstation base platform with dual Intel Xenon E5-2687W CPUs, 256 GB of RAM, and an Intel RS2WG160 RAID controller with 12 TB of SSD storage, running Windows 7 Professional 64-bit. Custom control software is written in LabVIEW (National Instruments) and Java to operate the microscope, execute the adaptive imaging workflow and run the online specimen tracking module. The adaptive imaging control framework is based on the AutoPilot automated control system (Royer et al., 2016), which was extended in several ways to facilitate long-term live imaging of post-implantation mouse embryos at high spatiotemporal resolution. New functionality in the extended AutoPilot system includes (1) continuous measurement and stabilization of 3D position of the specimen, (2) continuous measurement of specimen size and adaptation of the imaging volume to dynamic size changes, (3) use of a dynamic geometrical model of the specimen for mapping optical properties and determining the corresponding optimal microscope parameters across the specimen volume, including the adaptive placement of reference planes and their position correction through integration with the online specimen tracking module, and (4) the design of a new hardware- and software-based AutoPilot degree of freedom for adaptive correction for spherical aberrations (enabled by the use of motorized tube lenses). The algorithms underlying modules (1)-(3), which were essential to enable the *in toto* imaging and image quality reported in this study (Videos S1B and S1C), are described in more detail in the next section.

Adaptive volume expansion, placement of reference planes and specimen tracking

In order to assess specimen movements as well as dynamic changes in specimen size and shape, the microscope control framework continuously computes and analyzes maximum-intensity projections of the imaging volume along multiple axes and in real time. The procedures described below are executed independently for each specimen view recorded by the microscope.

To facilitate specimen tracking and volume expansion along the imaging axis (*z*-axis), the imaging volume is projected onto the *z*-axis, producing the intensity profile $I(z)$, which is subjected to the following automated procedure after each volume acquisition:

1) The minimum and maximum intensity values, $I_{z,min}$ and $I_{z,max}$, of the *z*-axis intensity profile $I(z)$ are computed. From these values, a threshold $t = I_{z,min} + l \cdot (I_{z,max} - I_{z,min})$ is computed, using the adaptive threshold level $l \in [0, 1]$ ($l = 0.1$ for the experiments reported here). Using the *z*-axis intensity profile and threshold t, the *z* locations $z_{f,1}$ and $z_{f,2}$ along the profile are determined at the point in which the profile's intensity values cross this threshold first ($z_{f,1}$, rising edge of transition from background to foreground) and last ($z_{f,2}$, falling edge of transition from foreground to background), respectively.

2) The locations $z_{r,i}$ along the *z* axis of the current set of n_r reference planes used for adaptive imaging are compared to $z_{f,1}$ and $z_{f,2}$ to determine the distance $d_{r,1}$ between the first reference plane ($z_{r,1}$, lowest value of *z*) and $z_{f,1}$ and the distance $d_{r,2}$ between the last reference plane ($z_{r,n}$, highest value of *z*) and $z_{f,2}$. If this is the first time point of the time-lapse recording, the distances $d_{r,1}$ and $d_{r,2}$ are stored for use as reference values at later time points ($d_{r,1,\mathrm{ref}}$ and $d_{r,2,\mathrm{ref}}$).

3) The distances $d_{b,1}$ and $d_{b,2}$ between the foreground/background transition coordinates $z_{f,1}$ and $z_{f,2}$ and the boundaries $z_{b,1}$ and $z_{b,2}$ of the current imaging volume along the *z*-axis are determined. These distances $d_{b,1} = z_{f,1} - z_{b,1}$ and $d_{b,2} = z_{b,2} - z_{f,2}$ represent the size of the background margin at each end of the imaging volume and are subsequently normalized to the size of the imaging volume: $f_{b,1} = d_{b,1}/(z_{b,2} - z_{b,1})$ and $f_{b,2} = d_{b,2}/(z_{b,2} - z_{b,1})$. If this is the first time point of the time-lapse recording, these fractions $f_{b,1}$ and $f_{b,2}$ are stored for use as reference values at later time points ($f_{b,1,\mathrm{ref}}$ and $f_{b,2,\mathrm{ref}}$). Irrespective of the initial configuration of the imaging volume, however, the control framework will not allow these fractions to fall below 0.05, to prevent loss of data in the event of rapid changes in specimen location or size. The only exception to this rule is the exhaustion of the physical *z*-range supported by the stage system of the microscope. In this latter scenario, the control software will suspend all attempts at adjusting the respective end of the specimen but will still try to keep the other end intact.

4) The current location of the center of the specimen along the *z*-axis is determined as $z_c = \sum_i z_i \cdot I(z_i) / \sum_i I(z_i)$, where z_i is the location of the image plane *i* in the imaging volume. If this is the first time point of the time-lapse recording, the location z_{center} is stored for use as a reference value at later time points $(z_{c,\mathrm{ref}})$.

5) The new size of the imaging volume is computed, and, if deemed appropriate, new image planes are added on either side of the current imaging volume. This decision is made by comparing the current values $f_{b,1}$ and $f_{b,2}$ to the respective stored reference values. If $f_{b,1}$ is smaller than $f_{b,1,\mathrm{ref}}$, then the size of the imaging volume is expanded by adding new image planes before the first plane to

increase the background margin to $f_{b,1,\mathrm{ref}}$. Likewise, if $f_{b,2}$ is smaller than $f_{b,2,\mathrm{ref}}$, then the size of the imaging volume is expanded by adding new image planes after the last plane to increase the background margin to $f_{b,2,\mathrm{ref}}$.

6) The current specimen drift z_Δ along the z-axis is measured as the distance between z_c and $z_{c,\mathrm{ref}}$.

7) The current average spacing s between reference planes used for adaptive imaging is computed, and the need for placement of new reference planes or repositioning of existing reference planes is evaluated as follows:

7a) If the current number of reference planes n_r is smaller than the maximum allowed number of reference planes n_{max}, the control framework determines whether the distance between the first reference plane located at $z_{r,1}$ (respectively, last reference plane located at $z_{r,n}$) and the location of the background/foreground boundary $z_{f,1}$ (respectively, $z_{f,2}$) exceeds $d_{b,1}+s$ (respectively, $d_{b,2}+s$). If so, then a new reference plane is added at the location $z_{r,1}-s$ (respectively, $z_{r,n}+s$). If conditions are met, it is possible that new reference planes are added on both ends of the imaging volume in this step.

7b) If the current number of reference planes n_r is equal to or larger than the maximum allowed number of reference planes n_{max}, the control framework shifts all reference planes such that the distance between the first reference plane and the volume boundary $z_{b,1}$ becomes equal to $d_{r,1,\mathrm{ref}}$, the distance between the last reference plane and the volume boundary $z_{b,2}$ becomes equal to $d_{r,2,\mathrm{ref}}$, and the ratio of distances between any two pairs of reference planes remains unchanged (i.e. the spatial arrangement of reference planes is uniformly rescaled to adapt to the new size of the imaging volume).

7c) The z-coordinates of all image planes and reference planes in the imaging volume are corrected to compensate for the specimen drift z_Δ.

To facilitate specimen tracking along the axes perpendicular to the imaging axis (x- and y-axis), the imaging volume is projected onto the x-y-plane, producing the maximum-intensity image $I_{x,y}$. We implemented and evaluated several different algorithms for converting the information encoded in $I_{x,y}$ into an estimated movement vector of the specimen: center-of-mass global (COM global), center-of-mass filtered (COM filtered), foreground centered (FG centered), foreground negative edge (FG –edge), foreground positive edge (FG +edge), and orthogonal edge tracking modes.

The COM global algorithm first performs a background subtraction of $I_{x,y}$ and then determines the x- and y-coordinates of the fluorescence center-of-mass. The center-of-mass determined at the first time point serves as a reference position and the specimen is moved after each volume acquisition to compensate for the displacement between the current center-of-mass and this reference position. This is the primary mode of specimen tracking along the x- and y-dimensions used in the mouse live imaging experiments presented in this study.

All other algorithms are incorporated into the imaging workflow via the same automated procedure after each volume acquisition:

1) The image $I_{x,y}$ is background corrected, thresholded according to a manually set threshold level and filtered to retain only 8-connected foreground objects with a minimum pixel count of c_{min} (set to 5,000 when imaging mouse embryos).

2) If the largest connected object is at least three times larger than the second-largest object, all other foreground objects are deleted; otherwise, they are kept. The former situation is typically encountered when an embryo expressing ubiquitous fluorescent labels is surrounded by auto-fluorescent background objects. The latter situation is typically encountered when using sparse fluorescent markers that "fragment" the foreground occupied by the embryo upon thresholding.

3) For FG –edge and FG +edge tracking modes, if more than one connected foreground object remains that meets the criteria defined in step (2), the object closest to the designated edge (see below) is selected as the reference object for further analysis. For FG center and COM filtered tracking modes, if more than one connected foreground object remains, only the largest object is kept for further analysis.

4) The final computation performed with the remaining foreground object(s) is then specific to each mode:

4a) COM filtered: x- and y-coordinates of the fluorescence center-of-mass of the image region in $I_{x,y}$ corresponding to the largest foreground object are determined. Coordinates at the first time point serve as a reference and drift correction is performed by compensating for the displacement between the current COM coordinates and reference COM coordinates.

4b) FG centered: the geometrical center of the largest foreground object is computed and is kept centered within the field-of-view.

4c) FG –edge: the $-x$ or $-y$ boundary of the reference object selected in step (3) is kept at a constant distance from the edge of the field-of-view.

4d) FG +edge: the $+x$ or $+y$ boundary of the reference object selected in step (3) is kept at a constant distance from the edge of the field-of-view.

4e) Orthogonal edge: when defining x (y) as the reference dimension, the algorithm identifies the first or last x (y) coordinate at which the largest foreground object has a y-length (x-length) that exceeds 25% of its maximum y-length (x-length). The binary y-COM (x-COM) is computed at this coordinate and kept centered within the field-of-view. Moreover, the $+y$ or $-y$ ($+x$ or $-x$) boundary of this object is kept at a constant distance from the edge of the field-of-view. This mode is useful if a curved edge of the specimen demarks an important region in the specimen, if this edge is expected to remain easy to discern during the experiment, or if the specimen is expected to eventually grow to a size exceeding the size of the field-of-view.

Configuration of the framework for adaptive imaging

Please see Methods S1A for a detailed list of the parameter settings for adaptive imaging of mouse development used throughout this study. The most important parameters of the adaptive imaging framework relate to the spatial and temporal sampling of aberrations measurements. Spatial sampling is defined by the placement of reference planes across the embryo with a certain pairwise distance between planes, and temporal sampling is defined by an update frequency measured in units of time points of the time-lapse imaging

experiment. Higher spatial and temporal sampling has the potential to further reduce defocus errors, but it also exposes the sample to more light and requires a larger fraction of the data acquisition bandwidth of the microscope. We thus optimize sampling parameters based on our resolution requirements for cell tracking as follows:

- Spatial sampling of detection defocus aberrations (pertaining to degrees of freedom *D*/*I* of the adaptive imaging framework): We set the spacing between reference planes to the maximum value that still ensures that the average difference in focus values between neighboring planes does not exceed the confocal parameter of the detection objective. The confocal parameter of the detection objective used in most experiments in this study is 1.93 μm (Nikon 16x/0.8W). By using a spacing of 50 μm between reference planes at the onset of imaging, the average difference in focus values is reduced to 1.61 ± 0.39 μm in post-implantation mouse development (n = 20 time-lapse experiments). Note that this is not equivalent to the residual defocus error, since we are interpolating optimal parameter settings between focus planes (the residual defocus error is 0.06 ± 0.04 μm, Table S1). Rather, by reducing pairwise differences in the detection focal shift to a value below the confocal parameter, we ensure that in a worst-case scenario, where linear interpolation of focus settings in space does not fully recover locally optimal focus conditions, we still acquire mostly in-focus images.
- Temporal sampling of detection defocus aberrations (pertaining to degrees of freedom *D*/*I* of the adaptive imaging framework): In contrast to spatial variations in defocus aberrations across the embryo, which can be captured and largely compensated for by interpolating a set of spatially sparse measurements (Table S1), temporal changes are more difficult to predict (requiring extrapolation instead of interpolation) and thus demand more aggressive sampling criteria than the spatial sampling approach discussed above. By using a temporal sampling frequency of 5 time points between the acquisition of full complements of our detection defocus aberration measurements, we ensure that the maximum axial change in focus position (0.30 ± 0.15 μm, n = 20 time-lapse experiments) remains below the lateral system resolution (0.36 μm).
- Temporal sampling of optimal illumination beam waist position (pertaining to degree of freedom *Y* of the adaptive imaging framework): The position of light sheets along the illumination axis must be updated over time to adapt to the change in specimen size resulting from the continuous growth of the embryo (thus tracking the movement of the fluorescent tissues resulting from embryo expansion). We thus set the temporal frequency of illumination focus positions update to the minimum value needed to ensure that the embryo does not growth by more than 20% of the length of the light sheets between successive updates, thereby maintaining optimal optical sectioning and axial resolution. The post-implantation embryo grows at an average rate of 3.20 ± 0.76 μm per time point along each dimension (n = 19 time-lapse experiments, 1 time point = 5 min). By using a temporal sampling frequency of 11 time points between the measurement of full complements of optimal illumination beam waist positions, we reduce the spatial shift of beam waist along the illumination axis to 18% of the light sheet length (200 μm at an illumination NA of 0.045).

Limitations of the adaptive imaging framework

We would like to note three primary limitations of our imaging method. The first limitation concerns the maximum embryo size supported by the detection optics. Based on the working distance of our detection objectives, the maximum supported embryo size along the anteroposterior axis is 6 mm. The maximum supported size along the mediolateral axis is limited to 1 mm by the field-of-view of the detection system. Although these limits are not yet reached within the 48-hour imaging window reported in this study, they eventually become relevant when imaging at even later stages. If needed, these constrains can be partially overcome either by using optical tiling (which is supported by our control framework and extends the size limit to 8 mm) or in future work by using custom detection objectives, cameras with an even larger field-of-view and/or zoom optics that allow adapting the detection magnification during the experiment. The second limitation concerns the maximum speed with which the embryo's optical properties can be mapped by our adaptive imaging framework. Our framework is currently fast enough to perform all measurements and computations required to update the aberration correction model for the embryo once every 3 min, assuming that only 5% of the microscope acquisition bandwidth may be used for measuring the embryo's optical properties. While this is sufficient when employing our adaptive imaging method in a microscope with diffraction-limited resolution (the remaining defocus error arising from first- and second-order aberrations is 60 nm, as shown in Table S1), faster updates may be needed for super-resolution techniques that aim to deliver a spatial resolution on the order of a few hundred nanometers or less (considering the lower error-tolerance in this scenario). The third limitation is that light scattering and higher-order aberrations ultimately constrain the maximum depth at which high-resolution images can still be recorded (as discussed in the main text and shown e.g. in Figure 2B). We partially compensate for these effects through confocal line detection, which blocks out-of-focus light in the detection arm and reduces the negative impact on resolution and image contrast. A further reduction of light scattering is possible using near-infrared fluorescence reporters (Video S3B).

Large-scale image data management

Overview of data management

A typical image data set of mouse development recorded with the technique described in the Results consists of approximately 10 Terabytes of raw image data and includes multiple views of the embryo. To make all processing and analysis steps as efficient as possible, we apply three general pre-processing steps to our data. First, we detect the foreground regions in the image data,

discard pixels with no information content and convert the raw data to an improved version of our previously described KLB lossless compression format (Amat et al., 2015) (Data S1B). This treatment reduces data size by a factor of 20 and enables instantaneous access to any part of the image data due to the block-based architecture of our file format. Second, we register (see section "*Robust affine 3D registration of multi-view and time-lapse image data*" below; Data S1C) and combine the complementary information content of the multiple specimen views acquired by the microscope into a single image volume either by content-based fusion (Tomer et al., 2012) (when imaging two opposing views) or an improved version of our previously described multi-view deconvolution algorithm (Chhetri et al., 2015) (when imaging four orthogonal views, Video S2C) (Data S1D). Third, we register the image data in time. Since the mouse embryo is largely unconstrained physically during the imaging experiment and the online specimen tracking module applies position updates in a temporally discrete manner, the acquired images exhibit residual jitter of the specimen location. We thus utilize a block-matching image registration algorithm to stabilize the images in time and maximize accuracy of the automated cell tracking described in the next section (see "*Robust affine 3D registration of multi-view and time-lapse image data*" below; Data S1C).

Foreground detection and block-based lossless data compression

For efficient storage and processing of time-lapse image data sets of post-implantation mouse development, we perform automated adaptive foreground detection and selectively keep and compress only foreground regions in the image data using our block-based lossless compression file format KLB, as previously described (Amat et al., 2015). We improved the original KLB compression libraries and related infrastructure by eliminating bottlenecks relevant for efficient scaling to large data sets and improving cross-platform use (and specifically support of Linux-based operating systems) (Data S1B). This treatment reduces data size by approximately a factor of 20 (converting on average 10 TB of raw image data to 500 GB of KLB-compressed image data per experiment), without incurring a loss of information in the image foreground.

Large-scale data handling in the TGMM and multi-view deconvolution frameworks

The exceptionally large size of the image data (up to 20 GB per single-view image stack, 80 GB of multi-view data per time point, and typically 10 TB per time-lapse experiment) and the large number of cells (on average 20,000 per time point, with 500 time points per time-lapse experiment) observed in the mouse live imaging experiments present several challenges to the multi-view deconvolution (MVD) and TGMM-based cell tracking frameworks. Here, we briefly outline the modifications applied to the MVD and TGMM frameworks to enable their application to data of this size and complexity. The improvements in TGMM cell tracking performance and related algorithmic changes are described in the section "*TGMM 2.0 framework for cell nuclei segmentation and tracking*" below.

Modifications to the MVD framework related to large-scale image data handling

We used our GPU-based multi-view deconvolution pipeline (Chhetri et al., 2015) as a starting point and modified the code to enable processing of input data too large to fit into available system memory (Data S1D). The task of deconvolving the image data for one time point is subdivided along the output z-dimension according to a configuration parameter *blockz*. Blocks are then distributed in parallel to all available GPUs in the processing workstation. For each block, the same deconvolution procedure is used by loading the input views, transforming and cropping them into the output space, and then preforming Lucy-Richardson deconvolution. The cropping is performed such that a boundary region is retained for each output block, thus avoiding boundary artifacts. Sub-dividing the problem in this manner enables processing larger data sets and speeds up computations through the parallel use of multiple GPUs, but it also incurs a small performance overhead since the input views must be read from the disk multiple times. This overhead is kept to a minimum through the use of the block-based KLB file format.

Modifications to the TGMM framework related to large-scale image data handling

TGMM was optimized to reduce memory usage, extend the number of trackable objects, and optimize for large numbers of potential lineage trees (Data S1F). Memory usage was reduced by 66% by more aggressively evicting cached input volumes while processing the time series. The number of trackable objects was originally limited to 65,535, so that a 2-byte identifier could be used to enumerate objects. By changing the identifier to be a 4-byte value across the code base, the number of trackable objects was extended to 4 billion. Tracking more objects with this modified scheme exposed several scaling issues with the TGMM implementation, such as possible faulting of recursive binary tree operations when recursions required more than the allocated stack size. These scaling issues were systematically identified and corrected.

Robust affine 3D registration of multi-view and time-lapse image data

In this section, we describe the principles and mathematical procedure underlying our 3D registration of multi-view image data sets (comprising four orthogonal views of the sample, acquired sequentially as two sets of opposing views, with a 90-degree physical rotation of the sample in between) and time-lapse image data sets (Data S1C).

Principle of multi-view registration

Let $\mathcal{I}$ be a set of 3D intensity-based images, where $\mathcal{I}_1$ is the first and reference image, $\mathcal{I}_2$ is the image acquired by the respective opposite camera, $\mathcal{I}_3$ and $\mathcal{I}_4$ are images acquired by the same cameras for the specimen orientation rotated by 90 degrees. We furthermore have the masked image $\mathcal{I}_1^M$ (resp. $\mathcal{I}_2^M$, $\mathcal{I}_3^M$, $\mathcal{I}_4^M$) for view 1 (resp. 2, 3, 4), which was created from $\mathcal{I}_1$ (resp. $\mathcal{I}_2$, $\mathcal{I}_3$, $\mathcal{I}_4$) by masking off the low-contrast image regions that are located at the respective far side of the image volume (corresponding approximately to the half of the embryo facing away from the camera) and that are captured in higher quality in the respective opposite camera view. The goal of the operation described here is then to register $\mathcal{I}_{\{2,3,4\}}$ to the frame of $\mathcal{I}_1$. To do so, we first register views 3 and 4 to the reference view, since

these views share more high-contrast information with the reference image than view 2. We then assemble a rough fusion of views 3 and 4 registered to the frame of view 1 to use them as a reference for registering view 2.

We thus wish to compute the three affine transformations $\mathcal{A}^*_{\{2,3,4\}\leftarrow 1}$ that facilitate the registration of the images of the three views 2, 3 and 4 to the frame of the image of the first view.

Construction of the affine transforms for multi-view registration

The three affine transformations described above are the result of the composition of one rigid and two affine transformations. These three transformations are the incremental refinement of the final transformation and are computed using a block-matching pyramidal algorithm (Guignard et al., 2014; Ourselin et al., 2000). This scheme is comparable to the ICP method (Besl and Mckay, 1992), except that instead of points, iconic primitives are matched. This block-matching algorithm iteratively refines the final transformation by integrating δT^j, $\mathcal{T}^{j+1}_{1\leftarrow i} = \delta T^j \circ \mathcal{T}^j_{1\leftarrow i}$, $i \in \{2,3,4\}$. To compute δT^j, blocks (or supervoxels) of the reference image (here $\mathcal{I}_1$) are matched to blocks of the floating image (here $\mathcal{I}_{\{2,3,4\}}$). These blocks are then paired by maximizing the sum of the normalized correlation between them. From this block pairing, a point pairing (p_1, p_i) is extracted (from the center of each block). Then, δT^j is estimated as the transformation (rigid or affine in our case) that minimizes the trimmed square distance between these points (using Least Trimmed Squares by Rousseeuw and Leroy, described in "Robust Regression and Outlier Detection", which allows the discarding of outliers):

$$\delta T^j = \underset{\delta T}{\operatorname{argmin}} \| p_i - \delta T \circ \mathcal{T}^{j-1}_{1\leftarrow i} \|^2 \tag{1}$$

To determine $\mathcal{A}^*_{i\leftarrow 1}$, $i \in \{3,4\}$, we first compute a coarse affine transformation using the unmasked image data in order to include the full global information of the embryo (overall size and location of specimen boundaries). This transformation is initialized by a rotation of $\pm(\pi/2)$ rad and registers view i to a frame close to that of view 1 (which we denote as (1)). We denote this transformation as $\mathcal{A}_{i\leftarrow(1)}$. $\mathcal{A}_{i\leftarrow(1)}$ is then refined using the masked images ($\mathcal{I}^M_1$ and $\mathcal{I}^M_i$, i.e. considering only the high-resolution image content), by computing an affine transformation $(\mathcal{A}_{(1)\leftarrow 1})$ that only allows small, local changes. This transformation enables a refinement of the previous transformation. The composition of these two transformations then yields the final transformation $\mathcal{A}^*_{i\leftarrow 1} = \mathcal{A}_{i\leftarrow(1)} \circ \mathcal{A}_{(1)\leftarrow 1}$. We apply this transformation to $\mathcal{I}_i$ to create the registered image $\mathcal{I}_{i\rightarrow 1} = \mathcal{I}^M_i \circ \mathcal{A}^*_{i\leftarrow 1}$.

Once $\mathcal{A}^*_{3\leftarrow 1}$ and $\mathcal{A}^*_{4\leftarrow 1}$ have been computed, the two images $\mathcal{I}^M_3$ and $\mathcal{I}^M_4$ can be registered to the frame of the reference image $\mathcal{I}_1$. We then use this registration to create a coarse fusion of the two registered images $\mathcal{I}^M_3$ and $\mathcal{I}^M_4$ that we define as the voxel-wise maximum of the two registered images. We refer to this fused image as $\mathcal{I}_{\{3,4\}\rightarrow 1}$:

$$\mathcal{I}_{\{3,4\}\rightarrow 1}(v) = \max\left\{\mathcal{I}^M_3 \circ \mathcal{A}^*_{3\leftarrow 1}(v), \mathcal{I}^M_4 \circ \mathcal{A}^*_{4\leftarrow 1}(v)\right\} \tag{2}$$

We then compute $\mathcal{A}^*_{2\leftarrow 1}$ in the same way as the transformations for views 3 and 4 except that instead of using $\mathcal{I}_1$ and $\mathcal{I}^M_1$ as reference images, we use $\mathcal{I}_{\{3,4\}\rightarrow 1}$.

Time-lapse registration

In order to stabilize the time-lapse image data in time, we register the (fused) images from all time points to the frame of a reference time point. This time point is usually chosen as the time point that minimizes spatial distances relative to all other time points, such that transformations that map voxels across large spatial distances are avoided. This process thus registers an image $\mathcal{I}_{t_i}$ at any time point t_i to the frame of a reference image $\mathcal{I}_{t_{ref}}$. To compute the final transformation $\mathcal{T}_{t_i\leftarrow t_{ref}}$ we first compute all pairwise transformations for consecutive time points. When registering two consecutive time points t_i and t_{i+1} the choice of the reference image is determined according to whether $t_i < t_{ref}$ or not. If $t_i < t_{ref}$, we compute $\mathcal{T}_{t_{i+1}\leftarrow t_i}$, otherwise we compute $\mathcal{T}_{t_i\leftarrow t_{i+1}}$. Then, to determine the transformation $\mathcal{T}_{t_i\leftarrow t_{ref}}$ that registers $\mathcal{I}_{t_i}$ to the frame of $\mathcal{I}_{t_{ref}}$, we compose the respective consecutive transformations (shown here for the case $t_i < t_{ref}$):

$$\mathcal{T}_{t_i\leftarrow t_{ref}} = \mathcal{T}_{t_i\leftarrow t_{i+1}} \circ \ldots \circ \mathcal{T}_{t_{ref-1}\leftarrow t_{ref}} \tag{3}$$

This scheme was used for time-lapse registration of all embryos presented in this work (including TARDIS embryos B, C and D), with the exception of embryo A, which upon visual inspection of the raw image data exhibited such small residual motion that time-lapse registration was deemed unnecessary.

TGMM 2.0 framework for cell nuclei segmentation and tracking

Overview

We used our nuclei segmentation and tracking algorithm TGMM (Amat et al., 2014) as a starting point for the development of a method capable of cell tracking in our time-lapse imaging data of developing mouse embryos. TGMM 1.0 was originally developed for nuclei tracking in *Drosophila* and zebrafish embryos, where it offers a linkage accuracy of at least 96.0%. However, the size of the image volumes of these embryos are approximately 70-fold and 10-fold smaller, respectively, than those of the mouse embryos imaged in the present study. *Drosophila* and zebrafish embryos are also considerably less complex with respect to their growth dynamics and global changes in morphology, optical properties, and the heterogeneity in observed cell/nuclei shapes across the embryo. As far as the analysis of mouse development is concerned, the application space of TGMM 1.0 was limited to analyzing very short time sequences (less than 30 time points over a 2-hour period) of mouse development in early stages (E6.25) in which the size and optical complexity of the embryo are still comparable to those of a *Drosophila* embryo. The linkage accuracy of TGMM 1.0 in this scenario was reduced to 90.0% and accurate detection of cell division events proved impossible for all but the smallest and youngest

embryos. Owing to this poor performance and the scalability issues described in the section above, TGMM 1.0 cannot be used for automated cell tracking over meaningful time scales during the stages investigated in this manuscript. We addressed the latter issue (scalability) by rewriting the core of the TGMM framework, improving memory efficiency 3-fold and enabling the reconstruction of more complex cell lineage databases (see section "*Large-scale image data management*" above). Overcoming the former issue, however, cannot be resolved simply with better software engineering. Instead, the main limitation in TGMM accuracy had to be addressed: TGMM originally utilized heuristics for detecting cell divisions, which fail in advanced stages of mouse development when nuclear shapes can not only be relatively complex and diverse but change dramatically from one time point to the next. Thereby, cell divisions are falsely assigned to non-dividing nuclei and those nuclei become fragmented in TGMM object space or incorrectly linked to neighbors in time, which in turn causes over-segmentation of the image volume and reduces linkage accuracy.

We thus developed TGMM 2.0, which improves the algorithm by replacing the core heuristic rules for cell division detection in TGMM 1.0 with a machine learning strategy for cell division discrimination using image- and lineage-based spatiotemporal features (Data S1F). Overall, TGMM 2.0 reduces the number of false division detections in post-implantation mouse development by a factor of 12 while recovering 71% of all true divisions (albeit with limited precision, which triggered our development of the CNN-based division detector described below). In terms of *f*-score, overall division detection performance was improved by a factor of 3.1 (n = 2,083 manual cell division annotations). The performance improvements introduced in TGMM 2.0 also carry over to other model systems: in *Drosophila* and zebrafish embryos, the precision of cell division detection is improved on average from 0.34 to 0.74 (Methods S1B), without sacrificing recall rates or computational speed. As a result of the improved division classification in TGMM 2.0, false division detections are now much less frequent than linkage errors for post-implantation mouse development (1:5.4 ratio of division vs. linkage errors), whereas the situation was reversed in TGMM 1.0 (2.7:1 ratio of division vs. linkage errors). This in turn positively affects overall segmentation and linkage accuracy: TGMM 2.0 reduces over-segmentation by a factor of 2.3 (n = 30,962 manual nuclei annotations) and yields a linkage accuracy of 93.0%-94.8% (depending on the fluorescent marker strategy) from gastrulation to early organogenesis (Figure S4B). The linkage accuracy (or tracking accuracy) reports the probability of correctly linking a nucleus at time t to the same nucleus at time $t + 1$, whereas the segmentation accuracy reports the probability of correctly detecting a nucleus in the image data, following the definitions in (Amat et al., 2014). Manual cell and lineage annotations were performed using the Fiji plugin Massive Multi-view Tracker (or MaMuT) (Figures S3A–S3D) (Wolff et al., 2018). We observed higher segmentation and tracking accuracy for the mKate2-nls fluorescent marker compared to the H2B-eGFP fluorescent marker, due to the higher image quality achieved with this very bright, far-red reporter. Despite the improvements summarized here, it should also be noted that the accurate reconstruction of cell tracks remains the primary strength and application of TGMM 2.0. Owing to the limited precision of cell division detection in post-implantation mouse development, we do not recommend using TGMM 2.0 as a tool for the identification of cell divisions *per se*. For the latter task, considerably higher precision and recall rates are obtained with our convolutional neural network designed specifically for cell division detection (see section "*Convolutional neural network for automated detection of cell divisions*" below).

Summary of TGMM 2.0 design considerations

As discussed above, reducing false positives (without increasing false negatives) during cell division detection is expected to provide the largest impact on TGMM's overall accuracy and subsequent manual data curation time (if error-free lineage data are required). Keeping the false negative cell division rate very low was in fact a design choice in TGMM 1.0 to ensure faster data curation: in general, it is more difficult and time-consuming to find and add a new cell division than to remove an existing, incorrect linkage. However, allowing a large number of false positives also degrades overall data quality and limits the usefulness of the resulting data for fully automated analyses. With the improved division detector of TGMM 2.0, false positives are reduced by more than an order of magnitude (in data sets of mouse development) without sacrificing recall, computation speed or generalization of the method to multiple data sets.

Our original pipeline TGMM 1.0 performed cell division detection by (1) creating a list of cell division candidates based on the number of connected super-voxels within each TGMM object, and then (2) pruning this list using contextual spatiotemporal heuristic rules. However, it is difficult for the human to determine optimal features and thresholds in high-dimensional spaces for a complex event such as a cell division. By contrast, machine learning approaches have proven very helpful for these types of tasks. Our first investigation to this end focused on generating a cell division classifier using 3D Haar-like image-based ellipsoidal features extracted at a single time point of a time-lapse image data set (Amat and Keller, 2013). Other methods combined image features across multiple time points using inference in learned discriminative graphical models to improve accuracy (Huh et al., 2011). However, inference in these models can be relatively slow. Inspired by supervised animal behavior classification (Kabra et al., 2013), we extended our original 3D Haar-like ellipsoidal features to 3D+*t* by generating new features based on the temporal statistics of each Haar-like region in consecutive time points within a time window encompassing a putative cell division event (see section "*Description of spatiotemporal features*" below for details). Thus, instead of using a temporal graphical model to find dependencies between features, we incorporate temporal information directly into the final feature vector to be able to use standard machine learning classifiers. This method allows for preserving the accuracy of rich spatiotemporal rules without sacrificing computation speed. Complementing these image-based features, we also incorporate lineage-based features, such as the distance between putative daughter nuclei and change in nuclear volume, to increase the accuracy of the classifier.

The new cell division detection module of TGMM 2.0 is accompanied by pre-trained classifiers for mouse, zebrafish and *Drosophila* embryos as well as a graphical user interface and training protocol to make it easier and more efficient for users to obtain training samples and build classifiers for their respective data sets. Following this protocol, users can adapt the classifier to their particular

image data within typically just a few hours by simply answering a set of visual yes/no questions, without the need for parameter tuning (see section "*Classifier parameters*" below as well as section 8.3 of the TGMM 2.0 user guide included in Data S1F, see file "\TGMM2.0\doc\TGMM_UserGuide.docx").

Description of spatiotemporal features

If a cell division candidate is detected by TGMM at time point t based on super-voxel connectivity within the ellipsoid (Amat et al., 2014), we extract the local lineage data for that object between time points $t - w$ and $t + w$, where w is defined by the parameter *temporalWindowForLogicalRules* in the TGMM configuration file. We then calculate for each centroid at each time point a vector of image-based features following the 3D Haar-like ellipsoidal features described in (Amat and Keller, 2013). For time points after t, we have two centroids per time point (one per putative daughter) and we use the mid-point of the line connecting the two centroids as the origin for the feature calculation. The size and orientation of the ellipsoid (the precision matrix W) are kept constant across all time points. Their values are calculated on-the-fly as the element-wise average of W for all objects tracked by the TGMM algorithm at time point t. These values thus only need to be calculated once per time point, even if (typically) thousands of cell division candidates need to be classified.

Aside from the image-based features, we also calculate the following lineage-based features for each time point:

- *Distance to cell division mid-plane*: We determine the plane that is equidistant to the centers-of-mass of the two putative daughter nuclei at the first time point after the putative division event. This feature is then defined as the distance of the center-of-mass of a nucleus to this plane.
- *Distance between putative daughters*: Euclidean distance between the centers-of-mass of the putative siblings.
- *Relative change in volume*: The percentage increase or decrease of voxel count in the segmentation volume of the same nucleus at two consecutive time points.
- *Relative change in volume of the putative daughters*: Same as above, but measured between the two putative daughters of a cell (i.e. the putative siblings) at the same time point, instead of for the same nucleus at consecutive time points.
- *Jaccard distance*: Jaccard distance between the voxel-wise segmentation of the same nucleus at two consecutive time points.
- *Offset Jaccard distance*: Same as above, but measured after translating voxels to match the centers-of-mass at the two time points.
- *Offset Jaccard distance between daughters*: Same as above, but measured between two daughters of a cell at the same time point, instead of the same nucleus at consecutive time points.
- *Centroid displacement*: Euclidean distance between the centers-of-mass of the same nucleus at two consecutive time points.

Once the above image- and lineage-based features have been extracted for each time point in the time interval centered on the putative cell division event, we generate the final set of features by computing temporal statistics for each of the above features. Thus, instead of using a temporal graphical model for identifying dependencies between features and performing inference (Huh et al., 2011), we incorporate the temporal information directly into the final feature vector. This approach allows us to use standard machine learning binary classifiers for the final decision, such as support vector machines or boosting. Specifically, we compute the following quantities:

- *Mean*: average values for three different time intervals: (1) The entire local lineage segment (from $t - w$ to $t + w$), (2) the pre-mitotic phase (from $t - w$ to t), and (3) the post-mitotic phase (from t to $t + w$).
- *Standard deviation*: Same as above, but computing the standard deviation.
- *Minimum value*: Same as above, but computing the minimum value.
- *Maximum value*: Same as above, but computing the maximum value.
- *Haar-like statistics*: For each of the features computed above, all possible linear combinations with weights $\{+1, 0, -1\}$. This can be understood as a gradient between different temporal segments.
- *Z-score*: The value of the feature at time point t after subtracting the mean over the entire local temporal window and normalizing by the standard deviation. Additional versions of this feature are computed by normalization using the pre- and post-mitotic phases, respectively.
- *Normalization using neighbors*: The value of the feature at time point t divided by any of the first four features described above. These normalizations serve the purpose of capturing relative changes over time within the local temporal window using a different normalization scheme than Z-scoring.

Since we do not know *a priori* the optimal size of the local temporal window used to calculate these statistics, we recursively reduce the value of *temporalWindowForLogicalRules* by a factor of two until the temporal window is smaller than five time points. This approach allows efficient calculation of features at different temporal scales. For example, if the user sets *temporalWindowForLogicalRules* initially to 5 (the value used for data analysis in this study), we calculate features for temporal windows of sizes 11 and 5. This particular configuration results in a total of 2,034 features for training the cell division classifier.

We also tested two alternative approaches to calculating centroids and ellipsoid sizes for the 3D Haar-like ellipsoidal features. First, we tested using a variable corresponding to the precision matrix W returned by the TGMM algorithm for each object. Second, we tested having two feature vectors at each time point after time point t (one for each putative daughter) and concatenating them in time. However, both modifications performed worse than the methodology described above, because the temporal statistics cannot

discriminate if the variability across time for each 3D Haar-like elliptical feature is due to a true cell division or due to a variation in the ellipsoid returned by the tracking algorithm.

Classifier parameters

We trained three cell division detectors for mouse, zebrafish and *Drosophila* embryonic development, respectively. Training and classification were performed as follows:

Classifier for mouse embryos: Three rounds of annotations were performed for a total of 1,795 training samples (1,332 negative, 463 positive). Training samples were selected from time points 0, 20, 100, 120, 280, 330, 350, 470 and 500, and were different from the samples used during testing. Following the protocol accompanying our training GUI (see below), we set the cell division mid-plane distance threshold to 12.4 and the threshold for the cell division classifier to 0.456. The classifier was trained with 100 weak classifiers, using a learning rate of 0.1 and 30 minimum elements per leaf.

Classifier for zebrafish embryos: Three rounds of annotations were performed for a total of 1,172 training samples (900 negative, 272 positive). Training samples were selected from time points 50, 85, 150, 185, 250, 285, 350, 385, 450, 485, 550, 585, 650 and 685 of the SiMView zebrafish time-lapse data set presented in (Amat et al., 2014), and were different from the samples used during testing. Following the protocol accompanying our training GUI (see below), we set the cell division mid-plane distance threshold to 8.0 and the threshold for the cell division classifier to 0.35. The classifier was trained with 100 weak classifiers, using a learning rate of 0.1 and 30 minimum elements per leaf. Please see Methods S1B for cell division detection accuracy in zebrafish embryos.

Classifier for Drosophila embryos: Three rounds of annotations were performed for a total of 1,719 training samples (1,437 negative, 282 positive). Training samples were selected from time points 15, 36, 55, 86, 106, 136, 156, 186, 206, 236, 256 and 286 of the SiMView *Drosophila* time-lapse data set presented in (Amat et al., 2014), and were different from the samples used during testing. Following the protocol accompanying our training GUI (see below), we set the cell division mid-plane distance threshold to 4.2 and the threshold for the cell division classifier to 0.25. The classifier was trained with 100 weak classifiers, using a learning rate of 0.1 and 30 minimum elements per leaf. Please see Methods S1B for cell division detection accuracy in *Drosophila* embryos.

TGMM parameter set used for reconstructing post-implantation mouse development

Based on the results of our TGMM parameter screen (Figures S4A and S4B), we used the parameter settings listed in Methods S1C to process all data sets presented in this study. Parameters defining the image background level and tau used for persistence-based clustering were individually adapted to each data set (as they depend on the choice of fluorescent markers) and were set to 25-100 and 10-15, respectively, across the data sets presented in this study.

Protocol for GUI-based training of TGMM 2.0 cell division classifiers

We included two Matlab graphical user interfaces (GUIs) in the TGMM 2.0 software repository to facilitate easy visual inspection of the tracking results and training of cell division classifiers. Detailed descriptions of these GUIs and the procedure of training new cell division classifiers are provided in sections 7 and 8 of the TGMM 2.0 user guide included in Data S1F (see file "\TGMM2.0\doc\TGMM_UserGuide.docx").

Long-term tracking of cell movements with Statistical Vector Flow (SVF)

Overview of the SVF method

The principle behind the SVF method is to (1) robustly estimate local cell displacement vectors from the raw TGMM cell tracking results by evaluating movements in a cell's direct neighborhood, and (2) use these displacement vectors to construct the statistical vector flow of cell movements as a function of time such that it recapitulates the true 4D trajectories of individual cells with a higher average precision than that provided by just the raw cell tracking data alone. However, instead of simply extracting a cell's neighborhood as a function of time and computing local movement statistics from this neighborhood, we first assess the level of neighborhood conservation over time and have this information enter the statistical computation in the form of weights that define the contribution of individual members of this neighborhood in determining local movement vectors (Figure 4A). If cells are direct neighbors over multiple consecutive time points, they influence each other's statistical movement estimates more than cells that reside in another cell's proximity for only a single time point. This scheme benefits substantially from TGMM's high segmentation accuracy (Table S1) and thus has the potential to further improve accuracy and, thereby, the overall impact of the SVF method; however, it also requires careful consideration as to what constitutes the optimal definition of a cell's neighborhood. Specifically, it requires that a cell's neighborhood should be uniquely defined, robust to small changes arising from tracking noise and it should also be independent of the variability in cell density observed across the mouse embryo as a function of space and time. These considerations prevented us from using classical methods, such as *k*-nearest neighbors or the Delaunay triangulation, and instead led us to choose the Gabriel graph for determining a cell's local neighborhood, which meets all criteria defined above. The SVF method first computes the global Gabriel graph from our TGMM solution, and then, starting at the last time point and working backwards in time, evaluates the level of neighborhood conservation in time for each cell, assigns the corresponding weights and then determines the median displacement vector as our most robust estimate of local cell movements (Figure 4A).

Principle and definitions

To segment and track cell nuclei in the post-implantation mouse embryo, we first applied the TGMM 2.0 algorithm described above. TGMM 2.0 performs well on average, providing 96.7%-97.6% segmentation accuracy and 93.0%-94.8% linkage accuracy across all annotated tissues (n = 8,982-30,962 annotated cell positions and linkages). However, the remaining linkage errors result in nearest

neighbor identity confusions and premature track terminations that compromise the ability to follow cells over long periods of time. To enable the tracking of cell movements over a period of 48 hr with undisrupted temporal coherence and small spatial errors, we developed a computational method that robustly estimates local displacement vectors from the raw processing results of TGMM 2.0 (Data S1G, Figure 4A). These estimated displacement vectors provide a statistical vector flow (**SVF**) of cell movements as a function of time that reconstructs the true 4D trajectories of individual cells with an average spatial precision of 34.2 μm across all annotated tissues (n = 285 fully annotated cell tracks), thus enabling the long-term tracking of individual cells and tissue boundaries with an average accuracy of 1-2 cell diameters.

To formally define **SVF** we introduce the following notations:

- Let $T \subset \mathbb{N}$ be the set of time points TGMM 2.0 has been applied to, with $t_b = 0$ being the first time point and $t_e = max(T)$ being the last time point of the time series.
- Let $\mathbf{TGMM}_t = \{p_i = (x, y, z)\}, t \in T$ be the set of positions of cells detected at time t by TGMM 2.0.
- Considering a cell c at a position $p_i \in \mathbf{TGMM}_t$, let $P_{t \to t-1}(p_i)$ be the position of this cell at the previous time point. We can extend this definition to positive and/or longer time differences: $P_{t \to t'}(p_i)$ is the position of cell c at time point t'. Note that $P_{t \to t'}(p_i)$ does not exist for every p_i in $\mathbf{TGMM}_t$.
- Let $\mathbf{VF}_t$ with $t \in T \setminus \{0\}$ be a function that maps a position p_i at time t to its corresponding displacement vector $\mathbf{v}_{p_i}$. $\mathbf{v}_{p_i}$ is defined as $\mathbf{v}_{p_i} = p_i - P_{t \to t-1}(p_i)$:

$$\mathbf{VF}_t : \mathbf{TGMM}_t \mapsto \mathbb{R}^3$$

$$p_i \to \mathbf{v}_{p_i} = p_i - P_{t \to t-1}(p_i)$$

Then, we define **SVF** as a set of functions $\{\mathbf{SVF}_t\}_{t \in T}$:

$$\mathbf{SVF}_t : \mathcal{D}_{\mathbf{SVF}_t} \subset \mathbb{R}^3 \mapsto \mathbb{R}^3 \tag{4}$$

$$p_i \to \overline{\mathbf{v}}_{p_i}$$

where $\mathcal{D}_{\mathbf{SVF}_t}$, the definition domain of $\mathbf{SVF}_t$ is defined using the image of $\mathbf{SVF}_{t+1}$, $\mathcal{I}m(\mathbf{SVF}_{t+1})$, recursively, from the last time point to the first, as follows:

$$\mathcal{D}_{\mathbf{SVF}_t} = \begin{cases} \mathbf{TGMM}_t & \text{if } t = t_e \\ \mathcal{I}m(\mathbf{SVF}_{t+1}) & \text{if } t < t_e \end{cases} \tag{5}$$

We then compute $\overline{\mathbf{v}}_{p_i} = \mathbf{SVF}_t(p_i)$ by first estimating the neighborhood $\mathcal{N}_t(p_i)$ of p_i in $\mathbf{TGMM}_t$, then weighting this neighborhood according to its consistency in time and by ultimately taking the weighted median vector of the set of TGMM vectors for this neighborhood:

$$\overline{\mathbf{v}}_{p_i} = \mathbf{VF}_t \left\{ \underset{p_j \in \mathcal{N}_t(p_i)}{\operatorname{argmin}} \sum_{p_k \in \mathcal{N}_t(p_i)} w_k \cdot \| \mathbf{VF}_t(p_j) - \mathbf{VF}_t(p_k) \|_2 \right\}, \tag{6}$$

where $||*||$ is the L2 norm of $*$. We chose the median vector over an average vector because it is more robust with respect to outliers and conserves edges (Astola et al., 1990; Liu, 2013). Next, we will define the neighborhood $\mathcal{N}_t(p_i)$ and the weights w_i.

Definition and computation of the neighborhood $\mathcal{N}_t(p_i)$ of a position p_i

Since we are assessing neighborhood conservation over time, the definition of these neighborhoods should be unique and robust to small changes arising from noise. Moreover, the definition should also be robust with respect to the cell density heterogeneity observed across the mouse embryo as a function of space and time. These considerations and constraints prevent us from using classical methods, such as k-nearest neighbors or the Delaunay triangulation. The neighborhood definition we chose to meet our requirements is a subgraph of the Delaunay triangulation, the Gabriel graph (Gabriel and Sokal, 1969). The Gabriel graph of a set of positions $\{p_i\}$ is a graph $\mathcal{GG} = (V, E)$, where $V = \{p_i\}$ (in our case $V = \mathbf{TGMM}_t$) is the set of nodes and $E \subset V \times V$ is the set of edges. E is built such that if (p_i, p_j) is an edge of the graph, then no vertex in V is included in the disk (or sphere in 3D) with diameter $[p_i, p_j]$ and center $(p_i - p_j)/2$:

$$E = \left\{ (p_i, p_j) \mid \nexists p_k \in V, \left\| \frac{p_i + p_j}{2} - p_k \right\|_2 < \frac{\| p_i - p_j \|_2}{2} \right\} \tag{7}$$

Unlike the Delaunay triangulation, the Gabriel graph of a set of points is unique. Moreover, as is the case for the Delaunay triangulation (and unlike the k-nearest neighbors), there are no *a priori* assumptions regarding the number of neighbors or the distance between neighbors. Ultimately, the Gabriel graph does not necessarily embed the convex hull of the set of points and it does not contain the long edges embedded by the Delaunay triangulation (Figure S4D). We define $\mathcal{GG}_t$ as the Gabriel graph of the set of points

$\mathbf{TGMM}_t$ and $\mathcal{N}_t(p_i)$ with $p_i \in \mathbf{TGMM}_t$ as the neighborhood of p_i in $\mathcal{GG}_t$. For simplicity, we will refer to the neighborhood at time t' of the cell with position p at time t: $\mathcal{N}_{t'}(P_{t\to t'}(p))$ as $\mathcal{N}_{t'}(p)$.

Definition of the weights and neighborhood conservation

To increase the predictive power of our statistical representative of a set of vectors, we weight each vector in the neighborhood according to a measure of neighborhood conservation over time. To do so, we first define three degrees of neighborhood conservation $D1$, $D2$, $D3$. If p_j at time t is in the first (respectively second, third) degree neighborhood of p_i at time t' then $p_j \in D1_{p_i}^{t'}$ (respectively $D2_{p_i}^{t'}$, $D3_{p_i}^{t'}$):

- 1st degree: Two cells are first-degree neighbors between time points t and t' if they are direct neighbors at time points t and t':

$$\forall (p_i, p_j) \in \mathbf{TGMM}_t^2, p_j \in \mathcal{N}_t(p_i) \text{ and } P_{t\to t'}(p_j) \in \mathcal{N}_{t'}(p_i) \Leftrightarrow p_j \in D1_{p_i}^{t'} \tag{8}$$

- 2nd degree: Two cells are second-degree neighbors between time points t and t' if they (1) are neighbors at time point t, (2) are not 1st degree neighbors and (3) share at least one common neighbor at time point t':

$$\forall (p_i, p_j) \in \mathbf{TGMM}_t^2, p_j \in \mathcal{N}_t(p_i) \text{ and } p_j \notin D1_{p_i}^{t'} \text{ and } \mathcal{N}_{t'}(p_j) \cap \mathcal{N}_{t'}(p_i) \neq \varnothing \Leftrightarrow p_j \in D2_{p_i}^{t'} \tag{9}$$

- 3rd degree: Two cells are third-degree neighbors between time points t and t' if they are neighbors at time point t and are neither 1st nor 2nd degree neighbors at time point t':

$$\forall (p_i, p_j) \in \mathbf{TGMM}_t^2, p_j \in \mathcal{N}_t(p_i) \text{ and } p_j \notin D1_{p_i}^{t'} \text{ and } p_j \notin D2_{p_i} \Leftrightarrow p_j \in D3_{p_i}^{t'} \tag{10}$$

Then, using these definitions, we compute a weight w_{ij} for each pair $(p_i, p_j), p_j \in \mathcal{N}_t(p_i)$. The score depends on the degree of conservation between time points t and $t+1$ and time points t and $t-1$. If the cells are first (respectively second, third) degree neighbors the weight w_{ij} is increased by w_1 (respectively w_2, w_3). For example, if $p_j \in D1_{p_i}^{t+1}, p_j \in D2_{p_i}^{t-1}$ then $w_{ij} = w_1 + w_2$. For all analyses presented here, we used $w_1 = 4$, $w_2 = 1$, $w_1 = 0$ based on the performance data obtained in our parameter screen (Figure S4C).

Complexity and computation

Gabriel graph: The Gabriel graph is a subset of the Delaunay triangulation. The Delaunay triangulation can be computed in $\mathcal{O}(n \log n)$ (n being the number of nodes) (Attali and Boissonnat, 2003). Since the Gabriel graph can be computed from the Delaunay triangulation in $\mathcal{O}(n)$ (Matula and Sokal, 1980), the complexity of the computation of the Gabriel graph is $\mathcal{O}(n \log n)$. The Gabriel graph is computed in parallel for all time points.

Median vector computation: The computation of the median vector involves the optimization described in Equation 6. The search space is defined by the neighborhood $\mathcal{N}_t(p_i)$, which in our case is relatively small (between 3 and 17 neighbors, and 7 neighbors on average). We therefore compute all possibilities and determine the position p_j that minimizes the cost function. Each neighborhood is evaluated independently and in parallel to speed up the computations.

Boundary case of cells entering the field of view

We defined the definition domain of $\mathbf{SVF}_t$, $\mathcal{D}_{\mathbf{SVF}_t}$, recursively, as the image of $\mathbf{SVF}_{t+1}$. This definition domain represents the locations at which $\mathbf{SVF}_t$ will be computed. The definition we provided to this end in Equation 5 prevents the computation of $\mathbf{SVF}_t$ for those cells that are not present (or whose progeny is not present) in the TGMM reconstruction at the last time point of the imaging experiment. It is possible that cells are not part of the TGMM solution of the last time point, e.g. if parts of the embryo grow out of the field of view or if cells move to a location inside the embryo that is so deep that it cannot be properly resolved with the light microscope. To allow the computation of displacement vectors for those cells that are not present in the TGMM solution computed for the last time point, we allow enriching $\mathcal{D}_{\mathbf{SVF}_t}$ by adding cell positions from $\mathbf{TGMM}_t$. Candidate cells to this end are identified by analyzing the average distance between the neighbors of a cell, using the Gabriel graph to define neighbors as described above. If the average distance between the neighbors of a cell $c \in \mathbf{TGMM}_t$ is 50% larger than the average distance between neighbors of that same cell in the scenario where this cell was included in $\mathbf{SVF}_t$, then we add this cell to $\mathcal{D}_{\mathbf{SVF}_t}$. We perform this operation before the computation of $\mathbf{SVF}$ for all $t \in T$.

Convolutional neural network for automated detection of cell divisions

Summary

We used deep learning to build a cell division detector capable of identifying the nuclei of dividing cells with high accuracy directly from the raw time-lapse image data of mouse embryonic development (Data S1E). To this end, we trained a 10-layer, 4-dimensional convolutional neural network (CNN), which predicted the presence or absence of a division at every voxel in the spatiotemporal volume. The network's one million parameters were optimized using stochastic gradient descent with the Adam optimizer (Kingma and Ba, 2014). Batch normalization (Ioffe and Szegedy, 2015) and gradient clipping were employed to stabilize training. Since there are many more voxels outside the neighborhoods of cell divisions than there are inside, there is a large class imbalance between the positive and negative classes. To overcome this class imbalance and encourage sampling of a diversity of hard negatives, we drew approximately half of the training samples in the neighborhood of a division and the other half uniformly at random. We additionally utilized voxel-wise loss reweighting as there was still significant class imbalance within each neighborhood. The network was

implemented with *Keras* (Chollet, GitHub) and *Tensorflow* (Abadi et al., 2016) and trained on a workstation using a TITAN X Pascal GPU. Finally, we filtered multiple detections of the same division using image processing techniques.

Our CNN-based method improves precision and recall of the cell division classification task 3.6-fold and 1.2-fold, respectively, compared to the machine-learning module of the TGMM 2.0 framework, yielding a precision of 0.81 and a recall of 0.81. Moreover, the CNN retrieves more than twice as many divisions (217%) from the image data than a human expert inspecting the image data at a rate of one image volume (i.e. one time point) per hour.

Network architecture

The architecture of our network is described in Methods S1D. All layers used the ReLU nonlinearity. There were 1,087,409 trainable parameters in total. Note that the network is fully convolutional and the output shapes shown in the table are a function of the filter and input sizes. The fully convolutional nature allows the network the be run on a volume of arbitrary size, while the prediction of each voxel depends on a contextual spatiotemporal window of 45×45 voxels in *xy,* 9 voxels in *z* and 7 time points.

All layers used *Keras'* 'Conv3D' layer type, even the dense layers 9-11. Expressing the dense layers as 1×1×1 convolutions kept the model fully convolutional and led to large efficiency gains when running the model over a volume. To construct a 4-dimensional network over space and time, we used 7 input channels to encode 7 time slices, effectively making the first convolutional layer 4D, while all subsequent convolutional layers are 3D.

Training procedure

We annotated the nuclei of all dividing and non-dividing cells in 11 image volumes and additionally created 2,083 pointwise division annotations across the entire time-lapse recording of a developing mouse embryo. We increased the robustness of our classifier by training set augmentation. We constructed virtual training samples on the fly by applying in-plane rotations around the z-axis and reflection across the xy plane to the original training samples. This corresponds to training set augmentation by the symmetry group $\boldsymbol{S}^1 \times \mathbb{Z}_2$. We trained our network by gradient descent with the Adam optimizer (Kingma and Ba, 2014), using an initial learning rate of 0.001. We employed batch normalization (Ioffe and Szegedy, 2015) and gradient clipping, clipping norms to 1 and values to 0.5, to stabilize training. Performance on a held-out validation set was used for early stopping of training and to manage the training rate. We halved the training rate if the validation loss plateaued for 5 epochs.

Overcoming class imbalance: Cell divisions are relatively rare in the image data. In a typical image volume with 3 billion voxels, which contains on the order of 20,000 cells, only on the order of 200 are dividing at any time. The labeled nuclei of these dividing cells occupy approximately one million voxels. If training samples are selected in a spatially uniform manner, the model learns the null strategy of never predicting divisions, which is correct 99.97% of the time. To overcome this, we drew approximately half of the training samples in the neighborhood of a division and the other half uniformly at random. We additionally utilized voxel-wise loss reweighting as there was still significant class imbalance within each sample. The inverse of the ground truth class probability, averaged over each batch, was used to reweight the loss.

Batch description: We exploited the fully convolutional nature of our model to efficiently train on spatially contiguous volumes larger than the contextual window of the model. That is, we choose volumes of a size such that the output of the model is a volume, not a single voxel, as compared to predicting each voxel one at a time. We refer to such chunked batches as volume batches. This strategy was used for fully annotated volumes, but could not be used for pointwise division annotations. We created a second batch type for pointwise annotations, and trained on single-voxel batches with probability 0.05 and volume batches with probability 0.95. We used a batch size of 128 for the single voxel batches. Volume batch samples had shape $(z, x, y) = (51, 81, 81)$ and a batch size of 2.

Post-processing

The output of the CNN division detector was a score for each time and location indicating the classifier's certainty that a division occurred then and there. The division classifier had positive predictions for a small window surrounding the true division. Because of this, we needed to integrate positive scores in small windows and suppress multiple detections within a window. We first filtered the volume with a Gaussian with covariance $diag[5, 5, 5, 2]$ to integrate positive predictions in a window. Then, we thresholded and suppressed multiple detections of the same division by finding local maxima and connected components. In order to make our analysis memory-efficient, we performed these operations on overlapping sub-spatiotemporal volumes and combined the results.

Analysis of orientation of cell divisions during neural tube morphogenesis

By mapping the cell division database onto the dynamic fate map in space and time, we achieved an average accuracy of 88% across all tissues in assigning cell type labels to each cell division (Table S1; $n = 2{,}806$ manually annotated cell divisions). To correlate division orientation with neural tube development we first mapped the progression of neural tube closure as a function of time by manually annotating the RC extent of the neural plate and the degree of bending by marking the dorsal, ventral, and medio-lateral edges (Figure 7C). Closure was denoted once the dorsal most edges of the neural folds fused together at the midline. From this we constructed a neural tube folding map that represents the degree of folding across the embryo in space and time (Figure 7D). We enriched the information provided by our division database for this analysis by manually annotating the locations of parent (at time *t*) and daughter nuclei (at time $t + 1$) in 3D space. From these annotations, performed for three different embryos, we determined the 3D orientation of cell divisions. Notably, more than 96% of divisions in all labeled tissues (neural tube, somitic mesoderm, and notochord) occurred along the tangential ML-RC plane (Figure 7E). We systematically annotated cell divisions across the entire extent of the rostral-caudal axis and for the entire period of neural tube elongation and closure, measuring the orientation of divisions relative to the RC (0°) and ML (90°) axes (Figure 7F).

TARDIS algorithm for spatiotemporal registration of multiple embryos

Overview of the TARDIS methods

To map one embryo onto another (the reference embryo), we first create sets of spatiotemporal landmarks for both embryos by manual image annotation, such as the locations of the anterior and posterior extent of each embryo, the notochord, heart field and anterior intestinal portal, node and node-streak boundary, the boundaries of the epiblast, lateral plate and neural tube, as well as the location and time of condensation of each somite (Methods S1E, Data S1Q). From these landmarks TARDIS then computes the actual transformation function, which comprises a spatial and a temporal component. Temporal registration is achieved by retrieving the temporal information encoded in the landmarks, aligning these time labels across the two embryos and interpolating between them to build a piece-wise linear temporal transformation map. Spatial registration comprises three main steps (Figure 5A). First, we perform a time-dependent alignment of the 3D orientation of the two embryos by using the spatial information encoded in the manual landmarks. Second, we improve local spatial correspondence of the two embryos by aligning anatomical landmarks along the anterior-posterior axis as a function of time, matching the relative anterior-posterior coordinates of the somites and the anterior intestinal portal. Third, we eliminate remaining geometrical mismatches arising from differences in embryo size and shape. For this final step, we build, as a function of time, a shell for each embryo that reflects the spatial distribution of cells segmented by TGMM. By examining the embryos from the perspective of a spherical coordinate system (with the geometrical centers of the embryos corresponding to the center of the coordinate system), we compute a radial deformation map that maps the shell of the first embryo onto the shell of the reference embryo. A ventral-view visualization of the individual stages of this procedure is shown as a time-lapse video in Video S6C.

To identify the optimal number of landmarks needed in the initial annotation step, we determined registration errors throughout the embryo as a function of landmark count (Figures S5C and S5D). Registration precision reaches a plateau at around 50 landmarks with an average registration error of 41.5 μm across the embryo. By doubling the landmark count, this error can only be slightly reduced to 36.6 μm (Figure S5D). For each embryo, we thus annotated on average 50 landmarks in intervals of 50 time points (which ensures that landmarks do not move by a distance larger than the registration error on average), creating in total 250-450 landmarks per embryo. Please see Methods S1E for a detailed list of landmark annotations included in the average embryo.

Principle and definitions

The purpose of the TARDIS framework (Data S1I) is to register, in space and time, the geometry of two embryos, a reference embryo $\mathcal{E}^{ref}$ and a floating embryo $\mathcal{E}^{flo}$. This registration is accomplished by building the transformation $\mathcal{T}_{ref \leftarrow flo}$ that maps spatiotemporal (4D) points from the geometry $\mathcal{E}^{flo}$ onto the geometry of the reference embryo $\mathcal{E}^{ref}$ (Figure 5A). $\mathcal{T}_{ref \leftarrow flo}$ has two main components, a spatial component, $\mathcal{S}_{ref \leftarrow flo}$, and a temporal component, $\tau_{ref \leftarrow flo}$. The spatial component is the composition of three main transformations:

$$\mathcal{T}_{ref \leftarrow flo} = \mathcal{S}_{ref \leftarrow flo} \circ \tau_{ref \leftarrow flo} \tag{11}$$

$$\mathcal{S}_{ref \leftarrow flo} = \mathcal{NL} \circ \mathcal{R}o_{(c, \overrightarrow{n})} \circ \mathcal{R}i \tag{12}$$

Note that $\mathcal{S}_{ref \leftarrow flo}$ and $\tau_{ref \leftarrow flo}$ are independent from each other and could therefore be applied in any order. In our case, we first build $\tau_{ref \leftarrow flo}$ and apply it to $\mathcal{E}^{flo}$, so that we can spatially compare the two embryos in order to deduce the remaining spatial transformations.

$\tau_{ref \leftarrow flo}$ is a piecewise linear transformation that maps the time frame of $\mathcal{E}^{flo}$ onto the time frame of $\mathcal{E}^{ref}$. $\mathcal{R}i$ is a rigid transformation (rotation and translation) that registers the spatial frame of the floating embryo onto the frame of the reference embryo. $\mathcal{R}o_{(c, \overrightarrow{n})}$ is a spatially variant rotation around axis $\overrightarrow{n}$ and relative to center c, where c is the barycenter of the reference embryo and $\overrightarrow{n}$ is the normal to the medial bilateral symmetry plane of the reference embryo. $\mathcal{R}o_{(c, \overrightarrow{n})}$ allows a coarse alignment of the geometry of the two embryos along the midline using the somites and the anterior intestinal portal as reference landmarks. $\mathcal{NL}$ is a final non-linear deformation that accounts for the differences in size and shape of the two embryos. $\mathcal{T}_{ref \leftarrow flo}$ is the composition of these four transformations (one temporal, three spatial) that incrementally refine the registration of the geometries of the two embryos in time and space.

To build $\mathcal{T}_{ref \leftarrow flo}$, TARDIS requires several types of spatiotemporal landmarks for both embryos. These landmarks are split into two classes. The first class contains landmarks that can be unambiguously matched between two different embryos (for example, the first somite on the left side). We call this first set L_P (for "paired landmarks"). The second class represents semi-paired landmarks, which belong to a specific tissue but do not necessarily reside in a well-defined (unique) spatial location within this tissue (for example, a location at the spatial boundary of the lateral plate mesoderm). We call this second set of landmarks L_G (for "geometrical landmarks"). L_P contains the positions of all somites at the earliest time point in which they have clearly condensed and broken off from the pre-somitic mesoderm. The positions of somites between somite condensation events are determined by linear interpolation.L_P also contains the coordinates of the anterior-most position of the notochord before reaching the anterior intestinal portal and the coordinates of the posterior-most position of the heard field (each for at least three time points across the time line of the recording). L_G contains sparsely mapped outlines of different tissues such as the epiblast, lateral plate mesoderm, paraxial mesoderm, heart field, and neural ectoderm. These tissues are also labeled at corresponding time points between the two embryos.

Definition and computation of τ

τ facilitates the temporal registration of $\mathcal{E}^{flo}$ onto the temporal frame of $\mathcal{E}^{ref}$, such that thereafter both embryos can be spatially examined side-by-side. Temporal registration is achieved by aligning sparse temporal landmarks and interpolating between them linearly. As landmarks we use the corresponding time points obtained by pairing the temporal components of geometrical landmarks L_G^{ref} and L_G^{flo} before condensation of the first somite. After condensation of the first somite, we use the consecutive somite condensation events for temporal alignment.

Definition and computation of $\mathcal{R}i$

Definition: $\mathcal{R}$i transforms $\mathcal{E}^{flo}$ into the spatial frame of $\mathcal{E}^{ref}$ by minimizing the sum of squared distances between tissue outlines identified in L_G^{ref} and L_G^{flo}. To determine $\mathcal{R}$i, we first need to pair the landmarks in L_G^{ref} and L_G^{flo}. The temporal component of this pairing is already covered by τ. The spatial component is determined by first applying a global coarse rigid transformation defined by the minimization of the sum of squared distances between the paired landmarks in L_P^{ref} and L_P^{flo}. Then, for each tissue type and for each set of paired time points, we find the pairing that minimizes the sum of squared distances between landmarks in L_G^{ref} and L_G^{flo}. Note that not all landmarks might have a counterpart in the respective other set of annotations. As the result of this procedure, we obtain a set of paired landmarks for each annotated time point t. Finally, we compute the rigid transformation ri_t that minimizes the sum of squared distances between the paired landmarks at t. $\mathcal{R}$i is defined as the piecewise linear interpolation in time between all transformations ri_t.

Computation: To find the optimal rigid transformations we use the method described in (Besl and Mckay, 1992). We decompose $\mathcal{R}$i into rotation and translation. The translation is given by the vector between the centroids of the two point clouds. The rotation is given by the singular value decomposition of the familiar covariance matrix between the centered landmarks H:

$$H = \sum_{i=1}^{N} \left(l_i^{ref} - c_{ref}\right) \cdot \left(l_i^{flo} - c_{flo}\right)^T \tag{13}$$

l_i^{ref} and l_i^{flo} are paired landmarks in L_G^{ref} and L_G^{flo}. Provided the singular value decomposition of $H = U\Sigma V^*$, the rotation is $ri_t = V^*U^T$. The temporal interpolation between the transformations ri_t is then computed by spherical linear interpolation (Slerp) (Shoemake, 1985). In order to find the optimal pairing between the geometrical landmarks, we use the Hungarian algorithm (Kuhn, 1955).

Definition and computation of $\mathcal{R}o_{(c,\overrightarrow{n})}$

$\mathcal{R}o_{(c,\overrightarrow{n})}$ serves the purpose of refining the spatial alignment of somites and anterior intestinal portal coordinates between the two embryos, using the sets of paired landmarks L_P^{ref} and L_P^{flo}. The positions of the landmarks are linearly interpolated in time, starting at the time of their first appearance. $\mathcal{R}o_{(c,\overrightarrow{n})}$ is then a rotation as a function of time and position of the points to rotate. This rotation is defined by a rotation center, axis and the angle of rotation. The center c is the barycenter of the embryo, which is defined manually at 10 time points across the data set and linearly interpolated between these time points. The axis of rotation is the normal of the symmetry plane P_s defined as the plane that, as a function of time, bisects the somites on the left and right side of the midline and contains c. The rotation angle of a given point, A, is determined by the position of A relative to its two surrounding landmarks $LM_{A,1}^{flo}$ and $LM_{A,2}^{flo}$, which are defined as the two landmarks closest to A towards anterior and posterior, respectively:

$$LM_{A,1}^{flo} = \underset{LM_i \in L_P^{flo}}{\operatorname{argmin}}\left(\angle LM_i cA\right) \tag{14}$$

$$LM_{A,2}^{flo} = \underset{LM_i \in L_P^{flo}}{\operatorname{argmin}}\left(\angle AcLM_i\right) \tag{15}$$

Using these two surrounding landmarks, we can compute the angles required for aligning the landmarks to their respective counterparts in the reference embryo. The two angles are $\alpha = \angle LM_{A,1}^{flo} cLM_{A,1}^{ref}$ and $\beta = \angle LM_{A,2}^{flo} cLM_{A,2}^{ref}$. From these two angles, the angle between $LM_{A,1}^{flo}$ and $LM_{A,2}^{flo}$, $\angle LM_{A,1}^{flo} cLM_{A,2}^{flo} = \gamma$ and the angle between $LM_{A,2}^{flo}$ and A, $\angle LM_{A,2}^{flo} cA = \zeta$, we can interpolate the rotation angle $\eta(A)$ that we need to apply to A (please see Figure S4E for a visualization of the geometry described in this paragraph):

$$\eta(A) = \beta + \zeta\frac{\alpha - \beta}{\gamma} \tag{16}$$

Definition and computation of $\mathcal{NL}$

The last transformation, $\mathcal{NL}$, serves the purpose of eliminating remaining geometrical mismatches arising from differences in embryo size and shape. The principles underlying $\mathcal{NL}$ take advantage of our knowledge that (1) $\mathcal{R}$i aligns the barycenters of the embryos and (2) in a spherical coordinate system with the coordinates (ρ, θ, ϕ), where ρ is the distance between the barycenter, $\mathcal{R}o_{(c,\overrightarrow{n})}$ accounts for deformation along θ. At this point, we then correct for differences along ρ. In this last step, a point $A = (\rho_A, \theta_A, \phi_A)$ is transformed in $\mathcal{NL}(A) = (\rho'_A, \theta_A, \phi_A)$, where only ρ_A was changed to ρ'_A.

To estimate this transformation we build the inner and outer shells of both embryos, $\mathcal{I}^{ref}$, $\mathcal{I}^{flo}$, $\mathcal{O}^{ref}$ and $\mathcal{O}^{flo}$ ($\mathcal{I}$ represents the inner shell, $\mathcal{O}$ the outer shell). From these shells, we build $\mathcal{NL}_{\mathcal{I},\mathcal{O}}$, which is defined for points that belong to $\mathcal{I}^{flo}$ or $\mathcal{O}^{flo}$:

$$\mathcal{NL}_{\mathcal{I},\mathcal{O}}\left(p_i^{flo}\right) = p_i^{ref} \tag{17}$$

Here, $p_i^{flo} = (\rho_i^{flo}, \theta_i, \phi_i)$ belongs to $\mathcal{I}^{flo}$ (respectively $\mathcal{O}^{flo}$) and $p_i^{ref} = (\rho_i^{ref}, \theta_i, \phi_i)$ belongs to $\mathcal{I}^{ref}$ (respectively $\mathcal{O}^{ref}$).

In order to compute $\mathcal{I}$ and $\mathcal{O}$ we use the distribution density of the points given by our segmentation algorithm TGMM. We discretize the space along ϕ and θ by sampling 200 values for ϕ in $[0, 2\pi]$ and, to obtain an evenly distributed discretization, sampling $1 + \phi(200 - 1/2\pi)$ values for θ in $[0, 2\pi]$. Then, for each pair (ϕ_i, θ_i) of our discretized space we need to find $\rho_i^{\mathcal{I}}$ and $\rho_i^{\mathcal{O}}$. To do so, we examine the distribution of points along the line L_i parametrized by the angles $\phi(t) = \phi_i$ and $\theta(t) = \theta_i$ and $\rho(t) = t$ for $t \in \mathbb{R}^+$, i.e. this is the line segment that starts from the center of our frame c at inclination ϕ_i and azimuth θ_i. We then compute the number of points n in TGMM within a distance D (here, we used $D = 60\ \mu m$) to the point $p_{i,\rho} = (\rho, \phi_i, \theta_i)$, as a function of the distance ρ to c:

$$n(\rho) = \mathrm{CARD}\left(\left\{p \mid p \in TGMM, \mathrm{proj}_{L_i}(p) = (\rho, \phi_i, \theta_i) \textbf{ and } \| p - p_{i,\rho} \|_2 < D\right\}\right) \tag{18}$$

Here, $\mathrm{proj}_{L_i}(p)$ is the orthogonal projection of p onto the line L_i. As an approximation, we define $\rho_i^{\mathcal{I}}$ as the $\mathrm{perc}_{\mathcal{I}}$ percentile of this distribution and $\rho_i^{\mathcal{O}}$ as the $\mathrm{perc}_{\mathcal{O}}$ percentile of this distribution (for our study we empirically chose $\mathrm{perc}_{\mathcal{I}} = 20^{\mathrm{th}}$ percentile and $\mathrm{perc}_{\mathcal{O}} = 80^{\mathrm{th}}$ percentile). Thereby, we have constructed a set of discretized points $\{\mathcal{O}_i\} \subset \mathcal{O}$ (respectively $\{\mathcal{I}_i\} \subset \mathcal{I}$). To compute $\mathcal{NL}(A)$, we extract the set of neighboring points of A, $\mathcal{N}_A^{\mathcal{O}} \subset \{\mathcal{O}_i^{flo}\}$ such that if a point p_i is in $\mathcal{N}_A^{\mathcal{O}}$ then p_i is a first-degree neighbor in the Gabriel graph of $\{\mathcal{O}_i^{flo}\}$ (see description of SVF) to one of the two closest points of A in $\{\mathcal{O}_i^{flo}\}$. We do the same for $\{\mathcal{I}_i^{flo}\}$. Then, $\mathcal{N}_A = \mathcal{N}_A^{\mathcal{O}} \cup \mathcal{N}_A^{\mathcal{I}}$. From these points, we define $\mathcal{NL}(A)$ as follows:

$$\mathcal{NL}(A) = A + \vec{v}_A$$

$$\vec{v}_A = \sum_{p_j \in \mathcal{N}_A} \| A - p_j \|_2 \sum_{p_j \in \mathcal{N}_A} \frac{\vec{v}_{p_j}}{\| A - p_j \|_2}$$

$$\vec{v}_{p_j} = p_j - \mathcal{NL}_{\mathcal{I},\mathcal{O}}(p_j) \tag{19}$$

Constructing an average embryo from multiple reconstructed embryos

Overview

We defined the average embryo as the average of the statistical vector flow (**SVF**) of a set of individual embryos reconstructed with **TGMM** and mapped into a common centroid frame (the average space). Previous work on the registration of human brain image data served as an inspiration for our strategy underlying the construction of an average transformation that registers individual embryos to the average space (Guimond et al., 2000). In this study, in total 4 **TGMM**-reconstructed embryos were used to build the average embryo. To define the centroid c of the contributing embryos, we used the centroid definition of a set of n-dimensional points $\mathcal{S} = \{p_i\}$:

$$c = \frac{1}{|\mathcal{S}|} \sum_{p_i \in \mathcal{S}} p_i \tag{20}$$

Using this definition, we derive the following property (Figure S4F):

$$\forall p_0 \in \mathcal{S}, \overrightarrow{p_0 c} = \frac{1}{|\mathcal{S}|} \sum_{p_i \in \mathcal{S}} \overrightarrow{p_0 p_i} \tag{21}$$

The points in the equation above represent the embryos, and the vectors represent transformations. For example, $\overrightarrow{p_0 p_i}$ is the transformation that registers the point/embryo p_0 onto the point/embryo p_i. Using this property, we can then utilize the embryo-to-embryo transformations computed with TARDIS to build the transformations from any embryo frame on the average frame as follows:

$$\overrightarrow{p_i c} = \overrightarrow{p_i p_0} + \overrightarrow{p_0 c} \tag{22}$$

We then applied these transformations to each embryo contributing to the construction of the average embryo. From the transformed embryos registered to the common average space, we build an average embryo for each time point by jointly analyzing the **TGMM** data of all embryos (accounting for the correct number and positions of cells in each data set). The average embryo is ultimately represented by an average statistical vector flow (**SVF**) data set, which is defined in the average space and built from the combined information of all transformed embryos.

Definition of the deformation field

The deformation field $\mathcal{T}_{ref \leftarrow flo}$ registers a position p^{flo} from a floating frame to its corresponding position p^{ref} in the reference frame. Specifically, $\mathcal{T}_{ref \leftarrow flo}$ is a continuous function (see below) that maps a 3D position p^{flo} to a 3D vector **v**, which corresponds to the displacement that needs to be applied to p^{flo} to transform it into the reference frame.

$\mathcal{T}_{ref \leftarrow flo}$ is based on a set of points $\mathcal{P}^{flo} = \{p_i^{flo}\}$ in the floating frame. A new position p_i^{ref} (which together form the set $\mathcal{P}^{ref}$) in the reference frame is associated with each point $p_i^{flo} \in \mathcal{P}^{flo}$ following the principles of the TARDIS registration method described above. This pairing between the floating and reference frames allows building a continuous function defined in $\mathbb{R}^3$:

$$\mathcal{T}_{ref \leftarrow flo}(p^{flo}) = \frac{1}{|\mathcal{N}(p^{flo})|} \sum_{p_i^{flo} \in \mathcal{N}(p^{flo})} \frac{p_i^{flo} - p_i^{ref}}{\| p^{flo} - p_i^{ref} \|_2} \tag{23}$$

Here, $\mathcal{N}(p^{flo})$ is the neighborhood of p^{flo} defined as the set of points in that are located at a distance of at most 20 μm (one cell diameter) of p^{flo} and $\|\cdot\|_2$ is the L2 norm.

Building deformation fields for registering individual embryos to the average space

To build a statistical vector field representing the average embryo, first all **TGMM** data are transformed to the average space (Figure S4F). For this step the transformations that register each respective embryo to the average space are required. For this purpose, we utilize the respective transformations to the reference embryo that can be computed using TARDIS as described above. Let $E = \{\mathcal{E}^j\}$ be our set of embryos and $\mathcal{E}^1$ the reference embryo. The following transformations are then available through application of TARDIS: $\mathcal{T}_{\mathcal{E}^1 \leftarrow \mathcal{E}^j}, \forall \mathcal{E}^j \in E$. From these transformations, we can compute the transformation $\mathcal{T}_{\mathcal{E}^{avg} \leftarrow \mathcal{E}^1}$ that registers embryo $\mathcal{E}^1$ to the average space:

$$\forall p_i^1 \in \mathcal{E}^1, \mathcal{T}_{\mathcal{E}^{avg} \leftarrow \mathcal{E}^1}(p_i^1) = \frac{1}{|E|} \sum_{\mathcal{E}^j \in E} \mathcal{T}_{\mathcal{E}^j \leftarrow \mathcal{E}^1}(p_i^1) \tag{24}$$

Here, $\mathcal{T}_{\mathcal{E}^j \leftarrow \mathcal{E}^1} = -\mathcal{T}_{\mathcal{E}^1 \leftarrow \mathcal{E}^j}$ and $\mathcal{T}_{\mathcal{E}^1 \leftarrow \mathcal{E}^1} = (0,0,0)^T$. The transformation from any embryo to the average space is then defined as the composition of the transformation from that embryo to the reference embryo $(\mathcal{E}^1)$ and the transformation from the reference embryo to the average space:

$$\mathcal{T}_{\mathcal{E}^{avg} \leftarrow \mathcal{E}^j} = \mathcal{T}_{\mathcal{E}^{avg} \leftarrow \mathcal{E}^1} \circ \mathcal{T}_{\mathcal{E}^1 \leftarrow \mathcal{E}^j} \tag{25}$$

The reference embryo is thus only used temporarily as an intermediate coordinate system and does not serve any purpose other than enabling the re-use of the TARDIS framework in facilitating the construction of the average embryo.

Building the average embryo in the average space

To construct the average statistical vector field (**SVF**), we first build an average **TGMM** embryo for each time point. For this purpose, the **TGMM** results of all embryos registered to the average space are used as statistical support. The average **TGMM** embryo, $\mathbf{TGMM}^{avg}$, is then used to compute the average **SVF** as described below. Two key constraints we wanted to impose on $\mathbf{TGMM}^{avg}$ are that cell number and local cell density in the average **TGMM** embryo should be identical to the average cell number and average local cell density of the embryos utilized to construct the average embryo.

Let $\mathbf{TGMM}^{i \rightarrow avg}$ be the **TGMM** result of registering embryo *i* to the average space. $\mathbf{TGMM}^{i \rightarrow avg}$ is a subset of the cells in $\mathbf{TGMM}^{i}$. To ensure that the average embryo does not violate basic geometrical constraints, only cells that had a minimum distance of 20 μm (one cell diameter) from the cells of all other embryos in the set *E* were transformed into the average space.

Then, to build $\mathbf{TGMM}^{avg}$ at a given time point *t*, we first determined the average number of cells, $\tilde{n}$, in all $\mathbf{TGMM}^{i \rightarrow avg}$. We randomly picked $\tilde{n}$ cells from the combined set of cells from all $\mathbf{TGMM}^{i \rightarrow avg}$ to create a first draft of the average embryo, $\mathbf{tgmm}_1^{avg}$. We built as many draft embryos $\mathbf{tgmm}_i^{avg}$ as there were original embryos, using this method, while ensuring that $\cap_i \mathbf{tgmm}_i^{avg} = \emptyset$. As a result, each draft embryo $\mathbf{tgmm}_i^{avg}$ thus had the desired number of cells $\tilde{n}$; however, since these cells were chosen randomly from the contributing embryos, local cell densities may not yet correspond to the average cell densities of the contributing embryos. To correct for this, we adjusted $\mathbf{tgmm}_1^{avg}$ with an iterative scheme that successively utilizes each of the $\mathbf{tgmm}_i^{avg}$. We first identified those cells in $\mathbf{tgmm}_i^{avg}$ that resided in locations where cell density in $\mathbf{tgmm}_1^{avg}$ was at least 25% lower than the average local density in the contributing embryos. We then added these cells from $\mathbf{tgmm}_i^{avg}$ to $\mathbf{tgmm}_1^{avg}$ while removing those cells in $\mathbf{tgmm}_1^{avg}$ that resided in overpopulated locations, starting with locations exhibiting the highest average local density. We repeated these steps by sequentially going through all $\mathbf{tgmm}_i^{avg}$ until local cell density values were fully corrected, resulting in the final solution $\mathbf{TGMM}^{avg}$.

Local average cell density at a position *p* was computed as the average of the local cell density for all contributing embryos. The local cell density $d^i(p)$ for an embryo $\mathbf{TGMM}^{i \rightarrow avg}$ was computed as follows:

$$d^i(p) = \frac{1}{n} \sum_{p_i \in \mathcal{N}_n(p)} p_i - p_2 \tag{26}$$

Overall, the resulting average embryo thus (1) has total cell counts as well as local cell densities that are identical to the respective average values measured across all contributing embryos and (2) is constructed through a procedure that recruits cells from the contributing registered embryos to the average embryo without selection bias.

Building the average statistical vector flow

The average statistical vector flow (**SVF**) was built as described in section "*Long-term tracking of cell movements with Statistical Vector Flow (SVF)*" above, using the information provided by $\mathbf{TGMM}^{avg}$. The only difference in the procedure used here is that, to build the statistically representative movement vectors, we considered the neighborhoods in all available embryos individually, so that each embryo provided a median vector. These vectors were averaged to compute the final vector used in the **SVF** solution. Finally, trajectories were smoothed in time using a Gaussian kernel with a sigma of 5 time points (25 min), as described above.

Probabilistic and statistical cell fate maps

Propagation of manual tissue labels

The first step in creating dynamic fate maps from the SVF cell tracking results is the manual, volumetric annotation of image data at the last time point t_e of the time-lapse data set. This annotation of the image data then allows us to assess for each spatial location (or cell) $p \in \mathcal{D}_{\mathbf{SVF}_{t_e}}$ which tissue it belongs to. We then propagated this information backward in time using our statistical vector flow data

(**SVF**). Let $\mathcal{U} = \{U_i\}$ be the set of tissues that have been annotated. Let us furthermore assume that if a location/cell $p \in \mathcal{D}_{\mathbf{SVF}_t}$ belongs to a tissue U_i then $p \in U_i$. We then define U_i as follows:

$$U_i = \{p \in \mathcal{D}_{\mathbf{SVF}_t},\ \forall t \in T \mid P_{t \to t_e}(p) \in U_i\} \tag{27}$$

Here, $P_{t \to t_e}$ is the SVF object at time t_e that is linked to position/cell $p \in \mathcal{D}_{\mathbf{SVF}_t}$ (as previously defined in the section describing **SVF**).

Cell fate density map

After propagating the tissue labels, we have not only the position of each cell as a function of time but also the corresponding tissue label, i.e. knowledge of the cell fate for each cell at each time point. From these we can compute, as a function of the location in the embryo, the density of cell fates with respect to each tissue. We define this density at time $t \in T$ for a given tissue U_i and a given volume $\mathcal{S}$ (a sphere with center O and radius r) as $D_{U_i,\mathcal{S},t}$. Let $C_{\mathcal{S}}$ be the set of coordinates belonging both to the sphere $\mathcal{S}$ and to $\mathcal{D}_{\mathbf{SVF}_t}$:

$$C_{\mathcal{S}} = \mathcal{D}_{\mathbf{SVF}_t} \cap U_i \tag{28}$$

Then:

$$D_{U_i,\mathcal{S},t} = \sum_{p \in C_{\mathcal{S}}} 1 - \frac{O - p_2}{r} \tag{29}$$

From this density, we can then compute the likelihood of a tissue U_i to be formed by the cells located within distance r of the position O in the embryo:

$$P(U_i \mid O, r, t) = \frac{D_{U_i,\mathcal{S},t}}{\sum_{U_j \in \mathcal{U}} D_{U_j,\mathcal{S},t}} \tag{30}$$

Visualizing the cell fate density map

To create a visualization of this map in space and time, we have to define the set of $\mathcal{S}$, for each time point, and thus define the set of $O_{\mathcal{S}}$ ($\mathcal{O}_t$) and the associated $r_{\mathcal{S}}$. We chose $\mathcal{O}_t = \mathcal{D}_{\mathbf{SVF}_t}$. The radii are defined as a function of the O they are associated with. To compute the radii, we first assume that $P(U_i \mid O, r, t) = 1$ for the five final time points of the time-lapse image data set (when tissues can be easily identified by eye and are manually annotated as described in the first step above). Therefore, assuming $p \in U_i$ (and by constructing U_i, $\forall t \in T$, $P_{t \to t_e}(p) \in U_i$):

$$r_p = \max(\{r \in \mathbb{R} \mid P(U_i \mid P_{t \to t_e}(p), r, t) = 1,\ \forall t \in [\![t_e - 5, t_e]\!]\}) \tag{31}$$

We then smooth these values in space and impose an upper bound so as not to over-estimate neighborhood size. To this end we define the upper bound as the median of $\mathcal{R} = \{r_p\}_{p \in \mathcal{D}_{\mathbf{SVF}_{t_e}}}$, which is the distribution of radii for the last time point as defined above. Spatial smoothing is achieved by taking the maximum r for the points in the volume around $p \in \mathcal{D}_{\mathbf{SVF}_t}$ defined by $\mathcal{S}_{p,r_{\min}}$, the sphere with center p and radius $r_{\min} = \min(\mathcal{R})$. The final radius r_p^* is therefore computed as follows:

$$r_p^* = \min\left(\overline{\mathcal{R}}, \max_{p_i \in \mathcal{S}_{p,r_{\min}}} \left(r_{P_{t \to t_e}(p_i)}\right)\right) \tag{32}$$

Using this definition, we can compute $P(U_i \mid p, r_p^*, t)$ for all $U_i \in \mathcal{U}$, $t \in T$, $p \in \mathcal{D}_{\mathbf{SVF}_t}$.

Computing a statistical cell fate map from multiple spatiotemporally registered embryos

To build a statistical cell fate map (visualized in Video S6E), we compute the average $\bar{P}(U_i \mid p, r_p^*, t)$, across all TARDIS-registered embryos as well as the reference embryo, of $P(U_i \mid p, r_p^*, t)$ using p and r_p^* from the reference embryo.

Computation of tissue morphodynamics maps

Outline of processing and visualization workflow

Since the embryo is normally curved in 3D space and the majority of cell movements occur within the relatively thin, curved geometry defined by the anterior-posterior and medio-lateral body axes, we adapted the visualization of our cell movement data to this geometry to make our data more intuitive and interpretable by eye. While the quantifications themselves are performed in 3D, we use Mercator projections to simplify the visualization of our tissue morphodynamics maps and flatten the embryo along a 2D plane (Figure 6E).

To build a tissue morphodynamics map that represent cell movements in a given tissue between time points t and $t + n$, we first perform a (global) rigid registration of the **SVF** data at time points t and $t + n$ in order to retain only the residual cell displacements between these time points (i.e. eliminating global embryo motion). We then project the 3D **SVF** data onto a 2D plane using a Mercator projection. From this point onwards, the tissues are treated individually. For each tissue, the continuous 2D space of the Mercator projection is downsampled to a grid (comprising 25 × 25 sectors in our case), and the average displacement and speed is computed for each sector of the grid. From this grid we then compute a streamline plot, which represents the final tissue morphodynamics map.

Rigid registration

Registration is performed by determining the rotation and translation ($\mathcal{R}^*$ for rigid) that minimizes the sum of squared distances between cell positions at time points t and $t + n$:

$$\mathcal{R}^* = \underset{\mathcal{R}}{\operatorname{argmin}} \sum_{p_i \in \mathcal{D}_{\mathbf{SVF}_t}} \| \mathcal{R}[P_{t \to t+n}(p_i)] - p_i \|_2^2 \tag{33}$$

Mercator projection

Given a point $p = (x, y, z)^T$ and the manually defined barycenter $b = (x_b, y_b, z_b)^T$ of the embryo, we compute the new coordinates of p ($p_M = (x_M^p, y_M^p)^T$) in the Mercator projection as follows:

$$p_M = \begin{pmatrix} x_M^p = 2 \cdot \tan^{-1}\left[\exp\left(\tan^{-1}\dfrac{y_b - y}{x_b - x}\right)\right] - \dfrac{\pi}{2} \\ y_M^p = \cos^{-1}\dfrac{z_b - z}{\| b - p \|_2} \end{pmatrix} \tag{34}$$

Computation of speed and displacement vectors in the projection grid

Using the Mercator-projected data, we can compute the average displacements in the downsampled space described above. Let g_i be a sector of our grid defined by its lower and upper bounds in the x and y dimensions of the Mercator projection: x_i, X_i, y_i, Y_i (note that for our grid, $0 < X_i - x_i = c_x$ and $0 < Y_i - y_i = c_y$, $\forall i$, i.e. the size of all sectors is the same, but length along x and y is not necessarily the same). Let g_{i,U_j} be the set of spatial locations (or cells) $p \in \mathbf{SVF}_t$ such that $p \in U_j$, $x_i < x_M \leq X_i$ and $y_i < y_M \leq Y_i$, corresponding to the cells that are positioning inside the respective sector and belong to the tissue U_j. We then compute the average displacement $(\delta_x^{i,U_j}, \delta_y^{i,U_j})$ and the average speed $\bar{v}^{i,U_j}$ for each sector of the grid and for each tissue:

$$\delta_x^{i,U_j} = \frac{1}{\left| g_{i,U_j} \right|} \sum_{p \in g_{i,U_j}} x_M^{\mathcal{R}^*[P_{t \to t+n}(p)]} - x_M^p \tag{35}$$

$$\delta_y^{i,U_j} = \frac{1}{\left| g_{i,U_j} \right|} \sum_{p \in g_{i,U_j}} y_M^{\mathcal{R}^*[P_{t \to t+n}(p)]} - y_M^p \tag{36}$$

$$\bar{v}^{i,U_j} = \frac{1}{\left| g_{i,U_j} \right|} \sum_{p \in g_{i,U_j}} \| \mathcal{R}^*[P_{t \to t+n}(p_i)] - p_i \|_2 \tag{37}$$

Note that the speed is intentionally computed in the original 3D coordinate system, since the Mercator projection does not preserve speed. We then compute, in the same way, the speed and displacement vectors for each tissue and each time point of the time-lapse data set. For visualization, the displacement vectors in the grid are represented using streamlines (see Fluid Mechanics by Granger).

A practical guide to the framework for analyzing image data of mouse development

Overview

In this section and in Methods S1F, we provide detailed instructions and example image data to demonstrate the practical application of our computational framework for image-based analysis of cell dynamics in developing mouse embryos (Figure 3A, Data S1B-S1J). In our step-by-step protocol (Methods S1F), we describe how to install, configure, and use each of our computational modules, how to provide input data and how to inspect and evaluate the results. We provide example image data, graphical user interfaces and preconfigured scripts for all software modules that allow the user to easily recapitulate all key analyses and computational capabilities presented in this paper, including time-lapse image registration, automated image segmentation with TGMM 2.0, automated cell tracking with TGMM 2.0, long-term cell flow tracking with SVF, backward propagation and dynamic fate mapping of any cell population of interest, registration of embryos and gene expression patterns to the dynamic atlas of mouse development with TARDIS, and automated detection of cell divisions. By following the instructions provided in this section and using our examples as teaching templates, the user can furthermore employ the software tools to process and analyze their own image data sets. The practical guide is accompanied by visual elements, referred to here as Guide Figure, which are provided at the end of the document "Guide.pdf" (included in the root folder of the guide's software and data repository). To start this walk-through, please download either the full version (26 GB; https://doi.org/10.17632/hy3y3z692g.1) or the compact version (11 GB; https://doi.org/10.17632/s24z62sbsf.1) of the guide's software and data repository, which contains all image data and the preconfigured software utilized in the step-by-step

protocol below. The respective download links are provided under the first bullet point of the section ''Setting up the software environment''. Once the software environment is set up, please read our step-by-step protocol for the data analysis framework provided in Methods S1F.

Setting up the software environment

Please download and unpack either the full version (26 GB; https://doi.org/10.17632/hy3y3z692g.1) or the compact version (11 GB; https://doi.org/10.17632/s24z62sbsf.1) of the guide's software and data repository at the Mendeley link: https://doi.org/10.17632/2589chvy9c.1. Both versions contain (a) 100 time points of a time-lapse image data set of mouse post-implantation development in the compressed KLB format, which is used to demonstrate time-lapse image registration, image segmentation, cell tracking, backward propagation, and dynamic cell fate mapping (data set A, folder *Images/DataSetA*; Guide Figure A, B), (b) 21 time points of a small sub-region of a time-lapse image data set of mouse post-implantation development in the compressed KLB format, which is used to demonstrate cell division detection (data set B, folder *Images/DataSetB*; Guide Figure C, D) and (c) all software modules, user interfaces, and preconfigured scripts needed to follow the step-by-step protocol below. The full version furthermore includes a copy of the processing results for all analyses described in the step-by-step guide (Methods S1F). Note that the compact version still allows the user to recapitulate all analyses but does not include a copy of the processing results themselves.

Our software modules for image registration, statistical vector flow (SVF), backward and forward propagation (dynamic fate mapping), SVF to MaMuT conversion (SVF2MaMuT), and mapping of 3D label masks to the average embryo database (I2AE) require a computer with a Linux distribution such as Ubuntu as well as Python 2.7. These modules are located in the folder *Python-Scripts* (and are also available from the GitHub repository https://github.com/leoguignard/standalone-Mouse). In addition, several third-party software packages must be installed. For a detailed description on how to install these packages, please follow our instructions provided in the *README.md* file located in the *Python-Scripts* folder. Alternatively, if you are using our software modules on Ubuntu, you have the option to automatically install all required software packages by executing the installation script *ubuntu-installation.sh* located in the *Python-Scripts* folder. The list of packages that must be installed to run our software modules is as follows:

- python-dev, cmake-curses-gui, pip and libhdf5 (Command: *sudo apt install python-dev pip libhdf5 cmake-curses-gui*)
- numpy, scipy, unittest2, matplotlib, PyWavelets (Command: *pip install numpy scipy unittest2 matplotlib PyWavelets [–user]*)
- pyklb (available from the GitHub repository https://github.com/bhoeckendorf/pyklb)
- TGMMlibraries, IO, BlockMatching, all of which are present in the *Python-Scripts* folder. Please refer to the respective *README.md* files for installation notes.

Our software module for cell tracking (Windows executable *TGMM.exe*) requires a CUDA-enabled graphics card (Maxwell architecture or higher).

Support for the .klb image file format is available as a plugin for the free image processing software Fiji that can be access through the Fiji update manager:

- To open .klb files in Fiji, please add the following update site to your Fiji installation: http://sites.imagej.net/SiMView
- Note that ImageJ2/Fiji requires SCIFIO to be enabled in order to be able to open .klb files by drag-and-drop. To ensure that SCIFIO is enabled go to Edit → Options → ImageJ2 → Check ''se SCIFIO when opening files''.

The software package BigDataViewer is available as a plugin for Fiji and used in the context of our image processing pipeline for the purpose of generating .xml files from .klb image series. These .xml files can then be viewed in MaMuT (see next item below):

- More information, installation instructions and a user guide to BigDataViewer are available here: https://imagej.net/BigDataViewer
- We provide an .xml file for the data included in the software/data repository (*TGMM/MousePipelineExample.xml*); however, in order to use this file on a different computer, the name of the folder containing the image data need to be changed to match the computer's local directory structure. To do so, open the respective .xml file and search for the string ''**'' (two asterisks). The search will then point you to the location in the .xml file where the file path adjustment is needed.
- Alternatively, to generate a new .xml file from a .klb image series in Fiji, follow these steps: Go to ''Plugins'', ''BigDataViewer'', and select ''Open KLB''. In the ''Template file'' field navigate to the last time point and color channel in the image series. Verify that the tags match the template file. Manually specify the pixel size. For example data set A provided in this software and data repository, the correct aspect ratio is obtained by entering X = 1.0, Y = 1.0 and Z = 6.25. Save the .xml file. This file can now be used to launch a new MaMuT annotation or as the image data location for importing TGMM results into MaMuT.

The software package MaMuT is available as a plugin for Fiji that can be accessed through the Fiji update manager. More information, installation instructions and a user guide to MaMuT are available on the following websites:

- https://imagej.net/MaMuT
- https://imagej.net/Getting_started_with_MaMuT

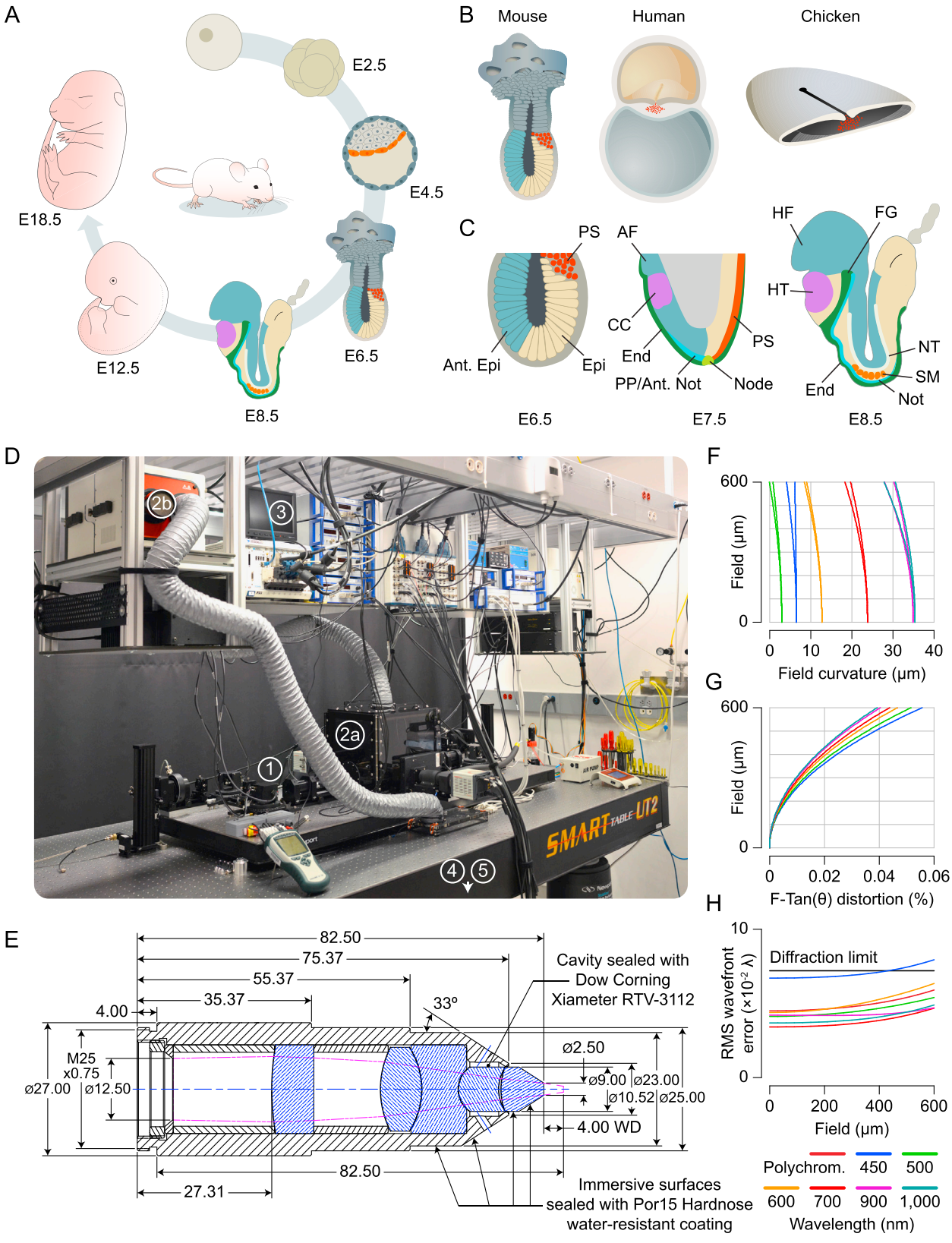
Our software module for cell division detection requires Docker for easy execution of the preconfigured framework. If you wish to use GPU acceleration, Nvidia Docker is needed as well:

- Docker is freely available here: https://docs.docker.com/engine/installation
- Nvidia Docker is freely available here: https://github.com/NVIDIA/nvidia-docker

Step-by-step protocol for the data analysis framework

Once the software environment is set up as described above, please read our step-by-step protocol for the data analysis framework provided in Methods S1F.

Supplemental Figures



(legend on next page)

Figure S1. Mouse Development, Microscope for Imaging Post-Implantation Development, and Custom Illumination Objective Design, Related to Figures 1 and 2 and STAR Methods

(A) Overview of mouse development from pre-implantation to birth. In the first four days of development (known as the pre-implantation stages, as this is the period of development before the embryo implants into the uterus) the mouse embryo develops from a single cell to an enlarged blastocyst of around 200 cells. Approximately two days later gastrulation begins with the formation of the primitive streak in the proximal posterior region of the cup-shaped epiblast, and over the next two days development rapidly progresses to the early organogenesis stages, and continues until birth around E18.5 d.p.c. to E20 d.p.c.

(B) Comparison of gastrulation in mouse, human, and chicken embryos. The shape of the gastrulating mouse embryo is unique compared to other amniotes such as human, rabbit, and chicken. The mouse embryo (left) forms a cup-shaped epiblast, and begins gastrulation in the primitive streak, a zone of epithelial-to-mesenchymal transition where cells travel out of the posterior epiblast and begin to migrate between the endoderm and epiblast (red cells). In human (middle) and chick (right) embryos, the embryo develops as a flat disc on top of a yolk sac, with gastrulation forming a trilaminar embryonic disc as cells egress from the primitive streak (red cells).

(C) Gastrulation to early organogenesis stages in the mouse embryo from E6.5 d.p.c. to E8.5 d.p.c. These are the stages covered primarily in this study, between when cells migrate out from the primitive streak into the mesodermal wings in between the epiblast and endoderm layers, and begin to assemble various tissues and organ systems such as the heart, foregut, neural folds and tube, and somites. Abbreviations: Ant. Epi, anterior epiblast; Epi, epiblast; PS, primitive streak; CC, cardiac crescent; End, endoderm; PP/Ant. Not, prechordal plate/anterior notochord; Node, node; HF's, head folds; HT, heart tube; FG, foregut; Not, notochord; SM, somites; NT, neural tube.

(D) Multi-view light-sheet microscope platform for adaptive imaging of mouse embryo development, including microscope optical components, mechanical components and cameras (1), Maus Haus environmental system (2a and 2b), real-time electronics and controllers (3). The laser systems (4) and computational equipment (5) are located outside the frame. Please see STAR Methods and Data S1A for a detailed description of the microscope's components and the full set of technical drawings.

(E) Technical drawing of the custom illumination objective with annotation of geometrical properties. This objective was designed for a narrow 33° nose taper, 4 mm working distance, and sealed with components ideally capable of withstanding repeated heating and cooling as well as chemical sterilization treatments. The objective has a numerical aperture of 0.2, a field-of-view of 1.2 mm, an effective focal length of 31.25 mm, a back aperture of 12.5 mm, a parfocal length of 82.5 mm and was designed for diffraction limited performance in the spectral band 450-1000 nm.

(F) Field curvature measurements for custom illumination objective, as a function of wavelength.

(G) F-Tan (θ) measurements for custom illumination objective, as a function of wavelength.

(H) Wavefront error (in units of lambda) for custom illumination objective, as a function of wavelength.

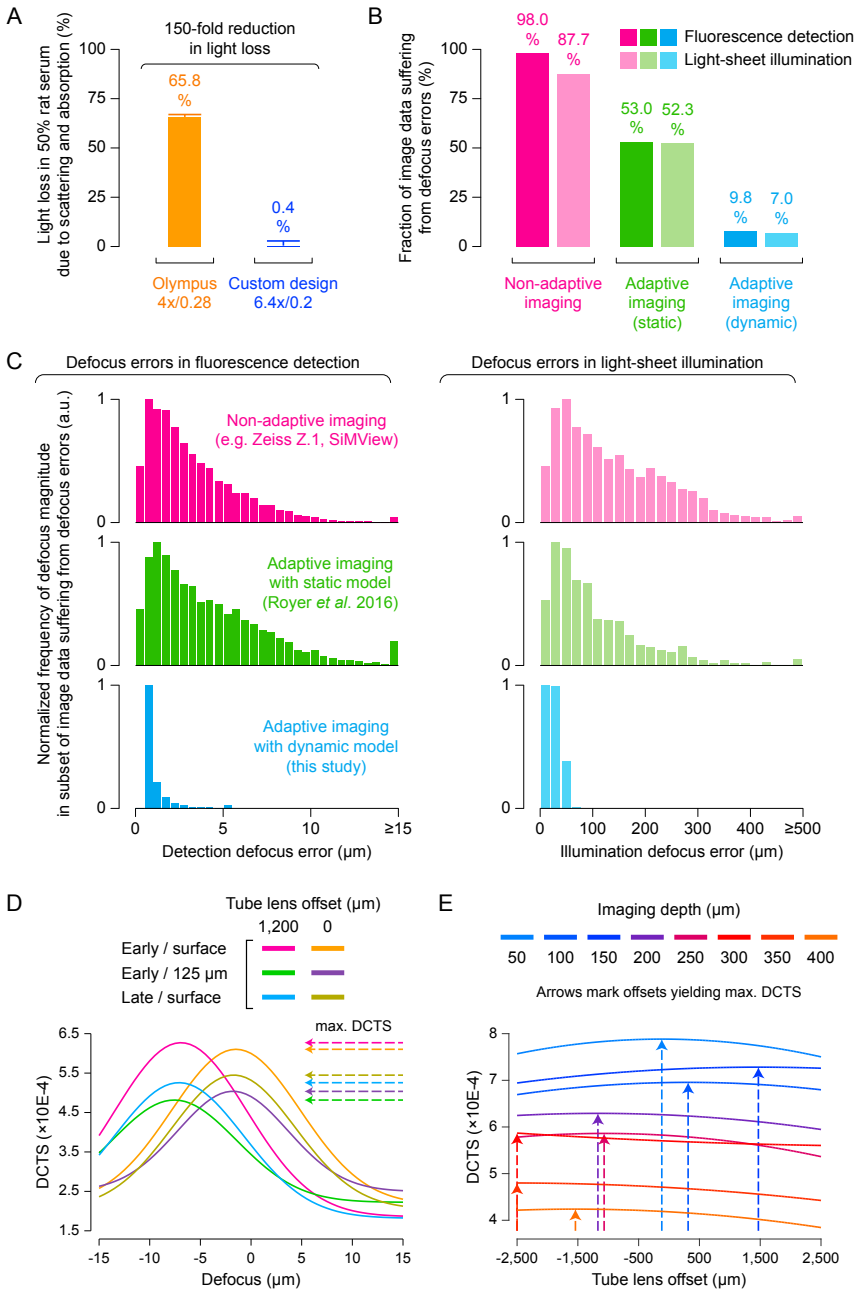


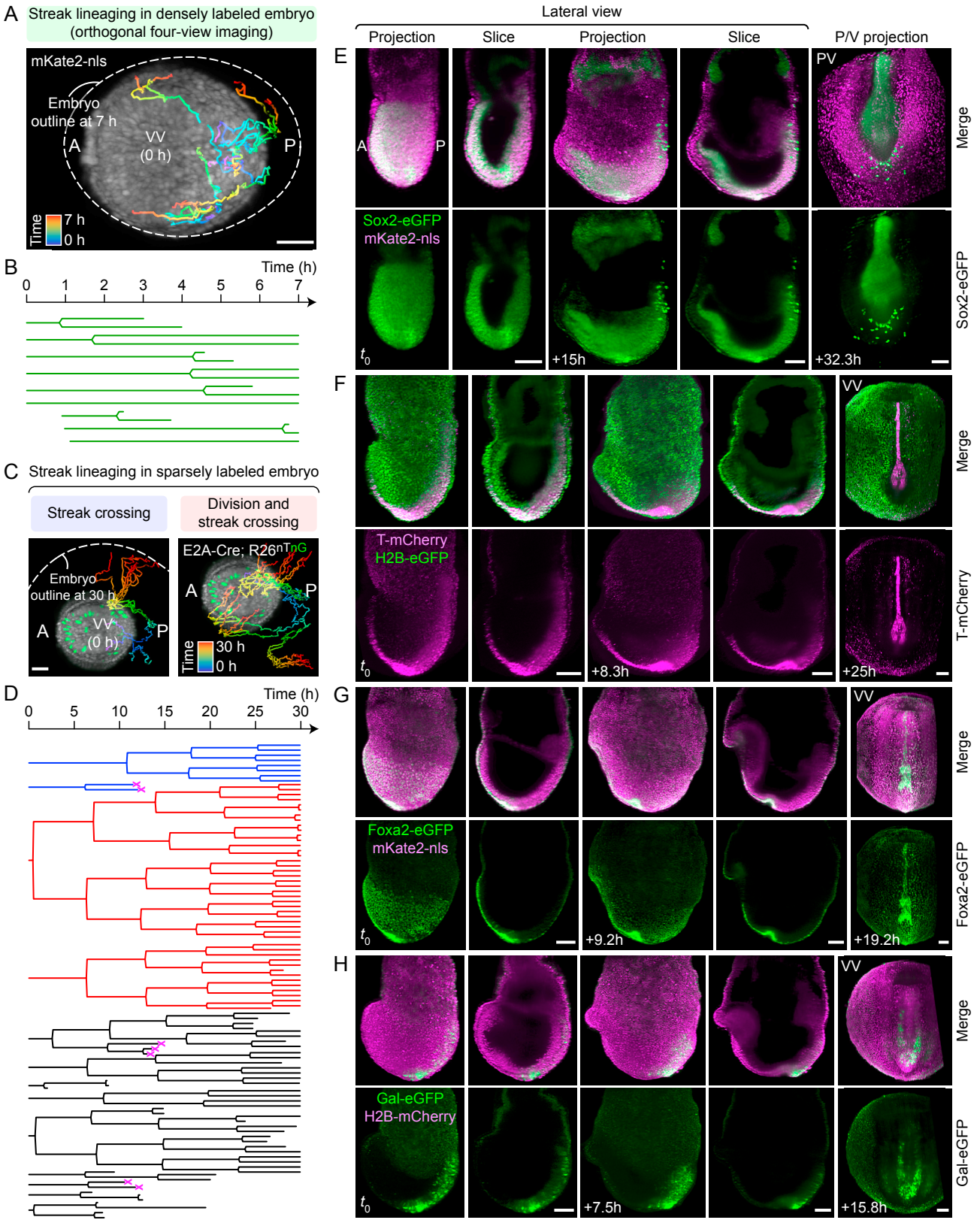
Figure S2. Quantification of Improvements in Light-Sheet Imaging Performance Using Custom Illumination Objectives and Adaptive Imaging Techniques, Related to Figures 1 and 2 and STAR Methods

(A) Performance comparison between our custom-designed 6.4x/0.2 water-dipping objective and a high-quality commercial illumination objective (Olympus XLFLUOR 4x/0.28). The Olympus objective previously served as the primary illumination objective in the original SiMView microscope design (Tomer et al., 2012) and was also chosen for this comparison here as the highest-quality commercial option available that is still geometrically compatible with both Nikon 16x/0.8 and Zeiss 20x/1.0 detection objectives (other commercial illumination objectives with shorter working distances have too wide nose pieces to allow their positioning at a 90 degrees angle relative to at least one of these two detection objectives). For this comparison, the two objectives were installed in the two illumination arms of the light-sheet microscope, which were then configured such that identical light sheets were produced in a water-filled specimen chamber (i.e., using the same illumination NA). The performance comparison was then carried out under identical conditions in standard mouse imaging culture media (50% rat serum in Fluorobrite DMEM). The comparison shows that a dramatic reduction in the amount of light lost to absorption and scattering is achieved with our custom illumination objectives (150-fold reduction compared to Olympus objective). The reduction of light loss to 0.4% can be attributed for the most part to the optimization of the optical path length inside the culture media achieved with the custom objective design.

(B and C) Side-by-side comparison of frequency and magnitude of defocus errors in fluorescence detection and light-sheet illumination when imaging developing mouse embryos with conventional, non-adaptive light-sheet microscopy (magenta), adaptive light-sheet microscopy using a static specimen model (Royer et al.,

(legend continued on next page)

2016) (green), and adaptive light-sheet microscopy using a dynamic specimen model (this study, blue). The measurements shown here are based on n = 20 time-lapse imaging experiments, in which post-implantation mouse development was recorded for on average 24 hr each and which yielded in total 680,000 DCTS-based focus measurements across the embryo volume and over time. When using non-adaptive imaging (such as in the case of SiMView or Zeiss Z.1 light-sheet microscopes), almost the entire specimen suffers from a non-zero amount of defocus aberrations (87.7%–98.0%) (B). Due to the rapid growth and continuous change in shape of the developing mouse embryo, even adaptive imaging with a static specimen model still suffers from defocus aberrations in more than half of the specimen (52.3%–53.0%), compromising the microscope's ability to acquire high-resolution image data for most of the embryo. By contrast, the adaptive imaging methodology based on dynamic mapping of sample geometry and optical properties developed in this study reduces the fraction of image data suffering from defocus aberrations to only 7.0%–9.8%. The corresponding distribution of the magnitude of these aberrations is shown in the histograms in (C). Note that these histograms specifically show the distribution of defocus errors in those parts of the specimen that suffer from defocus aberrations (i.e., only non-zero data points are included). Across the entire specimen, detection defocus errors are on average 2.99 ± 1.34 μm for non-adaptive light-sheet microscopy, 1.90 ± 0.81 μm for adaptive imaging with a static specimen model, and 0.06 ± 0.04 μm for adaptive imaging with a dynamic specimen model (Table S1). This means that both non-adaptive imaging and adaptive imaging with a static specimen model produce defocus errors that are on average larger than the axial depth of the detection focal volume itself (confocal parameters are 1.14-1.93 μm for a numerical aperture of 0.8-1.0), thus greatly decreasing resolution and image contrast. (D) By using motorized tube lenses in the detection arms of the microscope, the AutoPilot framework can jointly control objective and tube lens positions and thereby compensate for spherical aberrations in fluorescence imaging. To this end, the control software determines the tube lens settings needed for optimal image quality (defined as the maximum value of our AutoPilot image quality metric, the Shannon entropy of the Discrete Cosine Transform, DCTS). This feature is implemented as an optional module of the AutoPilot framework and is controlled separately from the core degrees of freedom *D*, *I* and *Y* used in all of our imaging experiments (controlling lateral and axial light-sheet offsets and the position of detection objectives). The DCTS plots shown here indicate that different tube lens settings are needed for optimal image quality in early versus late developmental stages and in superficial versus deep tissue regions. We note that the magnitude of this effect is relatively small compared to e.g., a 5 μm defocus correction (which has a 6-fold higher impact on DCTS in this example), however, and we thus typically restrict AutoPilot corrections to *D*, *I* and *Y* parameters to maximize performance improvements based on a minimal number of AutoPilot measurements. (E) Image quality (DCTS) versus tube lens offset as a function of imaging depth in the embryo. A 5-mm correction range is approximately sufficient to enable optimal parameter settings from the surface of the embryo to a depth of 400 μm (arrows mark optimal setting at each depth). Optimal settings change as a function of depth as expected approximately from the relative change in optical path lengths in the culture medium versus the inside of the embryo.



(legend on next page)

Figure S3. Reconstructed Lineages of Individual Epiblast Cells Traversing the Primitive Streak, and Live Expression of Fluorescent Genetic and Ubiquitous Nuclear Reporters, Related to Figure 2

(A) Visualization of cell tracks for individual epiblast cells as they migrate from the epiblast layer, through the primitive streak, and out into the mesendoderm or endoderm layers. Cells were tracked using a ubiquitous nuclear reporter (mKate2-nls). The shape of the embryo for the last time point tracked is outlined in white dashed lines.

(B) Reconstructed lineage trees for individual epiblast cells as they migrate through the primitive streak and beyond, for cells tracked using a ubiquitous reporter (A).

(C) Visualization of cell tracks as in (A) but using a mosaic reporter in which only a small subset of cells are labeled (E2A-Cre; $R26^{nT/nG}$). Epiblast cells can be seen crossing from the left or right side of the embryo through the streak to the opposite side of the embryo (left panel), or entering the streak, dividing, and each daughter crosses to the opposite side of the embryo (right panel). The shape of the embryo for the last time point tracked is outlined in white dashed lines.

(D) Reconstructed lineage trees as in (B), but for cells tracked using a mosaic reporter (C). Example lineages of epiblast cells that cross the streak are in blue, those that divide and then cross the streak to the opposite side of the embryo in red. The remainder of the lineages are for those epiblast cells that did not cross to the opposite side of the embryo after traversing the streak. Several lineages not only ended in cell death (denoted by a pink "X") but both progeny died within a few minutes of one another. Cell divisions are not synchronized across lineages, however some synchronization appears to occur within an individual lineage, as cells originating from the same parent epiblast cell will often divide within minutes of each other, even after having been separated for multiple generations.

(E) Sox2-eGFP expression from the mid/late-streak stage embryo to the early somite stage. Sox2-eGFP is robustly expressed in the neural ectoderm, chorion, and primordial germ cells (PGCs). PGCs can be visualized early in the posterior proximal region of the streak and are later drawn inward at the formation of the hindgut portal. The full time-lapse dataset is shown in Video S3C.

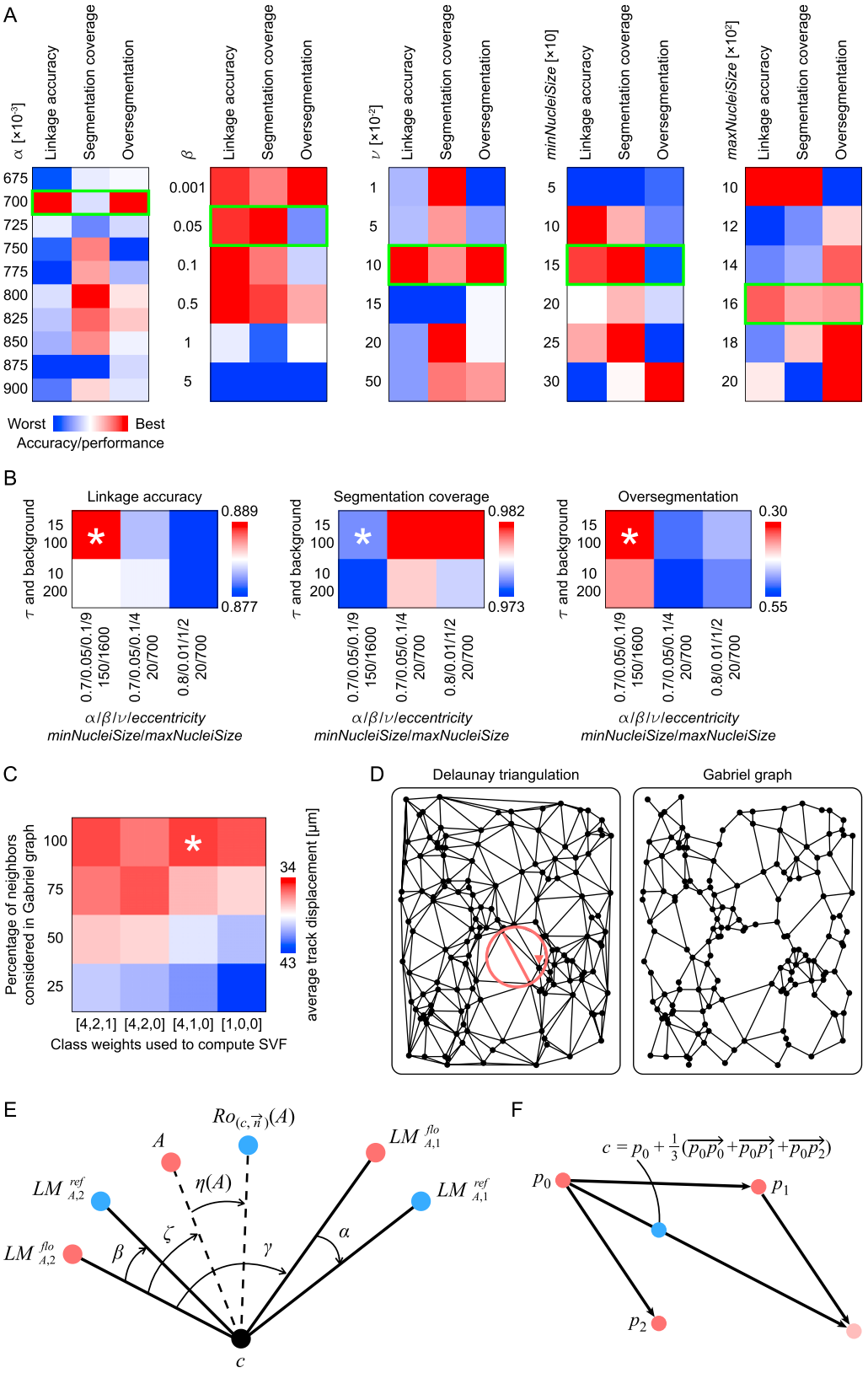
(F) Brachyury (T) mCherry is initially expressed in the streak, node, and anterior mesoderm which later converges into the anterior notochord. By the somite stage T-mCherry expression is largely restricted to the notochord and node.

(G) Expression of Foxa2-eGFP is seen in both node and notochord, as well as definitive endoderm.

(H) Galanin-eGFP is initially expressed in the node and primitive streak region, later is expressed in the node streak border, late primitive-streak, lateral plate and pre-somitic mesoderm, and some notochord. Expression in somites appears to be temporally regulated, as Gal-eGFP expression quickly diminishes once they have fully condensed.

A, anterior. P, posterior. VV, ventral view. PV, posterior view.

Scale bars, 50 μm (A, C), 100 μm (E-H).



(legend on next page)

Figure S4. Selected Steps of the TGMM 2.0 and SVF Parameter Screens, and Supplemental Illustrations for SVF and TARDIS Methods, Related to Figure 4 and STAR Methods

(A) Visualization of the results of an early-stage TGMM performance screening step used to determine the optimal range of parameter settings for α, β, υ, *minNucleiSize* and *maxNucleiSize* (see also STAR Methods, "*TGMM parameter set used for reconstructing post-implantation mouse development*"). Linkage accuracy, segmentation coverage, rate of oversegmentation and metrics related to cell division detection performance (not shown) were determined as a function of each parameter, while keeping all other parameters constant. This yielded initial estimates of the optimal range for each parameter, which were subsequently used as a starting point for combinatorial screens, in which multiple parameters were adjusted at the same time. An ideal result, though not always possible, is the optimization of performance (red) in all three benchmark categories simultaneously. Green boxes mark the settings that ultimately resulted as optimal in the subsequent combinatorial screens; note that these settings match the individually determined optima in 4 out of the 5 cases shown here (with the exception of β).

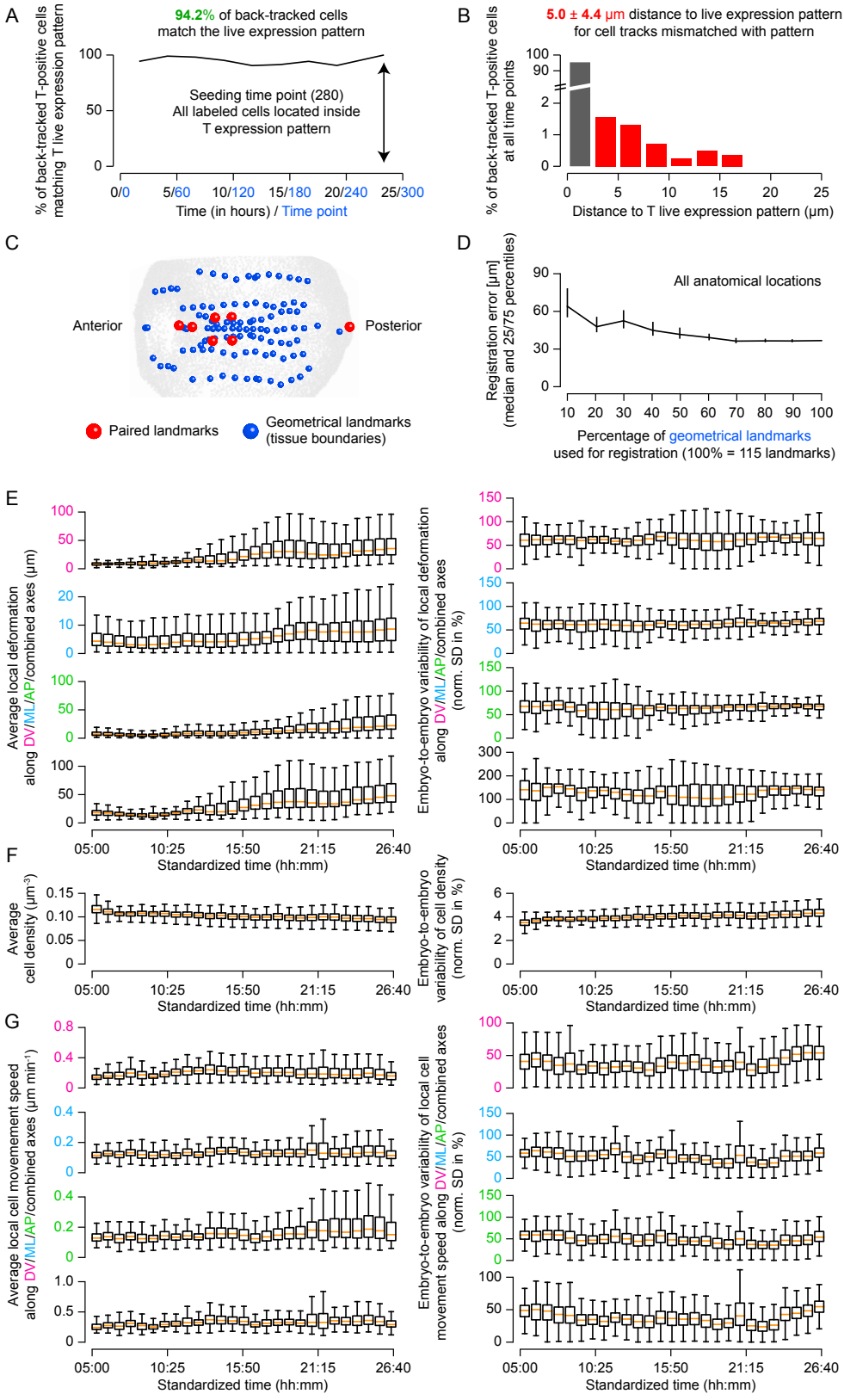
(B) Visualization of the results of a late-stage TGMM combinatorial performance screening step, in which segmentation parameters (τ and background threshold) and TGMM 2.0 parameters (α, β, υ, *minNucleiSize*, *maxNucleiSize* and *eccentricity*) were varied at the same time. The white asterisk marks the globally optimal parameter configuration used for TGMM data processing in this study. While parameter screening itself was performed based on a relatively small set of ground truth annotations, the final performance data reported in Table S1 is based on a large ground truth dataset spanning multiple tissues and developmental stages.

(C) Visualization of the results of the SVF performance screen used to determine the optimal settings of the two key parameters of the SVF framework, the percentage of neighbors considered in the Gabriel graph and the respective class weights used to define size and structure of local statistical ensembles for vector flow computation (see STAR Methods for details). Highest performance is indicated in red. The white asterisk marks the globally optimal parameter configuration used for SVF data processing in this study (100% neighborhood size, [4, 1, 0] class weights).

(D) Illustration of edges corresponding to Gabriel graph versus Delaunay triangulation for an example set of nodes. The red circle defined by the red edge highlighted in the Delaunay triangulation marks the region that is analyzed during the construction of the Gabriel graph to test for the presence of nodes inside the circle. Since there is indeed a node located inside the circle (red arrowhead), this example edge is pruned when computing the Gabriel graph.

(E) The locations and angles shown in this figure represent the respective parameters used in the STAR Methods section describing the principles of the TARDIS framework. $\mathscr{R}o(A)$ is the position of A after rotation by angle $\eta(A)$.

(F) Illustration of the computation of the centroid (c) of three points (p_0, p_1, p_2) using the average of the vectors of p_0 to each point, as described in Equation 21 of STAR Methods, which underlies the creation of the average embryo.



(legend on next page)

Figure S5. Accuracy of SVF-Reconstructed Cell Tracks Seeded by Brachyury Expression Pattern, Quantification of the Precision of the TARDIS Registration Framework as a Function of the Number of Landmark Annotations, and Stereotypy and Variability of Local Cell Dynamics in the Average Mouse Embryo, Related to Figures 4, 5, and 6 and STAR Methods

(A) Using an H2B-eGFP and Brachyury(T)-mCherry dual-expressing embryo, we reconstructed the tracks of cells expressing T-mCherry at a late time in the imaging experiment (time point 280) exclusively with tracking information obtained from the ubiquitous nuclear marker. After tracking cells in the H2B-eGFP channel with TGMM and SVF, a label mask obtained from the T-mCherry channel at time point 280 was used to select SVF objects within the T pattern at that point. The resulting time-lapse reconstruction of cell tracks and developmental origins of the T-positive cells is shown in Video S5C. In the plot shown here, we evaluated the accuracy of the SVF-reconstructed cell tracks from time point 0 to 280 by quantifying the spatial correspondence between the live expression pattern and the SVF cell positions as a function of time. To this end, we compared the location of the T-mCherry labeled SVF objects propagated backward in time to the actual location of the T-mCherry positive signal in the image data. On average 94.2% of T-mCherry positive SVF objects fell within the location of the actual T-mCherry expressing region of the embryo, confirming the accuracy of the SVF technique and ambivalence to the labeling strategy used (i.e., selecting cells based on morphological criteria, reporter expression, or any other form of arbitrary labeling).

(B) For the 5.8% of cell positions that did not overlap with the T-mCherry expressing region of the embryo (see (A)), we measured the distance between these cells and the T-mCherry expressing region. The results are shown here as a distance histogram. Complementing the quantification provided in (A), we found that cell tracks that did not overlap with the T-mCherry expressing were on average only 5.0 ± 4.4 μm away from a T-positive location. Thus, even for the small fraction of mismatched cell tracks, the spatial discrepancy with the T live expression pattern was below one cell diameter.

(C) Visualization of the location of paired and geometrical landmarks used in the TARDIS registration. Paired landmarks are point annotations that can be unambiguously located in each embryo, whereas geometrical landmarks correspond to spatially extended structures (such as tissue boundaries), which cannot be unambiguously matched with a point-based localization scheme (please see STAR Methods for a more detailed definition of these terms as well as details on their use in TARDIS). Most of the landmarks used by TARDIS fall into the category of geometrical landmarks (blue spheres), which comprises various anatomical regions and tissue boundaries such as the distal most midline region of the embryo, lateral most edges of the neural plate and lateral plate mesoderm, as well as locations for the heart field, headfold, central position of the dorsal aortae, posterior and anterior extremes, and extra-embryonic locations such as the allantois midline (please see STAR Methods and Data S1Q for a comprehensive list). Geometrical landmarks are used in the first step of the TARDIS registration workflow to align multiple embryos in space using a rigid transformation (Figure 5A). Paired landmarks (red spheres) are a series of landmarks such as the location of the somites, the posterior-proximal most region of the tailbud or posterior primitive streak, the anterior-most end of the notochord before the foregut portal, and the anterior and posterior most location of the developing heart field. Paired landmarks are used in the second step of the TARDIS registration workflow to align anatomical structures along the antero-posterior axis (Figure 5A). Since the vast majority of landmarks that can be identified and annotated in the developing embryo are geometrical landmarks, we evaluated the dependency of registration precision on landmark count (see (B)), thus determining both the overall precision of the TARDIS registration method as well as the landmark count that offers the optimal balance between registration precision and overall annotation effort. As a statistical basis for this analysis, we used the pool of 115 geometrical landmarks visualized here.

(D) Registration error of the TARDIS framework as a function of the number of geometrical landmarks. As expected, increasing the number of geometrical landmarks used for registration decreases registration error; however, the error stabilizes across the embryo at a landmark count of around 50-60 and diminishing returns are observed with further expansion of the data pool. We determined an average registration error of 41.5 μm, corresponding to a spatial mismatch on the order of about 2 cell diameters, for a landmark count of 50 and observed only a minor decrease to 36.6 μm when doubling this count. We thus recommend the use of approximately 50 landmarks per annotated time point (and ideally covering the classes defined in STAR Methods) when registering developing mouse embryos with TARDIS. Annotating the time-lapse data in intervals of 50 time points (assuming a temporal resolution of 5 min) furthermore ensures that landmarks on average do not move by a distance exceeding the registration error between annotated time points. These same guidelines were also systematically implemented throughout this study.

(E) Quantification of differences in local embryo shape across individuals, after their rigid alignment in space. This quantification was performed as a function of time by measuring the local deformations required to anatomically align four separate embryos into the average embryo (STAR Methods, Video S7A). The average spatial deformation across all embryos is shown to the left (as a total deformation as well as broken down into its components along the DV, ML and AP axes), whereas the variability of these deformations between embryos is shown to the right (quantified as the standard deviation normalized to the average). A time-lapse visualization of average and standard deviation of local deformations across the embryo is provided as Video S7B.

(F) As in (E), but for local cell density. A time-lapse visualization of average and standard deviation of local cell density across the embryo is provided as Video S7B.

(G) As in (E), but for local cell movement speed. A time-lapse visualization of average and standard deviation of local cell density across the embryo is provided as Video S7C.

DV = dorsoventral, ML = mediolateral, AP = anteroposterior, B = bud stages, HF = headfold stages, S = somite stages.