

www.sciencemag.org/cgi/content/full/1162493/DC1

Supporting Online Material for

Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy

Philipp J. Keller,* Annette D. Schmidt, Joachim Wittbrodt,* Ernst H. K. Stelzer

*To whom correspondence should be addressed. E-mail: keller@embl.de (P.J.K.),
wittbrodt@embl.de (J.W.)

Published 9 October 2008 on *Science Express*
DOI: 10.1126/science.1162493

This PDF file includes

Materials and Methods
Figs. S1 to S9
References

Other Supporting Online Material for this manuscript includes the following:

Movies S1 to S16

All movies can be downloaded as Apple QuickTime and DiVX versions on the public digital embryo repository website: <http://www.embl-heidelberg.de/digitalembryo>.

Supporting Online Material (SOM)

A) Supplementary Methods

Digital Scanned Laser Light Sheet Fluorescence Microscopy (DSLM)

The components of DSLM

The DSLM illumination system

The DSLM detection system

DSLM compared to conventional, confocal, two-photon and SPI microscopes

Comparison of the energy load in DSLM and confocal/two-photon microscopy

Image processing and data analysis

Overview

The segmentation processing pipeline

Segmentation efficiency and detection efficiency of migratory tracks

Data analysis

DSLM sample preparation

B) The public zebrafish digital embryo database

C) Supplementary Figures

Figure S1: Blueprint of the Digital Scanned Laser Light Sheet Fluorescence Microscope

Figure S2: DSLM multi-view imaging of large specimen

Figure S3: Statistical analysis of nuclear properties in the animal hemisphere

Figure S4: The image processing pipeline

Figure S5: Pre-processing/visualization of DSLM recordings of zebrafish embryogenesis

Figure S6: Global statistical analysis of cell divisions

Figure S7: Embryo-to-embryo variability of morphogenetic key parameters

Figure S8: Quantification of the properties of the peripheral cell division waves

Figure S9: Analysis of the patterns of mesendoderm internalization on the dorsal and ventral hemispheres

D) Movies

Movie 1: Schematic operation principle of DSLM

Movie 2: DSLM multi-view time-lapse recording of zebrafish embryonic development

Movie 3: Reconstruction of zebrafish embryonic development

Movie 4: DSLM multi-view time-lapse recording of *MZoep* mutant embryogenesis

Movie 5: Reconstruction of *MZoep* mutant embryogenesis

Movie 6: Visualization of the *MZoep* mutant defect in mesendoderm formation

Movie 7: DSLM time-lapse recording of zebrafish embryogenesis in the animal hemisphere

Movie 8: Reconstruction of zebrafish embryogenesis in the animal hemisphere

Movie 9: Global cell tracking and colour-indexing of morphogenetic domains

Movie 10: Global three-dimensional pattern of cell divisions

Movie 11: Reverse engineering of migratory tracks of retinal progenitor cells

Movie 12: DSLM time-lapse recording of zebrafish embryogenesis in the dorsal hemisphere

Movie 13: Reconstruction of zebrafish embryogenesis in the dorsal hemisphere

Movie 14: Global cell tracking in the dorsal hemisphere

Movie 15: Global cell tracking in the ventral hemisphere

Movie 16: Side-by-side comparison of hypoblast formation in the dorsal/ventral hemispheres

Supplementary Methods

Digital Scanned Laser Light Sheet Fluorescence Microscopy (DSLMM)

The components of DSLMM

A Digital Scanned Laser Light Sheet Fluorescence Microscope (DSLMM) consists of six sub-systems (see [Figure S1](#)): 1) a light source, 2) a beam shaping device, 3) a scanning illumination/excitation system, 4) a specimen translating/rotating device, 5) a (multi-channel) detection system and finally 6) the electronics, electrical devices, computer and the software. In contrast to standard fluorescence microscopes, e.g. conventional or confocal fluorescence microscopes, which use the same objective lens for excitation and emission detection (epifluorescence), the DSLMM excitation and fluorescence emission collection systems are operated independently.

The complete DSLMM illumination/excitation system (sub-systems 1-3 above) consists e.g. of a multi-line argon krypton laser (e.g. Melles Griot, 35 LTL 835-230), an acousto-optical tunable filter (AOTF; e.g. AA Opto-Electronic, AA.AOTF.nC-400-650nm-PV-TN) for laser wavelength selection and intensity control, a two-axis high-speed scan head (e.g. GSI Lumonics, VM500+), an f-theta lens (e.g. Sill Optics, S4LFT0061/065*) and a low-NA illumination objective lens (e.g. Carl Zeiss, Plan-Apochromat 5x/0.16) operated with a regular tube lens. The illumination/excitation objective lens is mounted on a piezo nanofocus (e.g. Physik Instrumente, P-725.CLQ), which can move the lens 400 μm along its optical axis. The specimen is placed inside a custom specimen chamber made e.g. from inert black Delrin. The specimen chamber features a temperature control system, which includes a temperature sensor inside the chamber and a heating foil attached below the chamber. Custom feedback electronics provide temperature control with a stability of ± 0.2 $^{\circ}\text{C}$ over several days.

The DSLMM detection system can take advantage of several detection objective lenses (e.g. from Carl Zeiss: Plan-Neofluar 2.5x/0.075, Fluor 5x/0.25, C-Apochromat 10x/0.45 W, Plan-Apochromat 20x/1.0 W or Plan-Apochromat 63x/1.0 W) mounted on a second independently operated piezo nano-focus, a beam splitter cube with exchangeable long-pass filters (e.g. from AHF: Q 565 LP; Semrock: RazorEdge RU 488 LP, RU 568 LP or RU 647 LP) and two detection tubes, equipped with two tube lenses and at least two cameras (e.g. the pco.2000 CCD camera from PCO AG). The DSLMM detection system currently records images with a dynamic range of ~8,000-16,000 gray levels (i.e. 14 bits) at a maximum speed of 15 frames per second and per camera and a full CCD image size of 2048 x 2048 pixels (using 2 A/D converters). Thus, the data rate generated by the dual-channel detection system is 240 MB/s at maximum speed (126 million pixels per second). For highest-quality recordings with a single A/D-converter per camera, the maximum recording speed of the DSLMM detection

system reduces to 63 million pixels per second (120 MB/s). For the experiments described in the main text the speed was typically 25 million pixels per second (49 MB/s).

To enhance its ease-of-use, our DSLM provides a secondary observation system based on a modified pivoting stereo microscope (Carl Zeiss, Stemi 2000-C) equipped with e.g. an infrared light source in the camera port for transmission illumination. The user can pivot the stereo microscope to focus on the centre of the DSLM specimen chamber, while simultaneously adjusting the specimen position via the DSLM control software and observing the camera image in the primary fluorescence and/or transmission detection system. The direct observation using the stereo microscope simplifies and speeds up specimen positioning and specimen orientation for optimal DSLM imaging conditions in the primary detection system (thereby also optimizing the energy load on the specimen).

The illumination and detection sub-systems are complemented by a specimen positioning system, comprising of a set of three linear translation stages (Physik Instrumente, M-111K028) and one micro-rotation stage (Physik Instrumente, M-116.DG). The rotation stage with its customized port provides connectors for anodized aluminium specimen holders that hold glass capillaries and plastic syringes.

The DSLM's computer provides four central processing units (CPUs; Intel, Core 2 Quad Q6600 2.4 GHz). The operating software uses two physical CPUs exclusively for the communication with the two camera pipelines, while the remaining two CPUs handle the graphical user interface (GUI), sub-routines for real-time image processing and the communication with the main electronics hub. The mass storage device is a high-performance hardware RAID controller (Promise, SuperTrak EX8350) with two RAID-0 dual-disk arrays (one per camera), each with 1.5 TB of disk space. The recorded images are written in an alternating manner to the independent RAID systems, thereby doubling the maximum manageable data rate and allowing for operation of the dual-camera detection system at the speed limit of the camera's A/D converters. Dual-camera data transfer is facilitated via a dual-camera-link controller card (National Instruments, PCIe-1430), installed in the acquisition computer. The other electronics controller cards are located in the DSLM electronics hub, which is attached to the computer via an Ethernet-linked bus extender (Hartmann Elektronik, StarFabric Bridge). The electronics hub includes the scan controller (GSI Lumonics, HC/3), a four-channel stage controller (Physik Instrumente, C-843.41), two multi-channel I/O controllers (National Instruments, PCI-6733) and the custom mainframe relay system. The AOTF beam control unit (AA Opto-Electronic, AA.MOD.8C-C**-75.158.24VDC) with an independent linear power supply (Kniel, CA 24.2,5), the dual-channel scanner drivers (GSI Lumonics, MiniSax) with dedicated linear power supplies (Kniel, CA 15.4) and the custom environmental control system for the specimen chamber are in a separate crate.

The custom DSLM control software was developed within the .NET framework 3.0 (Microsoft), using the programming language C# for the user interface and the high-level control layers and C++ for lower-level hardware communication.

The DSLM illumination system

One of the fundamental novel ideas of the DSLM concept is the use of laser scanners to create a two-dimensional sample illumination profile perpendicular to the detection axis. In the standard mode of DSLM operation, one of the scan mirrors of the scan head moves at a constant speed within a predefined angular range. An f-theta lens converts the angular scan range of the laser beam into a vertical set of parallel beams. Because of the scanning approach, a single diffraction-limited beam of light illuminates the sample at any time point along a well-defined line in space. Integration over time and space results in the illumination of an entire plane. Thus, in order to obtain a two-dimensional image, the camera integrates the signal while the laser scanners illuminate the respective plane in the detection system's field of view. By using a constant scan speed, a homogeneous light sheet-like profile is generated, i.e. all horizontal lines are illuminated with the same light intensity. Because of the high scan speed, the entire two-dimensional profile is created within less than 1 ms independent of the extent of the field-of-view.

At any time point during the illumination process, the entire laser intensity (as determined by the setting of the AOTF) is focused into a single diffraction-limited line. In previous light sheet microscope implementations (SPIM) apertures are necessary to shape the intensity profile (resulting in an average illumination efficiency of 3 %). In contrast, the DSLM illumination efficiency is 95 % is only limited by the losses at the mirrors and lenses, which amount to considerably less than 1 % per surface (and less than 5 % in total). In addition to the high illumination efficiency, the scanning approach provides a precise control over the vertical extent of the illumination profile.

The DSLM detection system

The main features of the DSLM detection system are a) a (quasi-)parallelized detection concept that uses two cameras simultaneously, b) the possibility to avoid moving elements such as shutters or filter wheels, c) cameras with a very large field of view, d) a minimal number of optical elements and e) an electronics concept that allows us to independently address all components digitally and in parallel. In combination, these features maximize the imaging speed.

The detection objective lens is positioned such that its focus is co-planar with the plane generated by the illumination/excitation system, i.e. the DSLM detection axis is oriented perpendicular to the illumination plane. Each DSLM detection arm consists of a light insulated tubing system with a tube lens and a camera. To prevent the formation of a dust layer on the protective glass of the camera, both the tube lens and the camera are directly attached to the metal tube cylinder. The fluorescence light collected by the detection objective lens is either transmitted or deflected by 90° at the central detection beam splitter. The beam splitter is optimized for maximum photon efficiency in double-marker experiments that generate fluorescence emission in the channels 488-568 nm and ≥ 568 nm (e.g. GFP and tdTomato). The fluorescence light emitted by the first marker is transmitted by the first tube

system, while the photons emitted by the second marker are deflected towards the second tube system. The beam splitter can be removed to obtain the maximum amount of light in single-channel experiments.

The cameras are triggered asynchronously by the operating software. In the two-channel DSLM recording mode, the two laser wavelengths required to excite the fluorescent marker molecules are activated sequentially. Ideally, while one camera records the signal of the molecules excited with the active laser line, the other camera transfers its data from the previous illumination cycle. Since the electronic down times of currently available CCD cameras due to A/D converter read-out and data transfer (>60 ms) are typically the main limitation in the recording speed, the asynchronous two-channel recording mode effectively doubles the DSLM acquisition rate. As an additional benefit, the concept of alternating acquisitions entirely avoids the need for spectral un-mixing of the two fluorescence signals. Our preferred strategy for specimens with two labels is to excite each fluorophore separately with the optimal wavelength. The fluorescence emission is then observed independently for each fluorophore in a filter arrangement that is optimized for the observation of exactly these two dyes. Our set-up would also allow us to excite two dyes at the same time using two or more laser lines and then split the emission with appropriate filters and record the images in parallel. However, the filter requirements become strict and considerable amounts of energy would not reach the detector. The loss in energy would have to be compensated by longer recording times to achieve an acceptable signal to noise ratio.

DSLM compared to conventional, confocal, two-photon and SPI fluorescence microscopes

1) Illumination efficiency

In DSLM, the entire laser intensity is focused into a single diffraction-limited beam of light, which illuminates the in-focus region of the specimen. Since no apertures are necessary to shape the intensity profile, the DSLM illumination efficiency is almost 100 %. The total losses at mirrors and lens surfaces amount to less than 5 %.

In contrast, the illumination efficiency of SPIM is on average only 3 %. This dramatic loss in illumination efficiency (not in detection efficiency) is caused by a set of horizontal and vertical slits that are necessary to crop the laser profile. The openings define the height and the thickness of the light sheet. This loss occurs both in high-magnification (40-100x) and low-magnification (1-10x) experiments. In high-magnification SPIM experiments a large illumination NA is required to obtain a thin light sheet, i.e. the thin laser beam has to be expanded. However, since the detected field is very small due to the high magnification, the laser profile must subsequently be severely cropped along the vertical dimension. Similarly in low-magnification SPIM experiments a low illumination NA is required to obtain a reasonably homogeneous light sheet, but the laser beam has to be expanded in order to illuminate the large detection field. In order to reduce the NA, the laser profile must, therefore, be cropped along the horizontal dimension. In conclusion, using the same light source, DSLM records images considerably faster in specimens with a low fluorescence yield

(e.g. due to a weak staining). Compared to SPIM a factor of 30 times more fluorescence light is obtained from the sample in the same time interval with a DSLM. Another major advantage is that the power requirements for the DSLM light source are much lower than for the SPIM light source. This allows one to use cheaper, simpler, lower-power and smaller lasers, i.e. gives access to a much broader range of lasers.

Confocal fluorescence microscopes perform optical sectioning by discriminating against out-of-focus light with the detection pinhole. The detection of the signal from a single plane requires the entire sample to be illuminated. Fluorescence light that is emitted from out-of-focus regions (the majority of the signal) is simply blocked by the detection pinhole. Depending on the imaging depth within the sample, the confocal fluorescence microscope's illumination efficiency is relatively low. Two-photon fluorescence microscopes are more efficient than confocal fluorescence microscopes, since a selection criterion is introduced by the probability density function of the two-photon excitation process. Only the specimen's fluorophores close to the focal region are sufficiently well excited.

2) Photo-bleaching

In order to achieve the same integrated intensity in the camera chip, DSLM and SPIM deposit the same amount of electro-magnetic energy in the specimen. Both microscopy techniques perform optical sectioning by illuminating single planes in the sample at a right angle to the detection system. While SPIM illuminates the entire plane with a constant illumination power, DSLM illuminates the plane section by section. It should be noted that the thickness of the illumination beam spans several lines in the camera and does not necessarily correspond to a single line or less in a camera. For example, at 10x magnification and a CCD pixel pitch of 7.4 μm , a DSLM section corresponds to six complete lines on the camera chip (covering a 4 μm stripe of the object plane). During the illumination period of a section, DSLM employs a higher laser power than SPIM (e.g. a factor of 300 at 10x magnification). However, only the illuminated section is exposed to the illumination light while the remainder of the specimen remains un-exposed. The required range of laser powers in the DSLM is typically between 10 μW and 1 mW. At the upper end of this range, we measured a maximum DSLM irradiance level of 2 kW/cm^2 in the focus region of the illumination profile, which is less than 20 % of the semi-saturation level of GFP (11 kW/cm^2 , [SI]). Thus, under the least favourable conditions the DSLM still operates in a linear photo-bleaching regime and the photo-bleaching rates in SPIM and DSLM are identical.

Since DSLM performs one-dimensional scanning, the time that is available for the light exposure of a single pixel in the image is several orders of magnitude longer than in confocal and two-photon fluorescence microscopes, which employ two-dimensional scanning. For example, during the 50 ms exposure time per image that is applied in our zebrafish experiments, each pixel in a DSLM image is exposed for ~0.5 ms. In contrast confocal and two-photon fluorescence microscopes require pixel dwell times of 5 μs and less. Therefore, DSLM can afford a factor of 100 less laser power in the illumination process compared to confocal and two-photon fluorescence microscopes.

A distinct advantage of DSLM and SPIM is that both implementations only illuminate the plane that is also observed. This advantage can be expressed as a ratio n , which is the total depth of the image stack required by an experiment divided by the light sheet thickness. Since confocal and conventional fluorescence microscopes illuminate the entire specimen when recording single planes, photo-bleaching rates in conventional fluorescence microscopes are at least by a factor of n higher than in DSLM. In confocal fluorescence microscope the situation is even worse. E.g., in a typical DSLM recording of zebrafish development, the light sheet thickness is 4-6 μm and the image stacks are 600 μm deep. Therefore, DSLM photo-bleaching is by a factor of 100-150 lower than in a conventional fluorescence microscope. Compared to confocal fluorescence microscopes, which operate in non-linear photo-bleaching regimes (caused by the high laser power), this factor is usually even larger. A detailed side-by-side comparison of light exposure energies and photo-bleaching rates in zebrafish live imaging with a confocal fluorescence microscope (Leica SP5) and with the DSLM is given below in the section “Comparison of the energy load in DSLM and confocal/two-photon microscopy”. Two-photon fluorescence microscopes require ultra-short high-energy light pulses in the illumination process due to the low two-photon excitation cross section and therefore operate in a non-linear photo-bleaching regime. In addition, two-photon optical sections exhibit a large overlap, which results in incremental photo-damage in three-dimensional imaging. Therefore, DSLM photo-bleaching rates are also considerably lower than those found in two-photon fluorescence microscopes.

3) Lateral and axial extents of the point spread function

The lateral and axial extents of the point spread functions of confocal, two-photon and light sheet based fluorescence microscopes are defined by the wavelengths of the excitation emission light, the numerical apertures of the objective lenses and, of course, by the signal to noise ratio in the image. Since the same objective lenses can be used in all four techniques, the advantage of one technique over another with respect to the lateral resolution in the recorded data is minor. The lateral performance of the conventional and the light sheet based fluorescence microscopes are identical. The confocal fluorescence microscope will perform $1/\sqrt{2}$ better and the two-photon fluorescence microscope $\sqrt{2}$ worse, i.e. the latter will perform even worse than a conventional fluorescence microscope [S2]. It is also well-known that the confocal fluorescence microscope cannot take advantage of this improved lateral performance, since it generally does not provide a sufficient signal-to-noise-ratio (e.g. [S3]).

The situation is different for the axial extents of the point spread functions. The axial extent of the point spread function of a conventional microscope is simply determined by the depth of focus of the detection objective lens. The axial extent of any optically sectioning fluorescence microscope is related to the depth of focus of the excitation and detection objective lenses. Light Sheet Based Fluorescent Microscopes (LSFM) perform optical sectioning with a diffraction-limited illumination profile and provide multi-view capability, i.e. the option to record multiple data sets of a sample along different angles. In single-view experiments that use lenses with relatively low numerical apertures, the LSFMs' axial extents of their point spread functions are mainly determined by the thickness of the illumination

profile. However, it is important to note that the illumination and detection systems in LSFM are decoupled. In contrast to conventional and confocal fluorescence microscopy the lateral and axial extents of the point spread functions can in fact be defined independently. Usually, the NAs of the LSFM illumination profiles are defined such that the profile thickness at the edge of the field is by a square root of two larger than the central thickness, which ensures reasonably homogeneous imaging conditions across the full field of view. Hence, axial resolutions are quite similar in SPIM and DSLM. In multi-view experiments the axial resolution can be further improved by fusing the information obtained along different recording angles. For some specimens, the fusion of a few angles already leads to an almost isotropic three-dimensional resolution, i.e. the considerably lower axial resolution becomes equal to the much better lateral resolution.

In summary, DSLM outperforms confocal and two-photon fluorescence microscopes for the entire range of detection NAs in multi-view experiments. In single-view experiments and for numerical apertures up to about 0.6-0.8 (depending on the field-of-view of the camera), DSLM provides an at least 50 % better axial extent of the point spread functions. At higher numerical apertures DSLM and SPIM still outperform two-photon microscope, while confocal fluorescence microscopes perform better than LSFMs with numerical apertures above 0.8 [S4].

4) Quality and flexibility of the illumination pattern

This comparison concerns DSLM and SPIM microscopes, which realize optical sectioning with thin diffraction-limited illumination profiles. The illumination profile of SPIM is generated by a set of beam-shaping apertures, i.e. a cylindrical lens and an illumination objective lens. In contrast, DSLM illumination employs laser scanners instead of a cylindrical lens. Thus, beam-shaping apertures are not required in DSLM. In fact, the optical components used in a DSLM are in principle identical to those used in a confocal fluorescence microscope such as e.g. the Carl Zeiss LSM 510 Meta. Hence the quality of the optical components is excellent.

In the standard DSLM illumination mode, a uniform illumination pattern is generated. The illumination laser intensities are identical for all lines of the recorded plane. The quality of the DSLM profile is only influenced by the field properties of the illumination optics. This is in contrast to SPIM, in which a light sheet profile is created by focusing an initially collimated Gaussian laser beam profile along one dimension with a cylindrical lens. The unfocused dimension retains the collimated beam properties and thus, e.g. the SPIM illumination intensity is different in the periphery of the field than in its centre. DSLM is, therefore, better suited for quantitative imaging and the imaging of large specimen than SPIM. Our current implementation of DSLM ensures a uniform illumination for samples up to 5 mm in height (see [Figure S2b](#)).

The beam-shaping apertures also affect the quality of the illumination beam. In DSLM, the unconstrained laser beam provides a diffraction-limited focal region. In SPIM, slits constrain the vertical and horizontal extent of the light sheet to obtain the desired light sheet thickness

and the required light sheet height at the same time. When imaging large specimens in SPIM ($> 200 \mu\text{m}$), the laser beam has to be expanded to achieve an acceptable level of homogeneity along the vertical dimension of the field and subsequently must be cropped along the horizontal dimension to reduce the illumination NA. The constraints imposed on the laser profile usually result in side lobes of the SPIM light sheet and increase the diffraction-limited light sheet thickness.

As an additional benefit of the laser scanning principle, DSLM provides the intrinsic capability of structured illumination and contrast enhancement. The scanning approach provides flexibility in the illumination system and e.g. allows modulating the intensity profile of the illuminating light sheet. An almost arbitrary intensity function can be realized by AOTF modulation and/or a respective scanner acceleration profile.

5) Imaging speed

DSLM currently provides two independent camera channels. The cameras are triggered in an alternating scheme and acquire images during the electronic down time periods of the other camera. While one camera processes and transfers an image to the computer, the other camera is exposed. At low exposure times ($< 50 \text{ ms}$), DSLM records data at essentially twice the frame rate of a single camera. At an exposure time of 200 ms , there is still a 60% increase in the acquisition rate, when operating e.g. the pco2000 camera with its full chip area. Importantly, the scanning approach does not introduce a limitation on the acquisition speed, since even an object plane of $5 \text{ mm} \times 5 \text{ mm}$ can be illuminated within 1 ms . Even the currently fastest EM-CCD cameras require at least 10 ms for data readout of a well-sampled image ($\geq 512 \times 512$ pixels) and thus, the camera speed is the limiting factor in the DSLM acquisition. In fact, for weakly-stained samples the DSLM scanning illumination actually increases the acquisition rate, since the laser light is focused into the sample at almost 100% efficiency and thus, a large amount of fluorescence light can be collected from the sample during a short time interval. Compared to our SPIM, our implementation of the DSLM detection system provides an at least two-fold higher image acquisition rate.

Confocal and two-photon fluorescence microscopes use two-dimensional scanning for sample illumination and photo-multipliers in the detection, i.e. the intensity level in the image is determined one pixel at a time. In a typical 512×512 confocal or two-photon scan field, the $262,144$ pixels of each image are illuminated sequentially. In contrast, the camera-based DSLM detection system collects intensity data in parallel for 6-8 entire lines on the camera chip, constituting $16,348$ pixels of the final image. The parallelized data acquisition provides DSLM with a dramatic speed advantage over confocal and two-photon fluorescence microscopes. At maximum speed, DSLM records 30 images per second (using 2 A/D-converters per camera), each 2048×2048 pixels in size, corresponding to an acquisition rate of 126 million pixels per second. State-of-the art confocal and two-photon fluorescence microscopes record images at a rate of 5 frames per second (and at 120 fps for the Zeiss LSM 5 live) at an image size of 512×512 pixels, corresponding to an acquisition rate of 1.3 million pixels per second (31 million pixels per second for the Zeiss LSM 5 live). In

conclusion, DSLM is by a factor of >90 faster than standard confocal and two-photon fluorescence microscopes and still even by a factor of four faster than the LSM 5 live.

6) Dynamic range

Both DSLM and SPIM feature camera-based detection and thereby benefit from the high dynamic range of CCDs and other semi-conductor based cameras. The DSLM detection system provides images with a dynamic range of 10,000 gray levels. Confocal and two-photon fluorescence microscopes typically provide data with a dynamic range of at most 6-8 bits (64-256 gray levels). One advantage of a high dynamic range is the possibility to perform image deconvolution to increase the resolution [S5]. More importantly, the high dynamic range is essential in time-lapse imaging of vertebrate development, since it allows us to reliably identify nuclei within a very large range of H2B-eGFP translation levels. We frequently observe the situation of nuclei exhibiting 100 times lower fluorescence intensities than the brightest nuclei in the image stack. Even at an excellent background level of only 1 %, these nuclei can only be distinguished from the background noise with a dynamic range of at least 1.000 gray levels in the detection system.

7) Cost-efficiency

Our design of DSLM is optimized with respect to the number of components and optical interfaces in the illumination system. Taking into account the number of degrees of freedom that are required to operate the elements in our instrument we are probably at most two elements short of using the minimal number of optical parts. In this sense the DSLM provides an almost optimal solution. We also took particular care to ensure a very compact implementation. The microscope currently has a footprint of 0.3 m² on the optical table (excluding the light source). The total price tag for our implementation of DSLM is 80,000 € (excluding the development of software and customized electronics).

Comparison of the energy load in DSLM and confocal/two-photon microscopy

We performed two experiments of zebrafish live imaging with a Leica SP5 confocal fluorescence microscope, using stage-matched specimens and applying comparable settings as in our DSLM recordings of embryonic development. We gave preference to this state-of-the-art confocal system over a spinning-disk confocal fluorescence microscope to obtain images of a reasonable quality, while keeping in mind the large-scale automated image processing that has to be performed for our experiments.

In the DSLM, H2B-eGFP mRNA-injected embryos were recorded as 600- μ m-stacks of 205 images (with 2048 x 2048 pixels each) within 30 seconds every 60 seconds, using a Carl Zeiss C-Apochromat 10x/0.45 water lens. In the Leica SP5 confocal fluorescence microscope, the embryos were recorded in stacks of 205 images (with 1024 x 1024 pixels each) every 130 seconds, applying a scan speed of 800 Hz and using the bidirectional scan mode. The latter configuration provided the highest recording speed in the confocal system (1.6 million voxels per second), still allowing the recording of the embryos in their entirety

(but with a considerably lower sampling than in the DSLM). However, the imaging speed in the confocal fluorescence microscope was still by a factor of 16 slower than in the DSLM, although the DSLM was not operating at maximum speed. We determined a maximum signal-to-noise ratio of 60:1 in the confocal recordings as compared to an average signal-to-noise ratio of 1000:1 in the DSLM. We used a Leica 10x/0.4 water lens in the Leica SP5 confocal fluorescence microscope as the closest match to the lens employed in the DSLM. The difference of 12.5 % in the numerical apertures of the two detection lenses (corresponding to a 27 % difference in the number of collected photons) was considered in the data analysis.

In order to obtain a single image with the same brightness, i.e. with the same nuclear fluorescence intensity levels, 46.8 μJ of electro-magnetic (EM) energy passed through an image plane in the confocal fluorescence microscope, while only 0.85 μJ were required in the DSLM. Thus a single plane was exposed to an EM energy that was a factor of 55 higher in the confocal fluorescence microscope compared to the situation in the DSLM. This is consistent with the fact that a) the confocal photo-multipliers (i.e. the detectors) have an a factor of 4-6 lower quantum efficiency than the CCD camera used in the DSLM and b) the confocal fluorescence microscope loses ~80-90 % of the signal photons at the confocal pinhole. Thus a theoretical factor of at least 40 can be expected in the EM energy requirements of a confocal fluorescence microscope.

When operating in the stack acquisition mode, i.e. when performing three-dimensional imaging (the mode required for our zebrafish experiments), we measured a factor of 5,000-6,000 higher EM energy passing through a single plane in the confocal fluorescence microscope as compared to the situation in the DSLM. This additional increase results from the fact that the excitation laser light passes through the entire specimen in the confocal fluorescence microscope, although only the fluorescence emitted from a single plane is recorded (see above). Therefore, each plane in the confocal recording is exposed to a factor of n more light as compared to the single image acquisition mode, where n is the total number of planes in the z-stack ($n = 205$ in our experiments). In contrast, the EM energy exposure in the DSLM stack acquisition mode increases only by a factor of 2 per plane (the thickness of the light sheet divided by the z-spacing in the image stack), independently of the total number of planes in the z-stack. In conclusion, an additional factor of ~100 more light passes through every plane in a confocal image stack as compared to the planes in a DSLM image stack.

Next, we determined the photo-bleaching rate in the confocal fluorescence microscope. As pointed out in the main text, the fluorescence intensity in DSLM time-lapse imaging of zebrafish embryos at a late stage (i.e. when GFP production has already ceased) remains almost constant for a time interval of at least 10 hours. In contrast, we determined photo-bleaching of at least half of the GFP molecules after one hour of confocal imaging in the same scenario. Therefore, more than 99.9 % of all GFP molecules are photo-bleached in the confocal fluorescence microscope within 10 hours, while photo-bleaching remains below 5 % in the DSLM over the same period of time.

In a second experimental comparison, we recorded zebrafish embryogenesis in two specimens with a Carl Zeiss LSM 510 Meta NLO two-photon fluorescence microscope. Again, we used stage-matched embryos and compared the energy load resulting in the two-photon fluorescence microscope to the DSLM measurements for images of identical brightness. The two-photon fluorescence microscope was equipped with a Coherent Chameleon Ultra 2 laser, a 500-550 nm band pass filter and a Carl Zeiss C-Apochromat 10x/0.45 W detection lens (the same detection lens as in the DSLM). The embryos were recorded in 600- μ m-stacks of 204 images (with 1024 x 1024 pixels each) in time intervals of 171 seconds, applying a pixel dwell time of 0.8 μ s and using the bidirectional scan mode. The latter configuration provided the highest possible recording speed in the two-photon system (1.25 million voxels per second), while still allowing the recording of the embryos in their entirety (but with a considerably lower sampling than in the DSLM). The imaging speed in the two-photon fluorescence microscope, however, was still by a factor of 20 slower than in the DSLM.

We measured an average EM energy of 8.4 mJ (at 930 nm) passing the embryo for each recorded plane in the two-photon fluorescence microscope. When operating in the stack acquisition mode, the energy load in each plane therefore resulted as 204 x 8.4 mJ = 1.7 J. The energy load in the two-photon fluorescence microscope for images of comparable brightness is thus by a factor of one million higher than in the DSLM. Importantly, the images that we recorded with the two-photon fluorescence microscope at an imaging speed of 1.25 million voxels per second were considerably noisier than those recorded with the confocal fluorescence microscope at a speed of 1.6 million voxels per second. The signal-to-noise-ratio in the two-photon fluorescence microscope was approximately 4:1. This has to be compared to a ratio of 60:1 in the confocal fluorescence microscope and to a ratio of 1000:1 in the DSLM. The ratio of 4:1 obtained with the two-photon microscope renders the data almost useless for a robust image segmentation. Moreover, although we observed photo-bleaching of the nuclear GFP in the two-photon recordings, it was impossible to quantify the photo-bleaching rate at the given data quality.

Typically, different biological effects result from illumination of tissues at different wavelengths. While we used an illumination wavelength of 488 nm in the confocal fluorescence microscope and in the DSLM, the experiments in the two-photon microscope were performed at an illumination wavelength of 930 nm. A comprehensive evaluation of comparative tissue damage in two-photon microscopy and in DSLM is thus non-trivial. However, we can directly estimate the relative energy absorption from the measured energy load. The embryo consists mainly of water, whose light absorbance coefficient is by a factor of \sim 100 higher at the wavelength 930 nm than at 488 nm [S6]. Thus, the heating effect of the illumination beam is by a factor of \sim 100 million higher in the two-photon microscope as compared to the DSLM. However, the biological effect should be the real issue when comparing different microscopes and different contrasting mechanisms. Energy load, i.e. the exposure of the specimen to a certain level of EM energy, is a very reasonable basis for a comparison, but most likely not entirely sufficient.

Automated 3D image processing and data analysis

Overview

DSLIM time-lapse recordings of zebrafish development at high spatiotemporal resolution produce large amounts of data. A typical 24 h time-lapse data set, recorded at a temporal resolution of 90 s, consists of 960 three-dimensional image stacks, each of which contains 400 planes with 2,048 x 2,048 pixels and a dynamic range of 14 bit per pixel, i.e. consuming 2 Byte per pixel or 8 MByte per image. Therefore, such a data set has a size of 960 x 400 x 8 MByte = 3 TByte and contains detailed information about the nuclear movements and cell divisions in the entire embryo, i.e. in a sample volume with the size 1,000 x 1,000 x 1,000 μm^3 . The animal hemisphere contains 16,000 nuclei after 18 hours of development and therefore, the positions and approximated morphologies of more than 10 million nuclei need to be extracted from a 24 h time-lapse data set. This task demands a simple and automated but also efficient and robust, computational solution.

We developed an automated segmentation and data analysis pipeline in Matlab (The Mathworks, Natick, MA, USA) for the processing of time-lapse data sets with a fully parallelized concept (see [Figure S4](#)). The processing is performed on the EMBL computer cluster, comprising 1,000 CPUs, and on a sub-grid of the computer cluster of the Karlsruhe Institute of Technology (KIT), comprising 4,000 CPUs. The pipeline consists of 64 processing modules, out of which 16 core programs are provided in the public digital embryo repository (<http://www.embl-heidelberg.de/digitalembryo>), including all modules of the segmentation sub-pipeline as well as the automated analysis modules for comprehensive cell tracking, detection of cell divisions and visualization of digital embryos. The segmentation pipeline consists of several modules for the crude segmentation as well as quality filters in the spatial and temporal domains. The segmentation quality filters include algorithms for the elimination of false positives as well as for the recovery of false negatives.

There are many different approaches to the segmentation of three-dimensional microscopy data. Our pipeline is based on the simple, but extremely time-efficient segmentation concept of a recursive analysis of object properties by adaptive and iterative thresholding with a simultaneous determination of the connectivity and morphology of the resulting binary data. This algorithm provides very robust segmentation results as long as the data provides a high dynamic range and an excellent signal-to-noise ratio (as is the case in the DSLIM). Importantly, the pipeline is able to fully segment and analyze a three terabyte large data set within 48 h of computation time on EMBL's cluster.

The segmentation processing pipeline

The raw DSLIM time-lapse data sets are cropped and converted into an image database (subroutine *createID*). During this initial step, pre-processed information on the experiment is collected, e.g. the time-resolved background level and the intensity histograms. The image

databases (typically 2-3 TByte of raw data per experiment) are stored on a fileserver in the EMBL computer cluster and are distributed for processing to the cluster cores in the form of small data slabs (typically 50 MB in size, including “overlap” slabs that represent the bounding volumes of adjacent data slabs). A subroutine *clusterNuclei* performs the initial segmentation of these data slabs.

In order to boost the segmentation efficiency (and to reduce filtering time), the segmentation routine is supported by *a priori* knowledge about the expected time-dependent minimum nucleus diameters. We obtained the time-course of this parameter by a manual analysis of 5,000 nuclei in three independently recorded time-lapse data sets of zebrafish development. The segmentation subroutine operates on the principle of applying step-wise increasing threshold levels, starting at the background level, and a subsequent detection of connected three-dimensional volumes in the thresholded image stacks. Each detected object is stored in a database and subject to further increasing intensity thresholds, thereby refining the information about the properties of the object and its internal morphology. If an increase in the threshold level yields several non-connected objects within a previously detected object’s volume, these non-connected objects are analysed independently and also further treated as independent objects. However, the refinement of the internal morphology is stopped as soon as the increase in intensity threshold yields an object volume equal to the known minimal nucleus size at the analyzed time point. The global thresholding is stopped when all detected objects have been refined to the pre-defined size limit or when the maximum intensity level of the image stack has been reached (indicating the presence of intensity-saturated objects). The characteristic properties of each object are computed, including its bounding box coordinates, the final thresholded volume size, the average intensity level in this volume, the gray-level based centre-of-mass coordinates, the intensity level at the centre-of-mass and the morphology of the final thresholded volume (in a three-dimensional binary representation). In order to obtain an estimate of the real object size, the algorithm analyzes the intensity profiles of lines through the object’s centre-of-mass along all three dimensions. For each object, the full-width-half-maximum (FWHM) values of these three profiles are stored (considering the background level at the respective time point). Objects that do not exhibit an intensity decrease to 50% are discarded as background-induced artefacts.

A data fusion module (*clusterCollect*) collects the raw segmentation data from the cluster nodes and fuses the information from related segmentation slabs, i.e. from those volumes that contribute to the image stack at a specific time point. The fusion module eliminates edge artefacts by discarding all edge-bounding objects and by evaluating segmentation data from “overlap” slabs at bounding coordinates.

The fused segmentation data is then passed on to the spatial- and temporal-domain filters of the processing pipeline (*clusterFilter* and *clusterCorrelate*). The spatial-domain filter analyzes the level of connectivity in the segmented object morphology as well as the level of geometrical isotropy of the object surface. The level of connectivity is determined by calculating the average number of neighbouring object pixels for all pixels that contribute to an object. Amorphous objects, i.e. objects with many hollow spaces, score a low connectivity, while solid objects score high values. A good approximation of nuclei shapes is

provided by ellipsoids, which score the highest levels of connectivity (with the maximum score corresponding to a sphere). Since the level of connectivity of ellipsoids and spheres depends on the surface-over-volume ratio, the connectivity-parameter of such objects is size-dependent (starting at zero for a one-pixel-sphere and converging to 26 for an infinite sphere/ellipsoid size). The spatial-domain filter applies a database of the size-dependency of the connectivity for perfect spheres to the known time-dependent minimum nucleus size to obtain the connectivity threshold level for nucleus-like objects. Morphologies that yield a too low connectivity, i.e. which do not match the shape of nuclei, are discarded. Similarly, the anisotropy filter discards objects that are stretched by a factor of three or more and thus significantly deviate from both an interphase nucleus (represented by a factor of one) and a metaphase nucleus (represented by an average factor of two; determined by manual analysis). The purpose of the anisotropy filter is the elimination of artefact objects, arising from the cytoplasmic background introduced by H2B-eGFP mRNA during the first two hours of embryonic development, which cannot always be recognized from their connectivity levels. After the first two hours, H2B-eGFP becomes exclusively nucleus-specific and thus, cytoplasmic artefacts do not occur. In order to save computation time, the anisotropy filter is, therefore, not used after this early developmental phase.

The temporal domain filter correlates the centre-of-mass positions of the detected objects within small time intervals and discards objects that fail to reappear consistently. Typically, a time window of eight time points (recorded in 45-90 s intervals) is evaluated for each object. The object is removed, if it is detected within a sphere with a 10 pixels radius (corresponding to 7.4 μm) for only 30 % of the time or less. This radius is sufficiently small to exclude a cross-correlation of independent objects within the time window. The criterion furthermore considers the behaviour of nuclei both during interphase and mitosis. In interphase, nucleus speeds of up to 3 $\mu\text{m}/\text{min}$ correspond to an expected 100 % success rate of identification in a 10 pixels radius for undirected (random) movements and to an expected 80 % success rate for directed movements. In metaphase, peak speeds of 8 $\mu\text{m}/\text{min}$ were observed (over maximum time periods of 2 min). Since both past and future movements are analyzed, the nucleus will either correlate along the forward time axis or along the backward time axis and thus, a minimum identification success rate of 40 % is expected (for directed movements; 50 % for undirected movements). The 30 % threshold level of the correlation filter therefore separates nuclei-like migrating objects from random segmentation artefacts. Since a small percentage of segmented nuclei are removed in this process (“false negatives”, considering a < 100 % segmentation efficiency and the resulting decrease in temporal correlation hits), we subsequently apply a processing module for the recovery of false negatives from the initial segmentation database (*clusterRecovery*).

The remaining objects in the filtered segmentation database are correlated with respect to their centre-of-mass position, their shape (represented by the bounding box) and their average intensity level. For each database entry, the object with the highest total correlation product in these seven parameters is determined both for the previous as well as the subsequent time point. The same calculations are performed on the data basis of the initial unfiltered

segmentation database. If an object with a better overall match can be retrieved from the initial database, this object is added to the filtered database (recovery of false negatives).

Segmentation efficiency and detection efficiency of migratory tracks

Manual controls were performed to obtain an estimate of the segmentation efficiency of the automated pipeline. For this purpose, nuclei counts and positions were manually determined for 13 time points in two data sets and compared to the data generated by the processing algorithms. The table below lists the fractions of false positives, i.e. segmented objects that do not represent nuclei of the embryo, as well as false negatives, i.e. nuclei of the embryo that were missed by the segmentation algorithm. The segmentation efficiency is then defined as 100 % (perfect segmentation) minus the sum of percentages of both false positives and false negatives.

Data set	Embryo age*	False positives	False negatives	Total efficiency
A, time point 10	110 mpf	0 %	1.56 %	98.44 %
A, time point 20	120 mpf	2.44 %	3.25 %	94.31 %
A, time point 50	150 mpf	2.89 %	0.36 %	96.75 %
A, time point 100	200 mpf	1.07 %	0.40 %	98.53 %
A, time point 200	300 mpf	1.12 %	0.75 %	98.13 %
A, time point 500	600 mpf	2.68 %	3.02 %	94.30 %
A, time point 1,000	1,100 mpf	2.94 %	6.62 %	90.44 %
B, time point 0	300 mpf	0.73 %	3.65 %	95.62 %
B, time point 200	500 mpf	1.47 %	6.57 %	91.96 %
B, time point 400	700 mpf	1.28 %	3.21 %	95.51 %
B, time point 600	900 mpf	2.63 %	6.14 %	91.23 %
B, time point 800	1,100 mpf	2.06 %	8.25 %	89.69 %
B, time point 1,000	1,300 mpf	2.27 %	6.82 %	90.91 %

(* mpf = minutes post fertilization)

In conclusion, we determined an average segmentation efficiency of 93.5 %. Not surprisingly, the segmentation efficiency depends on the developmental stage of the embryo. The segmentation algorithm performs best during early embryonic development up to the

early gastrula stages (97 % efficiency) and decreases during late gastrula and the somite-stages (90 % efficiency), owing to an increase in nuclei density at the embryonic body axis.

While the segmentation efficiency is equivalent to the detection efficiency (the percentage of nuclei of the embryo that were correctly identified, irrespective of their representation in the microscopy data) of nuclei during early embryogenesis, the detection efficiency starts to drop below the segmentation efficiency in late embryogenesis stages due to the increase in body axis thickness and the attenuation of the signal-to-noise ratio with penetration depth, intrinsic to the observation process with a light microscope. The average detected percentage of nuclei in the embryo results as 95 % until early gastrula stages (10 hpf, see above) and subsequently drops linearly to 85 % in late somitogenesis. The average overall detection efficiency of nuclei over the entire time course of the time-lapse experiment is thus 92 %.

We furthermore determined the tracking efficiency by manual controls and comparison to the data provided by the data analysis algorithms. The limitation in the tracking efficiency arises from the limited segmentation efficiency. If the segmentation of a nucleus fails, the tracking module may choose a different nucleus to continue the track, however, only if this nucleus happens to be sufficiently close to the expected nucleus position at the respective time point. Manual controls during late gastrulation, which constitutes the most difficult tracking scenario due to the high density of nuclei, revealed an average rate of one tracking mismatch per 200 time points, i.e. in average after 200 time points the tracking module chooses the closest neighbouring cell over the original identity. The resulting tracking efficiency of 99.5 % is more than sufficient for the analysis of global morphogenetic trends or the comparison of migratory behaviour in different tissues, but it introduces a limitation in the determination of a comprehensive cell lineage over the entire time course of embryogenesis. Although such a derivation of a comprehensive zebrafish cell lineage is not the topic of this study, we would like to provide an estimate of the tracking efficiency required for this objective. In order to obtain a cell lineage of the entire embryo with a fidelity of at least 90 % (i.e. no tracking error in 90 % of all cells), an average tracking mismatch rate of no more than 1 in 14,400 time points is required (the total number of time points multiplied by the inverse of the acceptable error fraction), corresponding to a tracking efficiency of 99.993 %.

Data analysis

The visualization of the virtual embryo, represented by the nuclei segmentation database, is facilitated by the module *clusterRendering*. This program provides options for a translation and rotation of the entire embryo as well as for a virtual sectioning. Thereby, the embryo can for example be reoriented with respect to the body axis and sectioned such that only nuclei in a slice through the body axis are visualized (see [Movies 14-16](#)). The processed database is converted into a POV-Ray script and rendered (POV-Ray 3.6, Persistence of Vision team), approximating the nuclei by spheres around the detected centre-of-mass coordinates and using an average radius computed from the nucleus morphology parameters (see above). The program also allows marking specific cells or entire tissues in different colours, thereby

highlighting cell migration and tissue formation. *clusterRendering* can be combined with the functionality of *clusterTracking*, *clusterMovement* and *clusterDivisions* in order to visualize the lineage of single tissues and organs or the three-dimensional pattern of cell-division in the entire embryo over time. The module *clusterCombination* facilitates the complete reconstruction of digital embryos, i.e. the fusion of two hemispheres, recorded from the same embryo in the DSLM multi-view imaging mode under 0° and 180° .

clusterCombination fuses two digital embryo hemispheres, obtained by recording the same embryo from two opposing directions. The program correlates the positions of all segmented nuclei in the overlap region of the two hemispheres (typically $700 \times 700 \times 100 \mu\text{m}^3$). Since the time shift between the recordings of the overlap region from the respective two angles is not larger than 10 seconds, any nuclear movements can be neglected. The optimal shift vector for the nuclear coordinates in the first hemisphere with respect to the coordinate system of the second hemisphere is defined as the average vector providing the highest nuclear pair count within spheres of $7 \mu\text{m}$ around the nuclei of the reference hemisphere. Although the embryo potentially drifts over the recording interval of 24 h, such a drift will always affect both hemispheres in the same way and thus a single shift vector is sufficient for the fusion of the entire data set. In a final fusion step, *clusterCombination* removes the duplicates in the combined data set by analyzing nuclear distances in the overlap region. Any pair of nuclei originating from different hemispheres and exhibiting centre-of-mass distances of less than $7 \mu\text{m}$ is considered as a data entry with redundancy and consequently one nucleus per pair is removed. The choice of $7 \mu\text{m}$ as the boundary distance reflects the fact that this value corresponds to the average nucleus diameter and thus characterizes the domain of physically impossible centre-of-mass distances for any two nuclei in the embryo (considering non-zero cytoplasmic fractions). While the fusion module does not introduce false positives to the digital embryo database, there is a small chance of arriving at false negatives (removal of correctly identified nuclei). This situation only occurs, if a pair of nuclei a) exhibits a distance of less than $7 \mu\text{m}$ and b) was not detected at the same time in a single hemisphere. Manual inspection indicates that the probability of requirement (a) is smaller than 1 %, while the probability of requirement (b) results mathematically as less than 0.4 % (see above for the determination of the segmentation efficiency). Thus, less than 1 in 10,000 correctly segmented nuclei is discarded by the fusion module and the previous quantification of efficiency levels can also be directly applied to multi-view reconstructions.

clusterTracking performs the reverse engineering of cell lineages. To provide a seeding data set, the user simply picks the cells contributing to the tissue or organ of interest at an arbitrary time point (using the GUIs *pickNuclei*, *trackNuclei* or *pickVolume*). Starting with this time point and moving backwards in time, the program then correlates the spatial coordinates of the nuclei and tracks the identities. If a nucleus cannot be identified at a specific time point, the algorithm attempts to find a match at the previous time point. If a match cannot be obtained for this time point either, the corresponding lineage is flagged as being discontinued/unknown.

clusterMovement performs long-term tracking of all nuclei. The program correlates the movement vectors with the positions of the nuclei relative to the centre of the yolk cell.

Thereby, the movement vectors are resolved into their components directed towards the vegetal pole, the animal pole, the body's symmetry axis and the yolk cell's surface. The resulting sub-populations of nuclei with common movement patterns are statistically analyzed over time and flagged in a movement database for subsequent visualization by *clusterRendering*.

As an alternative to *clusterMovement*, collective cell migration behaviour and migration vectors can also be quantified and visualized by *clusterField*. This program operates directly on the microscopy data and determines local shift vectors by correlating the gray-level information in small data slabs. The shift vectors represent the average migration vectors of the nuclei in the respective slabs. Three different arrow sizes are employed for data visualization, expressing the level of collective cell migration. A large arrow indicates that at least 75 % of the average migration vectors in the neighbourhood of a slab point in the same direction (within a 20 % margin). A medium-sized arrow indicates a 50-75 % match in migration directionality and a small arrow corresponds to a less than 50 % overlap, suggesting non-collective cell migration in that region. The quantitative global cell movement analyses in [Movies 9, 14](#) and [15](#) were performed with *clusterMovement*.

The data analysis module *clusterMaps* generates three-dimensional nuclei density maps and cell migration maps from the segmentation database and operates on principles related to *clusterTracking*. The cell identities, however, are only determined for subsequent time points and not for the entire time-course of development. A matching of nucleus identities is accomplished by evaluating the correlation of spatial coordinates, nuclei morphologies and intensity levels. Based on the detected identities, the average cell migration speed is resolved spatially in three dimensions for the entire embryo and converted into a three-dimensional map of migrations speeds.

A population-statistical analysis of the entire embryo is performed by *clusterStatistics*. The program determines the number of nuclei in the embryo, the distribution of nuclei intensity levels, the distribution of nuclei migration speeds and the distribution of chromatin domain sizes as a function of time. Migration speeds are computed as in *clusterMaps*. The chromatin domain size is approximated by the volume of an ellipsoid model of the nucleus morphology.

clusterDivisions detects cell divisions in the digital embryo and analyzes their global three-dimensional pattern as well as the polarity of the division axes. Potential candidates must fulfil several criteria in order to be scored as dividing nuclei. First, the module determines all pairs of nuclei that spatially correlate to the same nucleus identity in the previous time point. The potential daughter nuclei are tracked forwards in time for seven time points (i.e. seven minutes), while the mother nucleus is tracked backwards in time for seven time points. In order to be scored as a cell division event, the volume of the mother nucleus must be constantly decreasing over this time interval, while the volumes of both daughter nuclei must be constantly increasing, reflecting the transition from interphase to metaphase in the mother nucleus and the relaxation back to interphase in the daughter nuclei. The division axis is then defined as the vector from the first daughter nucleus to the second daughter nucleus in the first time point after the division. The test interval of seven minutes has been chosen as the

time interval that equals half of the cell cycle length during the shortest synchronous cell divisions in the cleavage stages of the embryo. It is therefore also the longest possible time interval, for which a constant decrease in the mother nucleus volume and a constant increase in the daughter nuclei volumes can be generally expected. Thus, the statistical strictness of the criterion underlying cell division detection is maximized without compromising the detection efficiency by introducing a high fraction of false negatives. We performed manual controls in the time interval 100-300 mpf and obtained a cell division detection efficiency of 60 % for this high-fidelity module. The fraction of false positives was ~1 %, while the remaining 39 % were accounted for by false negatives, i.e. unrecognized cell division events. It should be noted that the detection efficiency of cell division events is unrelated to the nucleus tracking efficiency (99.5 %), since these tracking modules rely on the concept of reverse engineering (nuclear tracking backwards in time) and thus do not require an initial detection of cell division events. *clusterDivisions* was designed exclusively for the study of cell divisions *per se* and is not recommended for the reconstruction of tissue formation or for cell lineaging (which should be performed with *clusterTracking*).

DSLIM sample preparation

Fixed Medaka embryos with acetylated tubulin immuno-staining

For the recordings of the central and peripheral nervous system, Medaka embryos aged 5 and 10 days post fertilization were stained using acetylated tubulin primary antibodies and Alexa Fluor 546 goat anti mouse IgG secondary antibody. The fixed whole mount embryos were embedded in low-melting-temperature agarose (SeaPlaque® GTG® agarose, Cambrex) at a concentration of 1.25 % in 1x ERM buffer, using 1 ml syringes (BD Plastipak™) as specimen chamber compartments. The embryos were transferred to the agarose-loaded syringes and oriented inside the compartment with dissection needles and forceps. For multi-view experiments the embryo body axis was oriented collinear with the rotational axis of the DSLM stage system, ensuring optimal imaging conditions independently of the rotation angle. For Medaka experiments, the DSLM specimen chamber was loaded with 12 ml 1x ERM buffer.

Preparation for in vivo zebrafish imaging

Zebrafish embryos were injected at the one-cell-stage (in average 15 minutes post fertilization) with 30 ng/μl H2B-eGFP mRNA.

Subsequently, the embryos were dechorionated in 20 mg/ml pronase in 30 % Danieau's solution and embedded in 1 % low-melting-temperature agarose and 30 % Danieau's solution as described above. For multi-view recordings of zebrafish embryogenesis, glass capillaries (Transferpettor caps with Teflon piston rods, Brand GmbH, Germany) were used instead of plastic syringes, allowing for multi-view rotation due to the smaller diameter. The agarose was allowed to polymerize for 5 minutes, after which the syringes/capillaries were transferred to a beaker with 30 % Danieau's solution to prevent agarose drying. The DSLM specimen chamber was loaded with 12 ml 30 % Danieau's solution and equilibrated to 26.5 °C via the temperature control system 1 h prior to the start of the experiment. The syringes/capillaries with the embryos were then mounted in the DSLM stage system. For single-view recordings of a specific hemisphere, the embryos were oriented such that the respective hemisphere faced the detection objective under an angle of 20° (tilted towards the light sheet), providing optimal imaging conditions.

Controls for physiological development

After imaging (typically at 36 hpf), zebrafish embryos were removed from the DSLM agarose sample cylinder with forceps and incubated for one week at 26 °C in 30 % Danieau buffer. We verified physiological development for all recorded embryos.

The public zebrafish digital embryo database

We provide public access to all digital embryos reconstructed and presented in this study. The database provides the three-dimensional positions, shapes and fluorescence intensity levels of in total 55 million nuclei, obtained from six wild-type embryos and one mutant *MZoep* embryo, as well as all movies associated with this study. The database can be accessed via the website <http://www.embl-heidelberg.de/digitalembryo>.

Each digital embryo database is provided as a Matlab MAT-file, containing a multi-cell array (with each cell corresponding to one time point) and a two-dimensional parameter table in each cell characterizing the nuclei segmented at the respective time point. Each table has n rows (n being the number of nuclei at the respective time point) and 17 columns, which contain the following parameters:

Columns 1-3: x-y-z coordinates of the centre-of-mass of the nucleus

Columns 4/5: Average (column 4) and centre-of-mass (column 5) fluorescence intensity (in gray levels) of the nucleus. The average is formed for the central 10-200 voxels of the nucleus (scaling linearly with total nucleus size).

Columns 6-11: x-y-z start and end coordinates of the bounding box of the segmented nucleus core (from 6 to 11: x-start, y-start, z-start, x-end, y-end, z-end).

Columns 12-17: x-y-z FWHM coordinates of the three-dimensional intensity distribution around the centre-of-mass of the nucleus (considering the background level). These parameters indicate at which pixel positions along all three dimensions the fluorescence intensity first drops below 50 % of the level at the centre-of-mass position. For more than 97 % of all nuclei, these parameters yield a good approximation of the nucleus shape and the total volume. It should be noted, however, that at very late time points, some nuclei at the centre of the body axis are too densely packed to allow for a decrease in fluorescence of 50 % in between neighbouring nuclei. If the intensity spread could not be determined along one of the dimensions, the table indicates +Inf or -Inf values for the corresponding parameter.

Specification of the seven data sets provided in the public database:

- *Data set 1: Reconstruction of zebrafish wild-type development*
Number of time points and temporal sampling:
901 time points (100-1450 mpf), 90-sec-intervals
Number of nucleus entries: 9,613,526 (416 MB)
Related manuscript items: [Figures 2 and S6a](#); [Movies 2 and 3](#)
- *Data set 2: Reconstruction of the zebrafish animal hemisphere (embryo A)*
Number of time points and temporal sampling:
1456 time points (100-1555 mpf), 60-sec-intervals
Number of nucleus entries: 8,736,014 (431 MB)
Related manuscript items: [Figures 3, S3, S6b and S8](#); [Movies 7-11](#)
- *Data set 3: Reconstruction of the zebrafish animal hemisphere (embryo B)*
Number of time points and temporal sampling:
1501 time points (200-1700 mpf), 60-sec-intervals
Number of nucleus entries: 10,622,019 (518 MB)
- *Data set 4: Reconstruction of the zebrafish animal hemisphere (embryo C)*
Number of time points and temporal sampling:
821 time points (120-1350 mpf), 90-sec-intervals
Number of nucleus entries: 3,446,269 (177 MB)
- *Data set 5: Reconstruction of the zebrafish dorsal hemisphere*
Number of time points and temporal sampling:
1001 time points (300-1300 mpf), 60-sec-intervals
Number of nucleus entries: 5,629,925 (283 MB)
Related manuscript items: [Figures 5 and S9](#); [Movies 12-14, 16](#)
- *Data set 6: Reconstruction of the zebrafish ventral hemisphere*
Number of time points and temporal sampling:
1381 time points (230-1610 mpf), 60-sec-intervals
Number of nucleus entries: 6,981,065 (344 MB)
Related manuscript items: [Figures 5 and S9](#); [Movies 15 and 16](#)
- *Data set 7: Reconstruction of zebrafish MZoep mutant development*
Number of time points and temporal sampling:
697 time points (230-1274 mpf), 90-sec-intervals
Number of nucleus entries: 10,677,510 (554 MB)
Related manuscript items: [Movies 4-6](#)

Supplementary Figures

Figure S1: Blueprint of the Digital Scanned Laser Light Sheet Fluorescence Microscope

(a) Schematic illustration of the central opto-mechanics of the microscope, including the laser light source (1), the acousto-optical tunable filter (AOTF, 2), the laser scanner with the f-theta lens (3), the line-like laser beam illumination arm (4), the temperature-controlled specimen chamber (5), the detector housing with a beam-splitter cube (6) and two wide-field fluorescence detection arms with independently operated cameras (7).

(b) The technical blueprint shows the complete instrument, as described in the chapter “Digital Scanned Laser Light Sheet Fluorescence Microscopy”. Two helium-neon-lasers and a beam expander can be added to the illumination system, which provide two additional excitation wavelengths and further control over the light sheet thickness.

Figure S2: DSLM multi-view imaging of large specimen

(a) Maximum-intensity projections of a DSLM image stack of the head of a Medaka fish embryo with acetylated tubulin immuno-staining (providing a view of the fish’s nervous system). The image stack consists of 2,000 images and shows the forebrain, the midbrain, the hindbrain and the retina of the left eye. Detection lens: Carl Zeiss Plan-Apochromat 20x/1.0 W. Green/blue look-up-table. Scale-bar = 100 μm . The specimen was kindly provided by Lazaro Centanin (EMBL, Wittbrodt Group).

(b) Maximum-intensity projections of a DSLM multi-view data set, showing a 5 mm long Medaka juvenile with the same staining as in (a). In total, 24 angles were recorded. The maximum-intensity projection along each angle is based on 2048 images, i.e. in total 49,152 images were acquired from the specimen. Detection lens: Carl Zeiss Plan-Neofluar 2.5x/0.075 Air. Green/blue look-up-table. Scale-bar = 500 μm . The specimen was kindly provided by Lazaro Centanin (EMBL, Wittbrodt Group).

Figure S3: Statistical analysis of nuclear properties in the animal hemisphere

Nuclei counts (top), normalized core intensities (second from top), movement speeds (second from bottom) and chromatin domain sizes (bottom) for all nuclei on the animal hemisphere of the data set shown in [Movie 7](#). Segmented and expected nuclei counts during synchronous cell division stages are shown in black and red respectively (top, left), indicating an average segmentation efficiency of 95 % during early embryonic development. The measurement of nuclear intensities over time indicates that steady-state GFP concentrations are reached at approx. 12 hpf, due to the limited stability of the injected mRNA. However, despite continuous imaging for another 10 hours the fluorescence intensity levels remained constant. This indicates an effectively negligible photo-bleaching rate in DSLM high-speed live

imaging. The inset in the plot of the average nuclei movement speed shows an overlay of the speed and chromatin domain size, revealing a perfect anti-correlation until 3 hpf and demonstrating the good sub-cellular resolution of nuclear positions and shapes in the digital embryo. The maximum speed of nuclear movement is achieved during chromosome segregation in anaphase, while maximum domain sizes result from chromatin decondensation during interphase. Errors indicate the SD over the entire nuclear population in the embryo.

Figure S4: The image processing pipeline

The flow-diagram shows the core software modules of the image segmentation and data analysis pipeline, which were developed for the reconstruction of vertebrate embryonic development via nuclear markers. The algorithms are described in the Materials and Methods (chapter “Image processing and data analysis”). Most of the pre- and post-processing modules operate on a regular computer workstation linked to a low-cost mass-storage file server (orange). In contrast, the computation-intensive core modules of the segmentation pipeline demand large-scale computer clusters (European Molecular Biology Laboratory and Karlsruhe Institute of Technology, gray).

Figure S5: Pre-processing/visualization of DSLM recordings of zebrafish embryogenesis

(a) Maximum-intensity projection of a raw DSLM image stack of the nuclei in the animal hemisphere at 3 hpf. Carl Zeiss C-Apochromat 10x/0.45 W lens.

(b) Projection of the image stack in (a) after applying a gamma of 0.7. The dynamic range of the raw data is 14 bit, which allows us to extract precise information about the shape and position in particular also of those nuclei that either exhibit a weaker fluorescent signal or reside in a less well-exposed region on the hemisphere (i.e., the far left side in the raw image). However, the gamma-correction is only necessary to highlight those nuclei for the human observer (considering the intrinsic reduction of grayscale data to 8 bit on the computer screen) and is not applied in the image processing pipeline. Scale-bar = 100 μ m.

(c) Projection of the image stack in (a) after applying a three-dimensional Gauss convolution with a sigma of 0.7 pixels. The Gauss convolution is applied in the image processing pipeline to enhance the segmentation efficiency. Effectively, the convolution reduces noise and prevents segmentation artifacts in nuclei exhibiting a very weak signal.

(d) Projection of the image stack in (a) after applying a Lucy-Richardson deconvolution with 20 iterations. The DSLM data features an exceptionally high signal-to-noise ratio of 1000:1, rendering the data particularly amenable for image deconvolution. However, the deconvolution shown here is only used for advanced data visualisation and neither required nor applied in the image processing pipeline.

(e-h) Zoom-ins on the images shown in (a-d). The raw DSLM data are recorded at an image size of 2048 x 2048 pixels and therefore all Supplementary Movies were scaled down by a factor of four to reduce file size. This panel serves the purpose of showing the DSLM data in its original size and demonstrates the good sub-cellular resolution achieved in DSLM recordings of entire zebrafish embryos (see also [Figure 2e](#)). Scale-bar = 10 μm .

(i-l) Raw segmentation of the Gauss convolved DSLM data at different time points of embryogenesis, shown as maximum-intensity projections. The head region and the notochord (shown for the 18-h-old embryo) constitute particularly difficult segmentation scenarios owing to the high cell density and large nuclear fraction in the total cell volume. The core voxels of the segmented nuclei are shown and encoded in different colours to highlight different nucleus identities. As an example for the computational limitation in embryo coverage (providing 92 % of the nuclei in the digital embryo), the red arrows in the notochord panel indicate a region with low segmentation efficiency.

Figure S6: Global statistical analysis of cell divisions

(a) Global nuclear counts in the digital embryo ([Movies 2 and 3](#)) reveal two regimes with near-constant cell division rates (200-500 mpf and 550-750 mpf; mpf = minutes post fertilization). The latter regime overlaps precisely with the period of convergence in the embryonic development. The global cell proliferation rate in the embryo is increased by 70% during this interval as compared to the early build-up phase of the blastoderm (200-500 mpf).

(b) Histograms of cell division polarity for 4,500 cell divisions in the time interval 100-400 mpf. Angles α and β describe the orientation of the axes of division in three dimensions, as defined in [S7]. α represents the angle in the plane tangential to the yolk surface and is measured towards the line on the sphere's surface connecting the animal, vegetal, dorsal and ventral poles. β represents the angle in the plane normal to the yolk surface and is measured towards the normal vector to the yolk surface. The distribution of α is uniform, suggesting that cells do not divide along a preferential tangential axis during early embryogenesis before the onset of gastrulation. The histogram of β indicates that cells preferentially divide tangentially to the yolk surface rather than perpendicular to it.

Figure S7: Embryo-to-embryo variability of morphogenetic key parameters

Analysis of the embryo-to-embryo variability of the global cell migration speed (500-600 mpf and 1,200-1,300 mpf) and the chromatin domain size (1,300 mpf). The chromatin volumes were corrected for the DSLM point spread function. Assuming a spherical shape, a volume of $186 \mu\text{m}^3$ corresponds to a nuclear radius of $3.5 \mu\text{m}$. Black/white lines indicate mean values for individual/all embryos. Gray bars indicate SD within single embryos. Red bars correspond to SD of mean values.

Figure S8: Quantification of the properties of the peripheral cell division waves

Progression speeds of the peripheral (left) and radial waves (middle) as well as delay between radial and peripheral waves (right) for division cycles 7-12 (64 - 2000 cells) for the data set shown in [Movie 7](#). The time point of the first symmetry break is indicated in gray. Errors in the top and middle panels are indicated as SEM and calculated from the number of nuclei contributing to the respective waves. Errors in the bottom panel correspond to the uncertainty introduced by the temporal sampling in the measurements.

Figure S9: Analysis of the patterns of mesendoderm internalization on the dorsal and ventral hemispheres

(a) Internalization maps for the dorsal and ventral hemispheres. The graphic shows the internalization coordinates for each mesendoderm cell (represented as spheres, dorsal and ventral views on the embryo) and provides the respective time point of internalization in the spheres' colour code.

(b) Zoom-ins on the internalization maps close to the shield on the dorsal hemisphere and opposite of the shield on the ventral hemisphere. A comparison shows that while cell internalization is well-synchronized on the ventral hemisphere, the synchrony of the dorsal pattern is disrupted by asynchronous ingression/intercalation. Note the clear separation of colours (corresponding to the internalization time points) in the ventral part and the mixing of colours in the dorsal hemisphere.

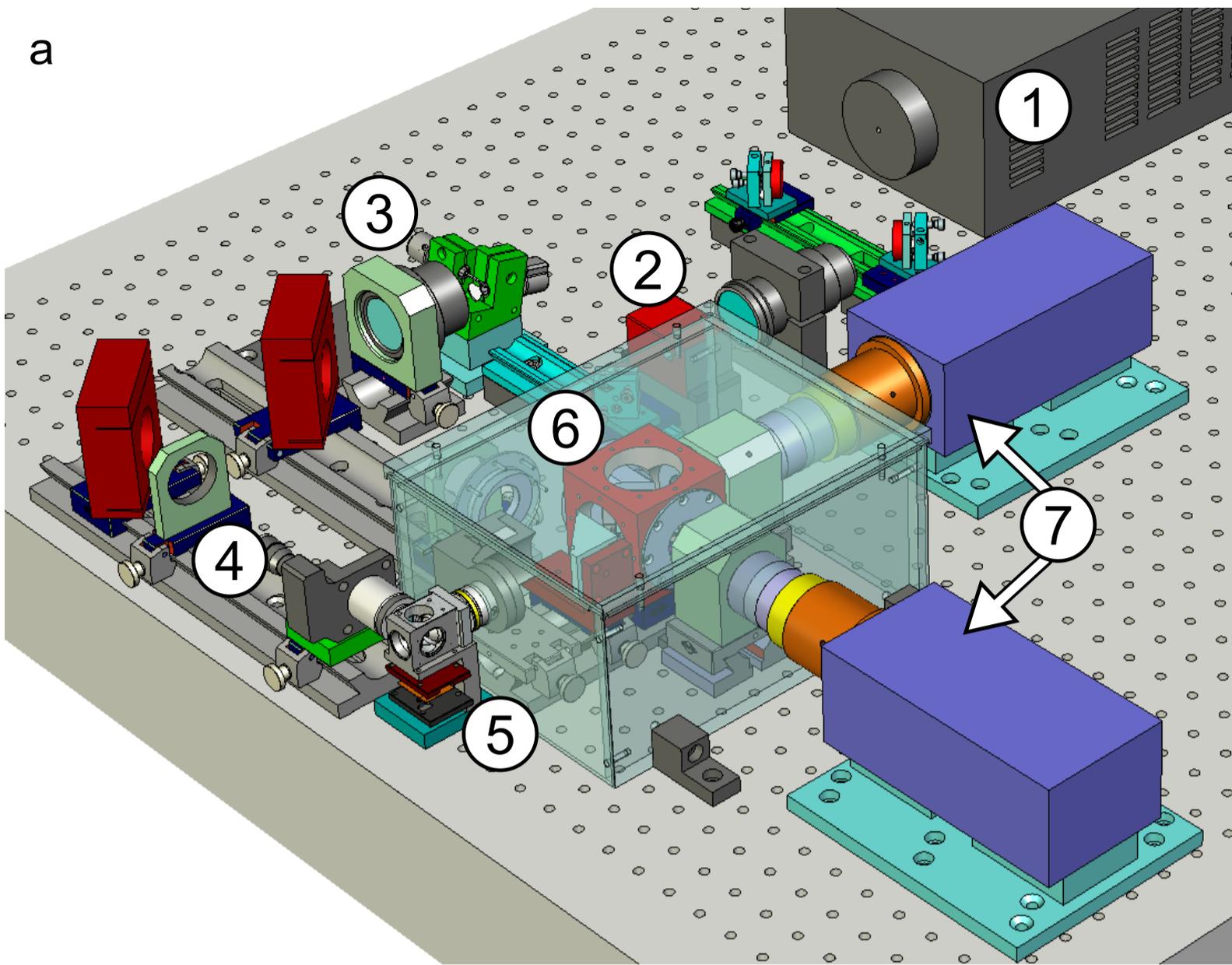
(c) Cell migration directions in the ventral and dorsal hemispheres during the synchronized cell internalization wave (emboly) and two hours later. Cell internalization occurs along the entire blastopore prior to 50 % epiboly, i.e. only in the animal hemisphere. See [Movies 14](#) and [15](#) for the complete reconstruction of cell migration during hypoblast formation.

(d) Cell internalization counts in the digital embryo between 6 hpf and 10 hpf. The embolic wave is synchronized in the dorsal (red) and ventral (blue) hemispheres and lasts approximately one hour. One third of all cells internalize and form the hypoblast (1,550 cells), while the remaining two thirds form the epiblast (3,000 cells).

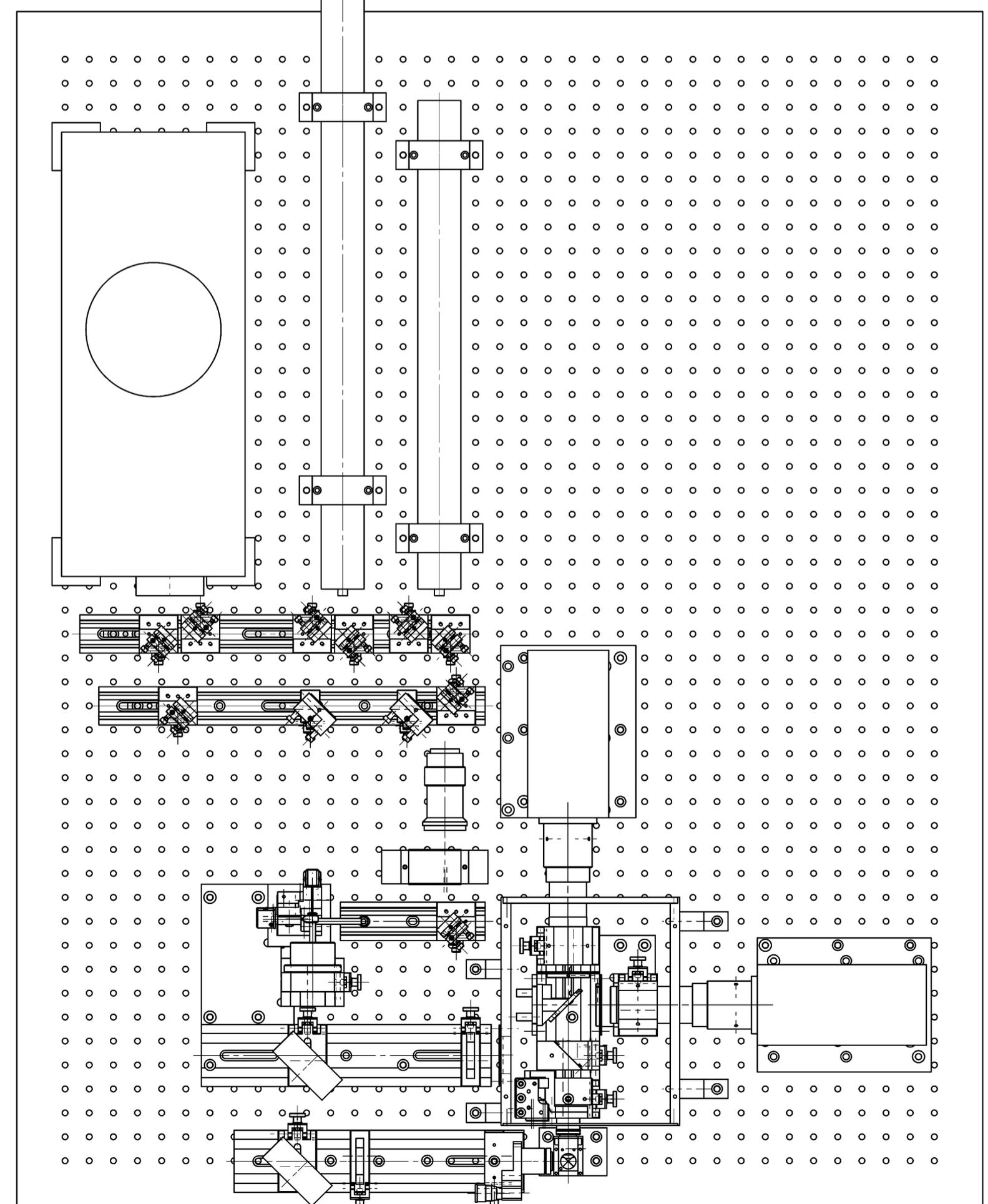
(e) "Internalization profiles" of the dorsal and ventral hemispheres, as well as of a Monte-Carlo simulation of tissue involution. The graphs show the average time difference in internalization of cell pairs as a function of the distance between the respective cell pairs. The statistics underlying the graphs consider all pairings of the cells provided by the internalization maps (a). While the graphs indicate a factor of 2-3 higher synchrony between neighbouring internalizing cells as compared to cell pairs that are further apart in the embolic wave, the global level of synchrony is by up to a factor of 2 higher on the ventral hemisphere as compared to the dorsal hemisphere. The Monte-Carlo scenario simulates synchronized tissue involution (or "synchronized ingression") in one embryonic hemisphere and considers 800 cell internalization events within a time interval of 90 minutes. These parameters were obtained from the experimental data of a wild-type zebrafish embryo and were subjected to

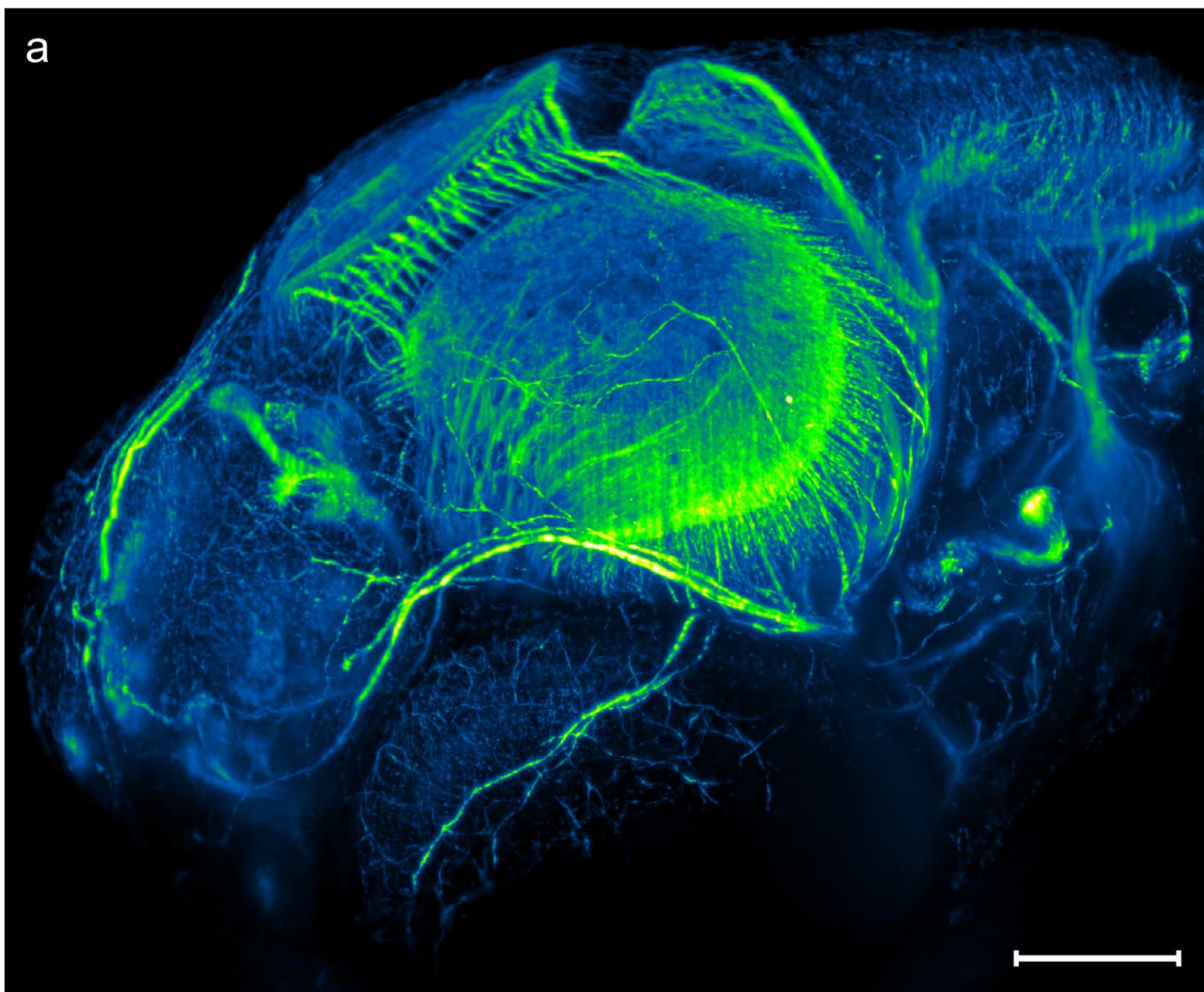
an average noise level of 15 minutes temporally and 7 μm spatially. Importantly, the simulated profile for involution matches the experimental profile of the ventral hemisphere quite well, but deviates considerably from the dorsal profile. We conclude that mesendoderm internalization occurs essentially via tissue involution/synchronized ingression ventrally and by mixed modes of internalization dorsally (involution, disrupted by independent ingression and radial intercalation). SEM is smaller than the plotted circles.

a

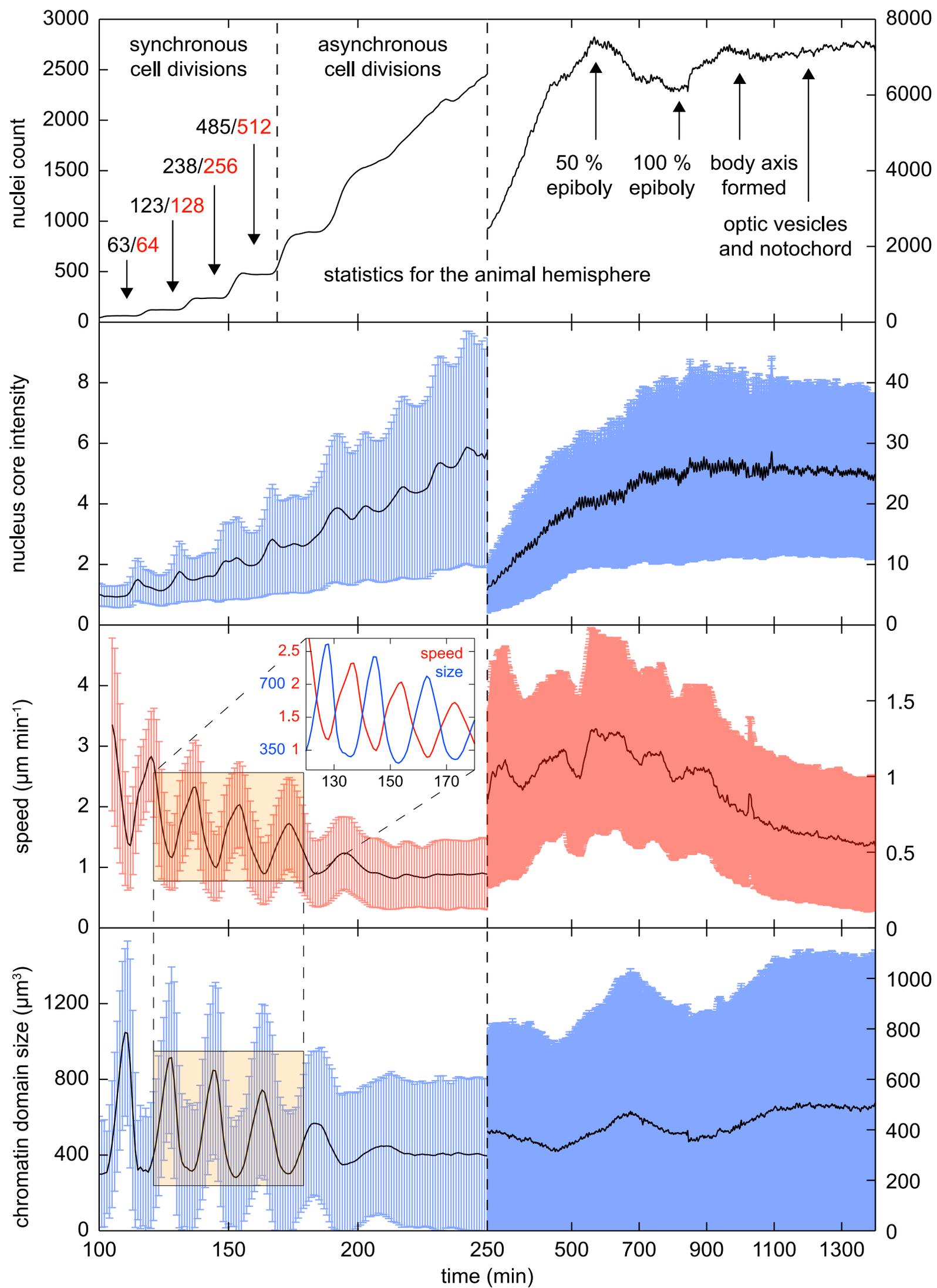


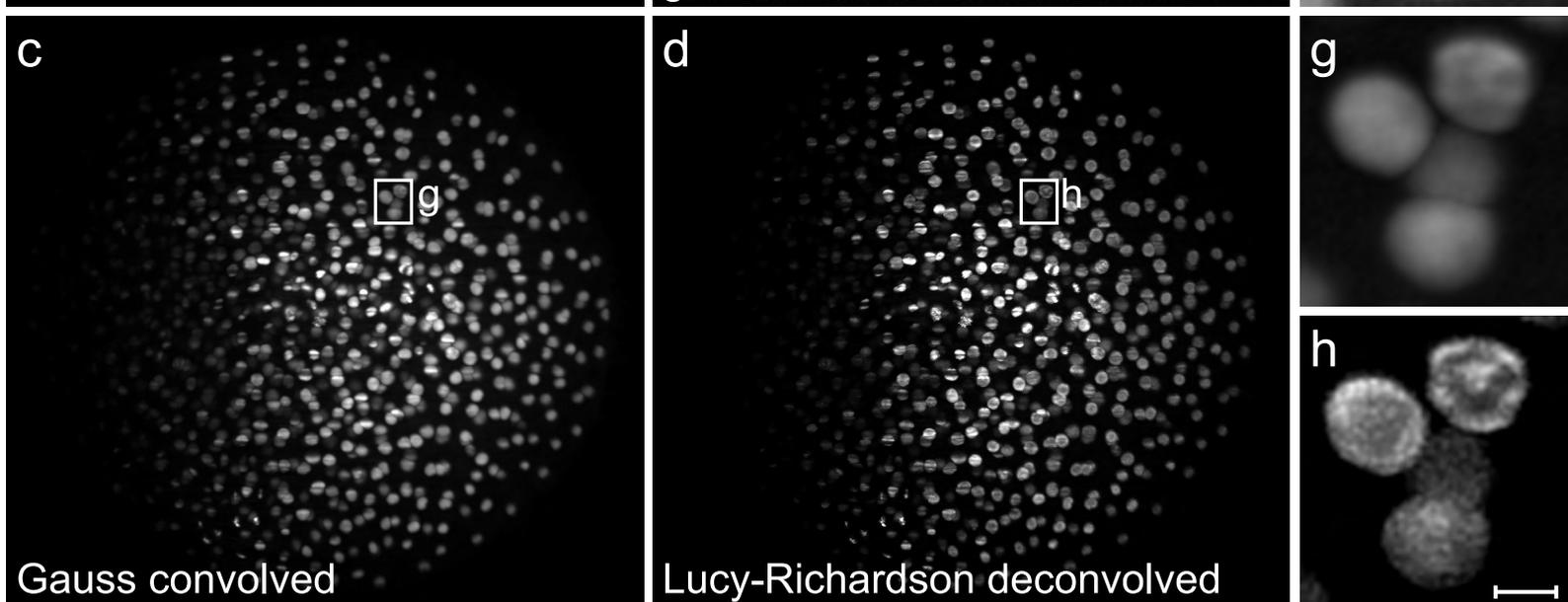
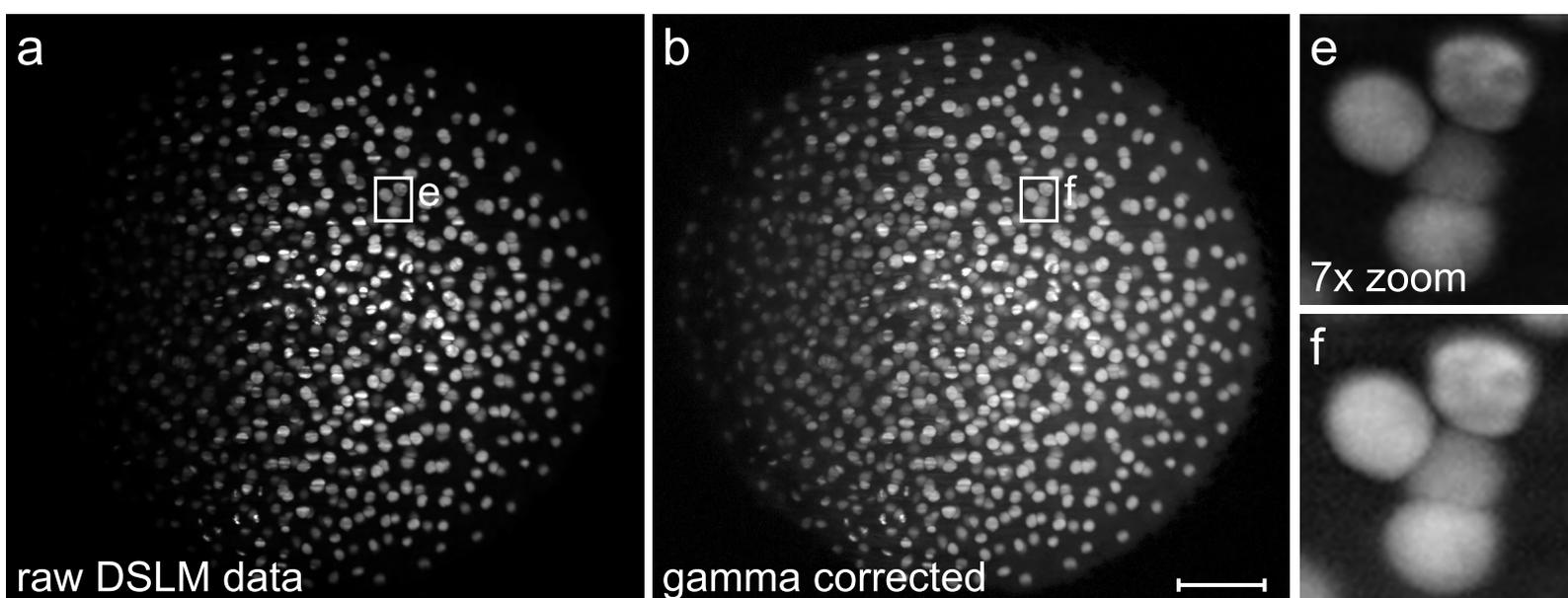
b





Keller *et al.* (2008); supplementary figure 3



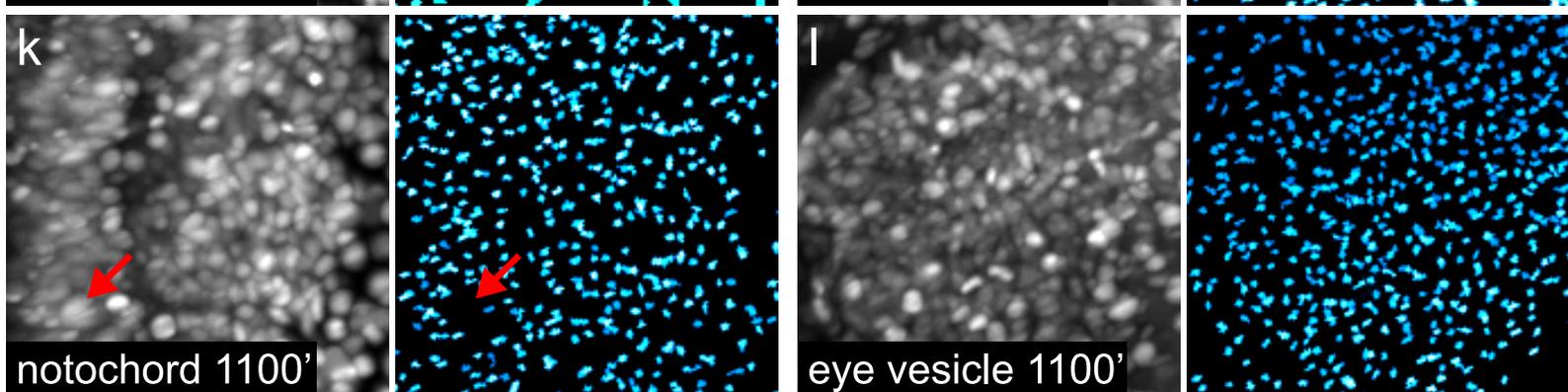
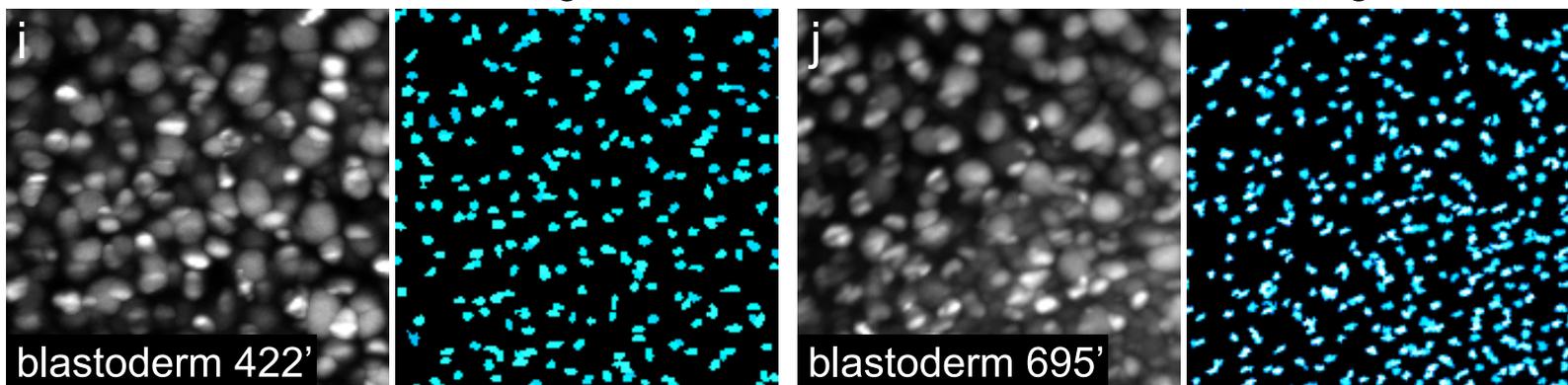


Gauss convolved

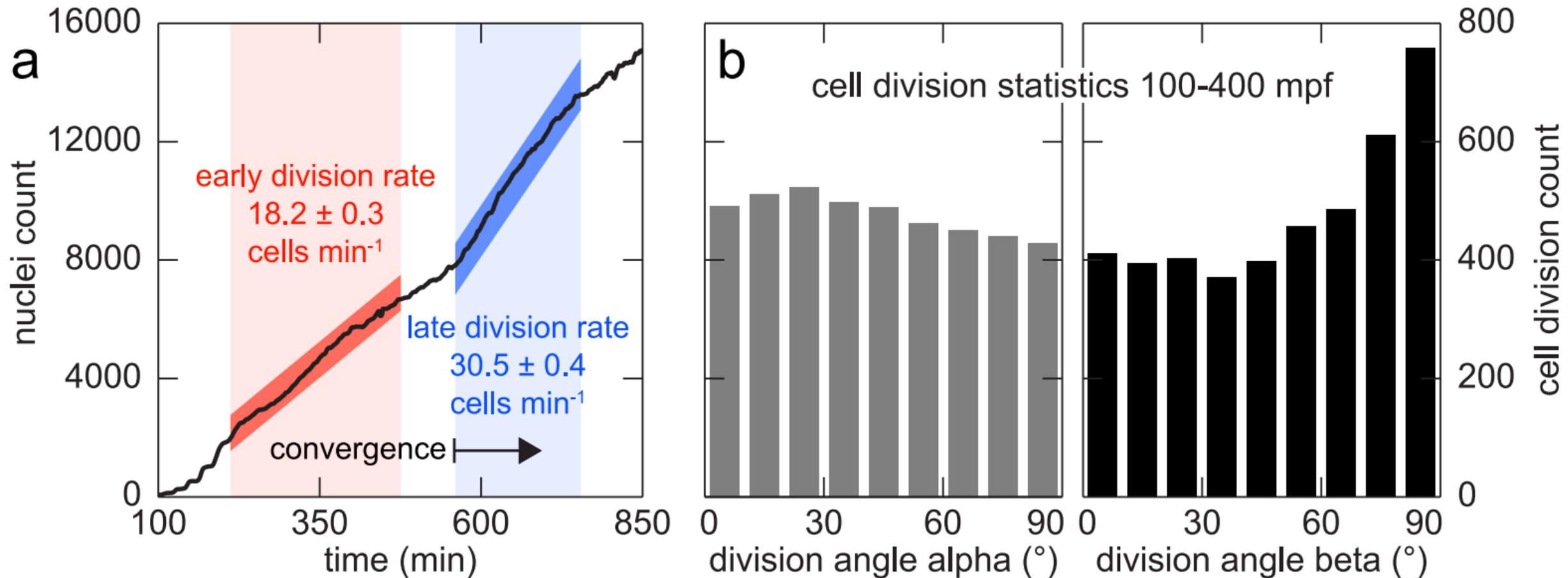
Core segmentation

Gauss convolved

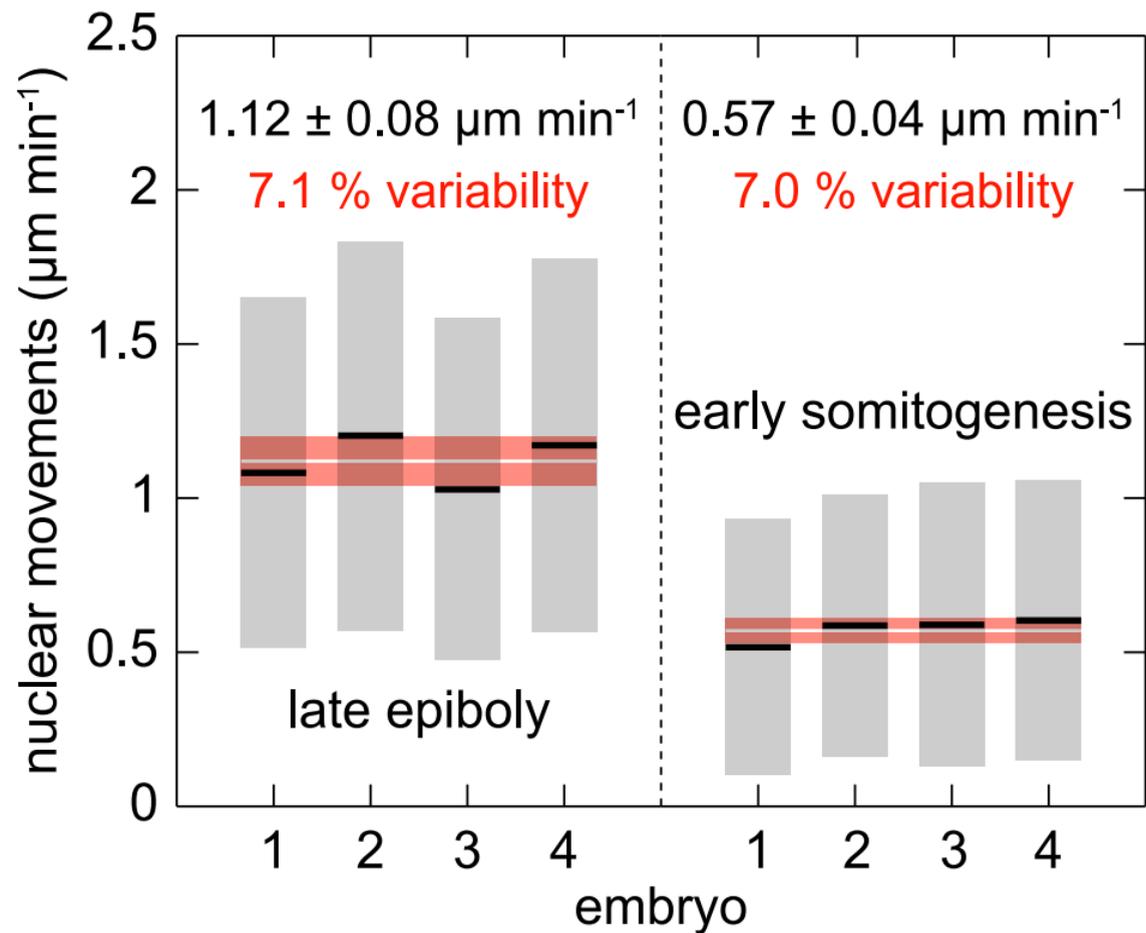
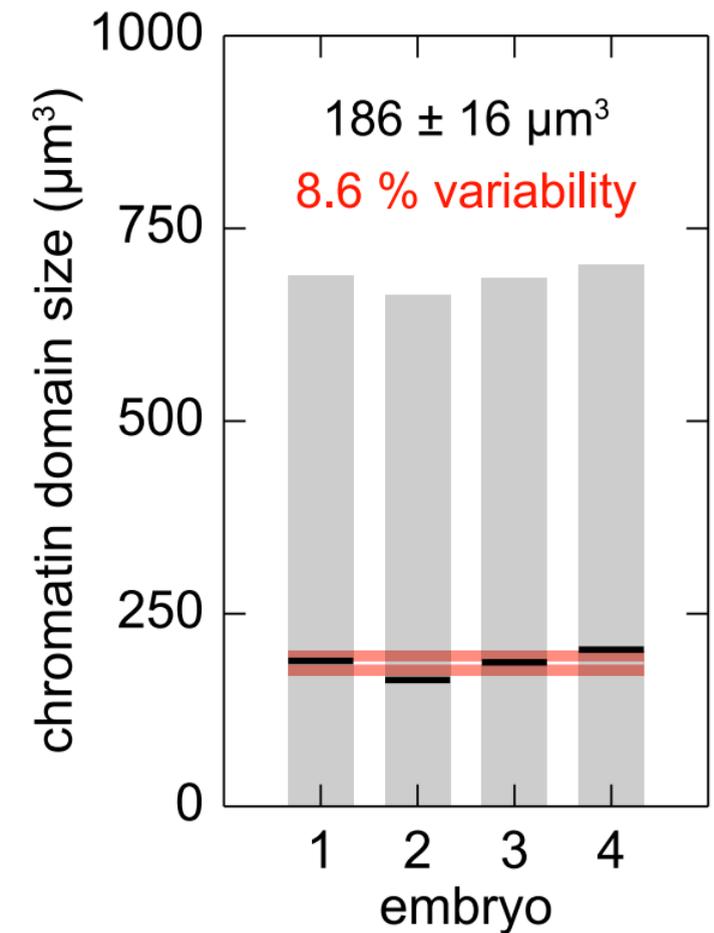
Core segmentation



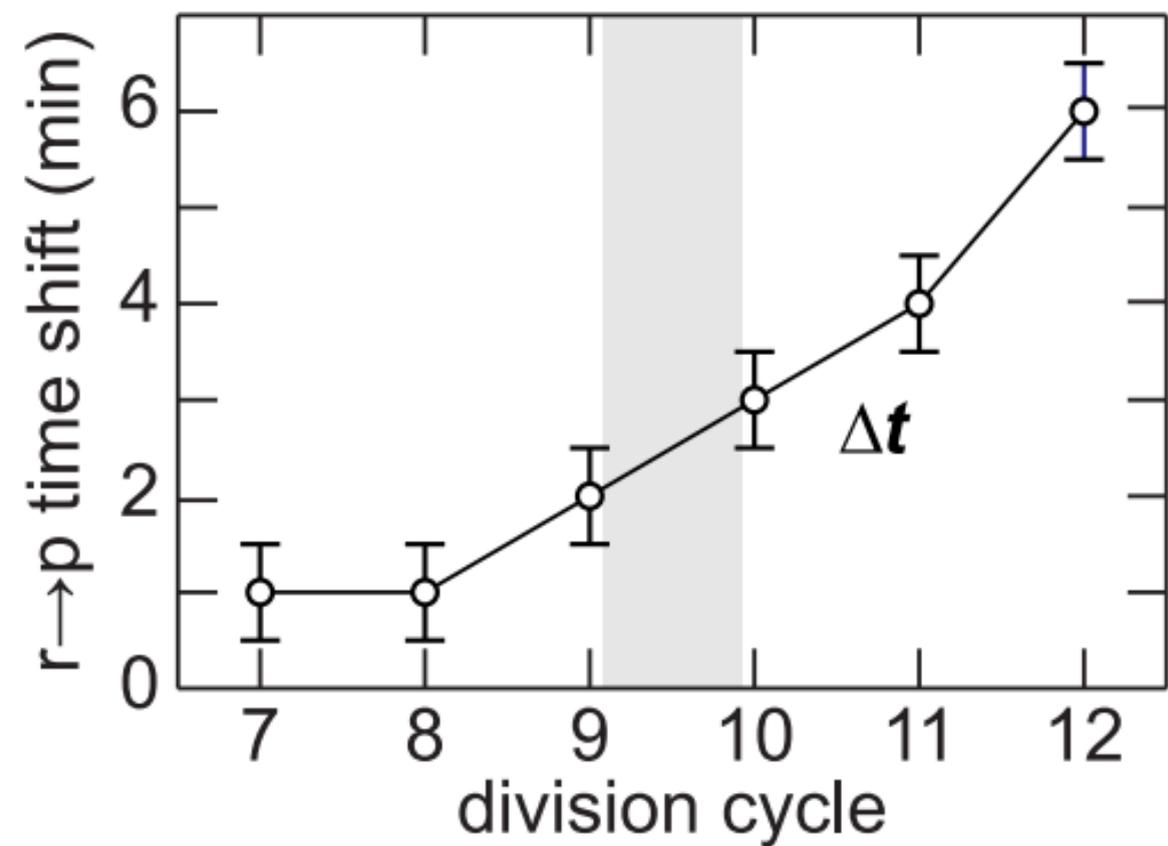
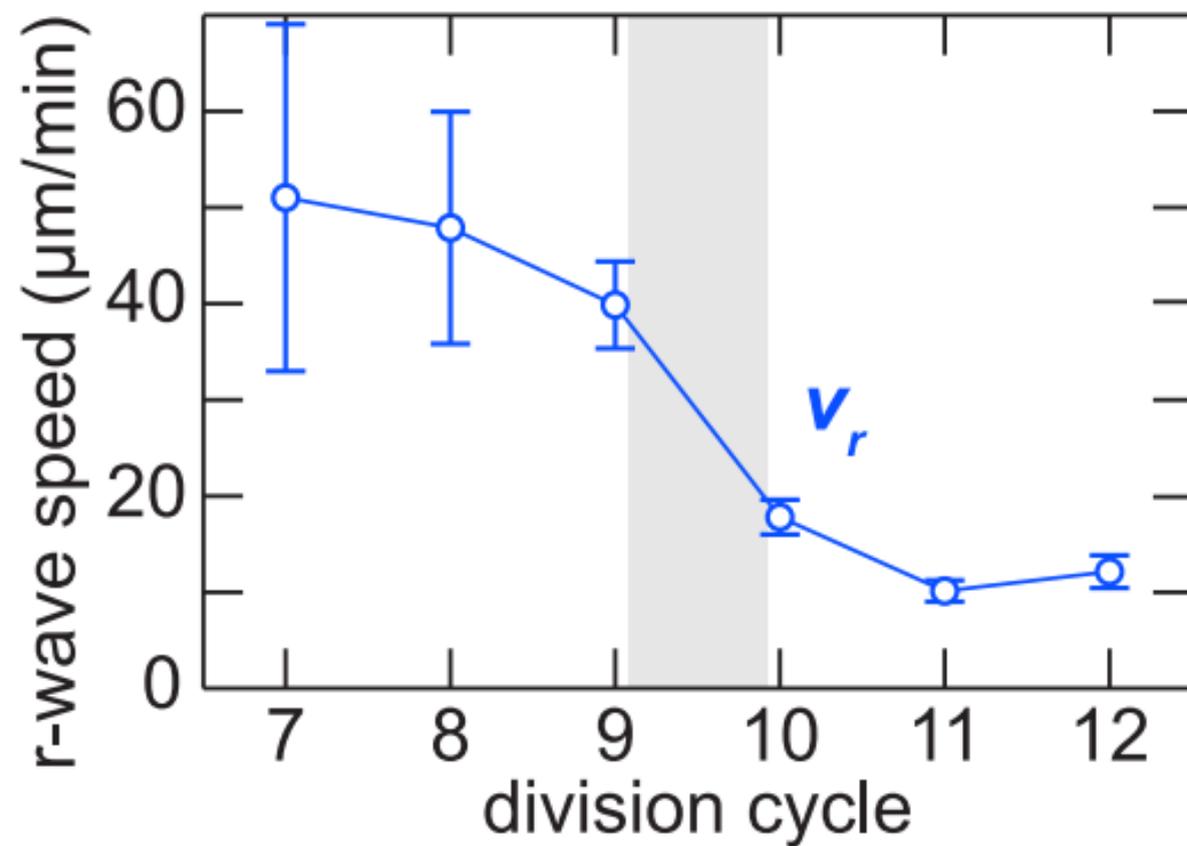
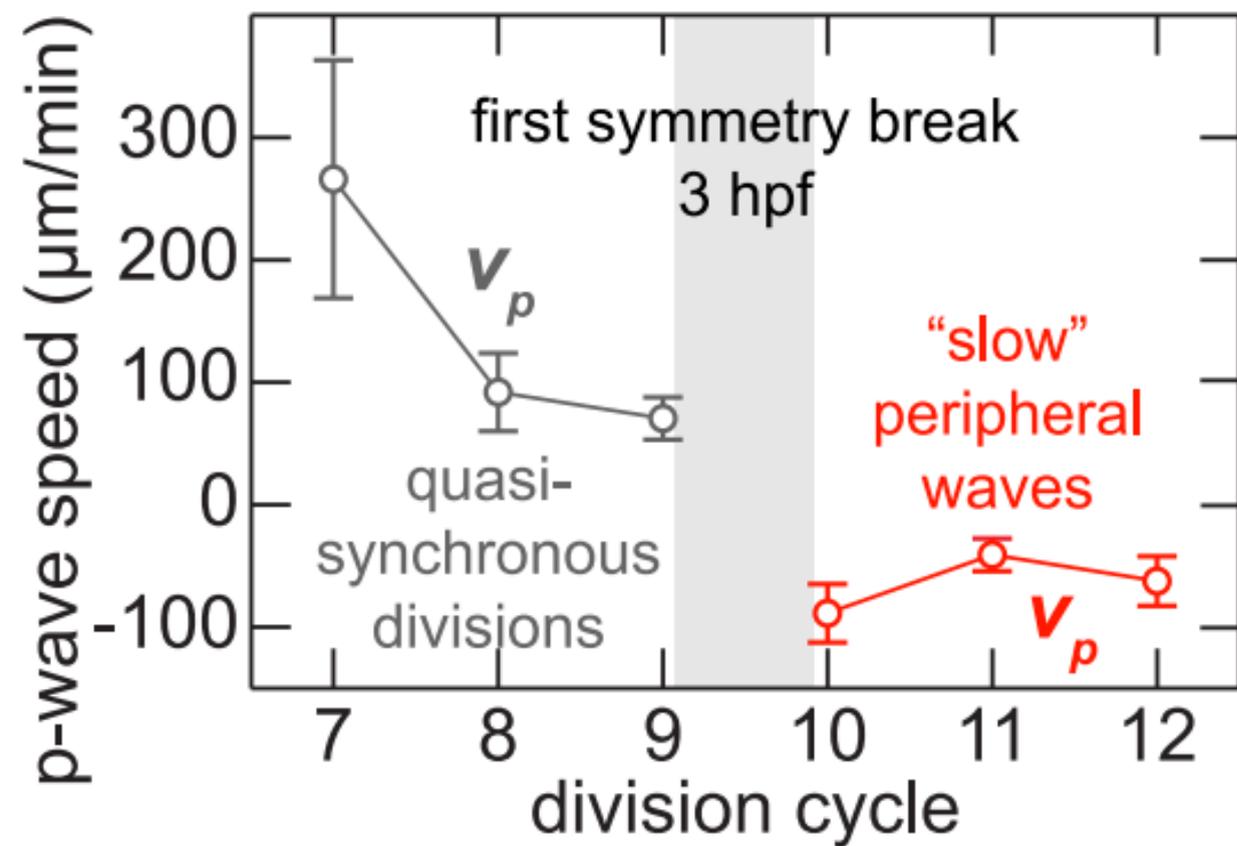
Keller *et al.* (2008); supplementary figure 6



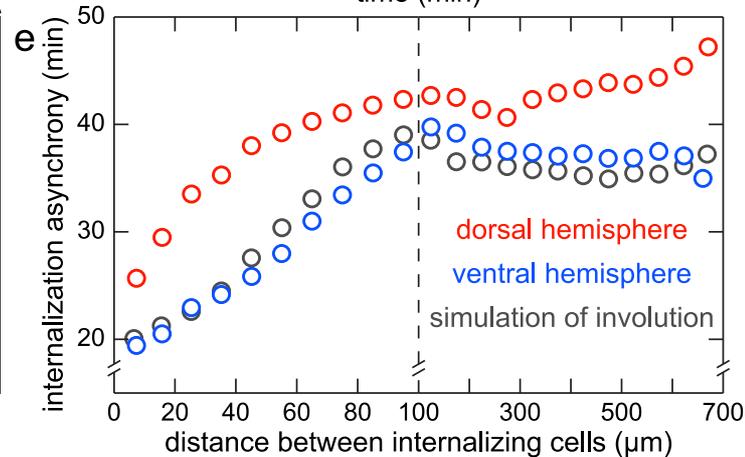
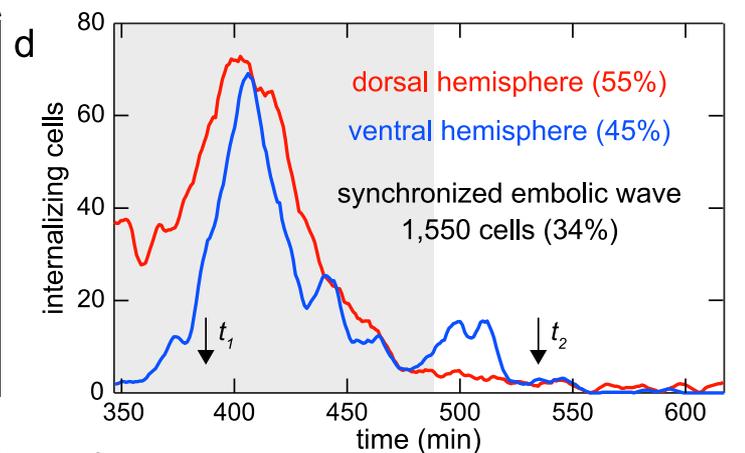
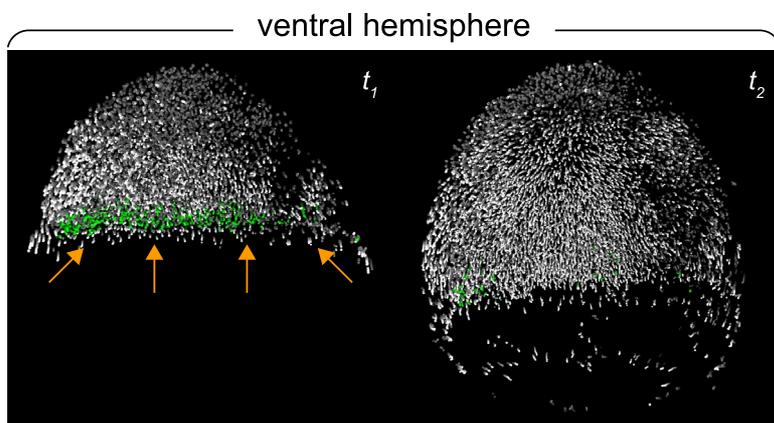
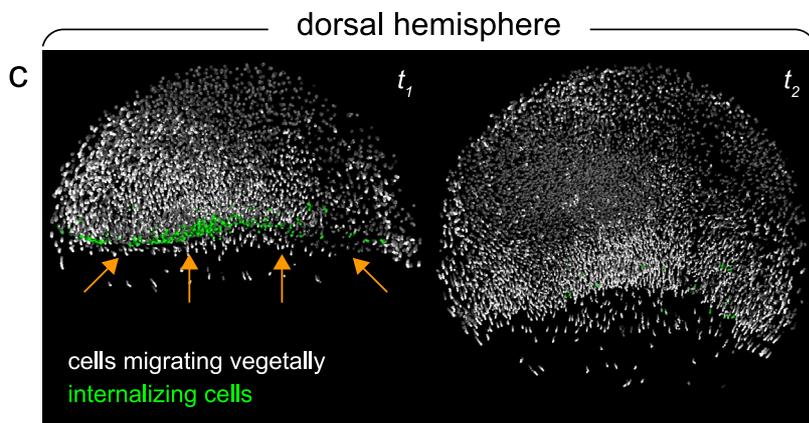
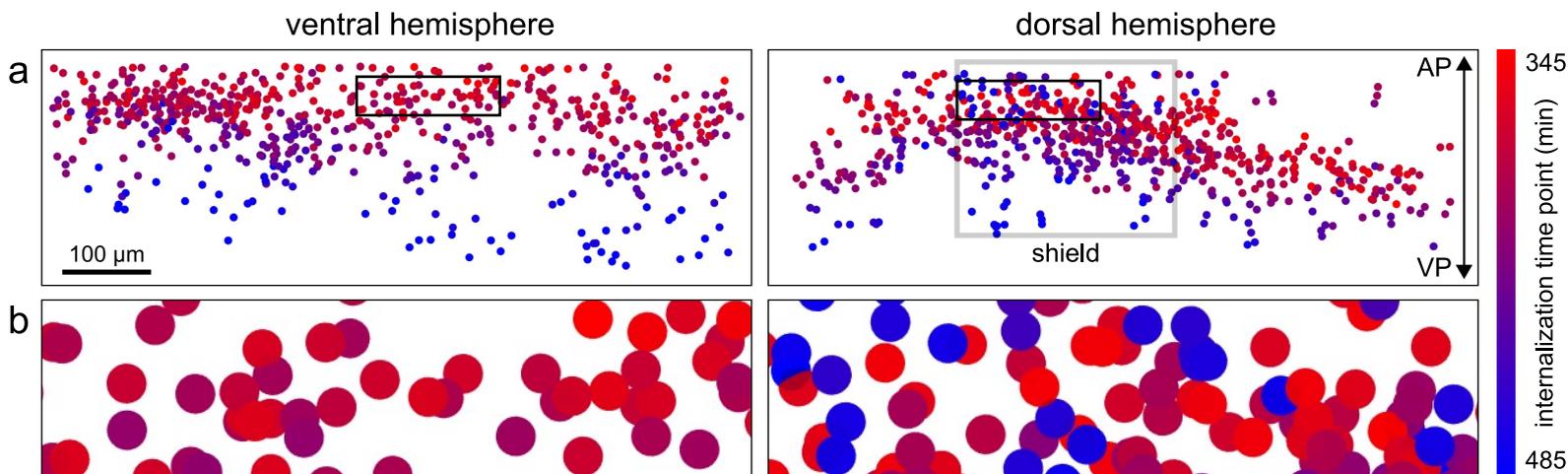
Keller *et al.* (2008); supplementary figure 7



Keller *et al.* (2008); supplementary figure 8



Keller *et al.* (2008); supplementary figure 9



Movies

Note: All movies can be downloaded as Apple QuickTime and DiVX versions on the public digital embryo repository website: <http://www.embl-heidelberg.de/digitalembryo>.

Movie 1: Schematic operation principle of DSLM

The movie shows a schematic illustration of specimen illumination and fluorescence light detection in the standard DSLM imaging mode. The laser beam illuminates the sample from the side and excites fluorophores along a single line inside the specimen. A pair of laser scanners moves the excitation laser line vertically and horizontally. An optically sectioned image is recorded by rapidly scanning an entire plane in the specimen and detecting the fluorescence at a right angle to the illumination axis. The top left inset shows a zoom-in on the illuminated fraction of the specimen, which is embedded in an agarose cylinder and placed inside the specimen chamber.

Movie 2: DSLM multi-view time-lapse recording of zebrafish embryonic development

The movie shows maximum-intensity projections of a DSLM time-lapse multi-view recording of zebrafish embryonic development, with a view on both the animal and vegetal hemispheres. The wild-type zebrafish embryo was injected with H2B-eGFP mRNA at the one cell stage. Imaging was started in the 64-cell stage. The entire data set consists of 1,226 time points, recorded in 90 s intervals. At each time point, 370 images with 2048 x 2048 pixels each were recorded with a z-spacing of 2.96 μm . The entire data set comprises of 453,620 images (3.5 TBytes). The images were deconvolved with the Lucy-Richardson-algorithm (10 iterations) and normalized for constant overall intensity, thus compensating for the increasing intensity over time due to GFP expression. Detection lens: Carl Zeiss C-Apochromat 10x/0.45 W.

Movie 3: Reconstruction of zebrafish embryonic development

The movie shows an orthographic rendering of the digital embryo obtained from the microscopy data shown in [Movie 2](#). In each frame, the reconstructions at ten time points, corresponding to a time interval of 15 min, are superimposed to indicate the nuclear movement directions. The darker shading corresponds to older time points. The colour-code indicates the current movement speed of each nucleus. A bright orange corresponds to an average movement speed of 1.2 $\mu\text{m}/\text{min}$ or more, while a bright cyan highlights effectively non-migratory nuclei. Intermediate speeds are shown by linear interpolation along the colour gradient.

Movie 4: DSLM multi-view time-lapse recording of *MZoep* mutant embryogenesis

The movie shows maximum-intensity projections of a DSLM time-lapse multi-view recording of zebrafish *MZoep* mutant embryonic development, with a view on the dorsal and ventral hemispheres. Preparation and microscopy settings as in [Movie 2](#).

Movie 5: Reconstruction of *MZoep* mutant embryogenesis

The movie shows an orthographic rendering of the *MZoep* mutant digital embryo obtained from the microscopy data shown in [Movie 4](#), with a dorsal view on the left and a lateral view on the right. Visualization settings as in [Movie 3](#).

Movie 6: Digital visualization of the *MZoep* mutant defect in mesendoderm formation

The movie shows cell tracking in a slice through the shield for the microscopy time-lapse data set presented in [Movie 4](#) (left: speed colour-code, right: directional colour-code). The 100- μm -thick slices are shown from a lateral view, with the dorsal side to the right. Speed colour-code as in [Movie 3](#). The nuclear movements were analyzed for 30 min intervals and colour-encoded as follows (see also sphere at the bottom): anterior movements (green), posterior movements (cyan), movements away from the body axis (yellow), movements towards the body axis (red) and movements toward the yolk cell (pink). The corresponding colours were assigned, if a movement of at least 20 μm along the respective direction was determined within a time interval of 30 min.

Movie 7: DSLM time-lapse recording of zebrafish embryogenesis in the animal hemisphere

The movie shows maximum-intensity projections of a DSLM time-lapse recording of zebrafish embryonic development, with a view on the animal hemisphere. The wild-type zebrafish embryo was injected with H2B-eGFP mRNA at the one cell stage. Imaging was started at 1.6 hours post fertilization, during the transition from the 32- to the 64-cell stage. The entire data set consists of 1,470 time points, recorded in 60 s intervals. For each time point, 205 images with 2048 x 2048 pixels each were recorded within 30 s with a z-spacing of 2.96 μm (i.e. a 600 μm deep volume). The entire data set comprises of 301,350 images (2.3 TBytes). The images were deconvolved with the Lucy-Richardson-algorithm (20 iterations) and normalized for constant overall intensity, thus compensating for the increasing intensity over time due to GFP-production. Since the illumination laser beam has to travel through the yolk in order to excite fluorophores on the far left side of the animal hemisphere, the average nuclear brightness is lower in this part of the embryo. However, because of the excellent signal-to-noise ratio and the dynamic range of 14 bit of the DSLM images, the corresponding nuclei can also be robustly identified (see [Figure S5](#)). Detection lens: Carl Zeiss C-Apochromat 10x/0.45 W.

Movie 8: Reconstruction of zebrafish embryogenesis in the animal hemisphere

The movie shows a perspective rendering of the reconstruction as well as nuclei population statistics for the zebrafish wild-type embryo in [Movie 7](#). Left: view on the animal pole from the centre of the yolk cell. Right: lateral view. In each frame, the reconstructions at ten time points, corresponding to a time interval of 10 min, are superimposed to indicate the nuclear movement directions. The darker shading corresponds to older time points.

Movie 9: Global cell tracking and colour-indexing of morphogenetic domains

We tracked all nuclei reconstructed from the microscopy time-lapse data set shown in [Movie 7](#). The digital embryo is visualized as a perspective rendering with a view on the animal pole from the centre of the yolk cell. The nuclear movements were analyzed for 30 min intervals and colour-encoded as follows (see also sphere to bottom left): anterior movements (green), posterior movements (cyan), movements away from the body axis (yellow), movements towards the body axis (red) and movements toward the yolk cell (pink). The corresponding colours were assigned, if a movement of at least 20 μm along the respective direction was determined within a time interval of 30 min.

Movie 10: Global three-dimensional pattern of cell divisions

The movie shows 4,500 cell division events in the data set presented in [Movie 7](#), detected by the automated processing pipeline at an efficiency of over 90 % for the time interval 100-400 mpf (minutes post fertilization at 26.5 °C). Dividing cells are marked in red and their corresponding daughter cells are shown in cyan. The movie illustrates the initial radial division waves, followed by the symmetry break and by slow peripheral waves in division cycles 10-13. Beyond that time interval, cell divisions appear in a complex pattern of asynchronous patches. The statistical analysis of the division axes of these 4,500 cell divisions is shown in [Figure S6b](#).

Movie 11: Reverse engineering of migratory tracks of retinal progenitor cells

Visualization of the tracking of retinal progenitor cells (red nuclei) in the digital embryo (gray nuclei). The cells were identified by their spatial position in the last time point, i.e. when the optic vesicle had already formed. A reverse engineering of their migratory tracks over a time interval of six hours provides information about the morphogenetic blueprint of eye organogenesis. The automatically computed tracks are in good agreement with the manual reconstruction presented in [S8]. The movie illustrates the principle of organ lineaging using our digital embryos.

Movie 12: DSLM time-lapse recording of zebrafish embryogenesis in the dorsal hemisphere

The movie shows maximum-intensity projections of a DSLM time-lapse recording of zebrafish embryonic development, with a view on the embryonic shield. The wild-type zebrafish embryo was injected with H2B-eGFP mRNA at the one cell stage. Imaging was started 5 hours post fertilization at approximately 35 % epiboly. The entire data set consists of 1,000 time points, recorded in 60 s intervals. At each time point, 205 images with 2048 x 2048 pixels each were recorded within 30 s with a z-spacing of 2.96 μm (i.e. a 600 μm deep volume). The images were deconvolved with the Lucy-Richardson-algorithm (10 iterations) and normalized for constant overall intensity, thus compensating for the increasing intensity over time due to GFP-production. Detection lens: Carl Zeiss C-Apochromat 10x/0.45 W.

Movie 13: Reconstruction of zebrafish embryogenesis in the dorsal hemisphere

Reconstruction and nuclei population statistics for the embryo shown in [Movie 12](#). Settings as in [Movie 8](#).

Movie 14: Global cell tracking in the dorsal hemisphere

The movie shows global cell tracking for the microscopy time-lapse data set presented in [Movie 12](#). The parameter settings are identical to those used in [Movie 9](#). The nuclear movements are colour-encoded as follows: towards the vegetal pole (cyan), towards the animal pole (yellow), towards the body axis (green) and towards the yolk cell (pink).

Movie 15: Global cell tracking in the ventral hemisphere

Global cell tracking with parameter settings identical to those used in [Movie 9](#). The nuclear movements are colour-encoded as follows: towards the vegetal pole (cyan), towards the animal pole (yellow), towards the body axis (green) and towards the yolk cell (pink).

Movie 16: Side-by-side comparison of hypoblast formation in the dorsal/ventral hemispheres

Frontal and lateral view on embryonic slices on the dorsal and ventral hemispheres of the digital embryo (right: slice close to the shield, left: slice opposite of the shield; complete reconstructions are shown in [Movies 14](#) and [15](#)). The montage visualizes cell internalization in the form of a wave of mesendodermal progenitor cells folding under cell layers of ectodermal progenitors (emboly) between 40 and 50 % epiboly. In the reconstruction, four cell populations were tracked: green nuclei at the tip of the early wave, yellow nuclei in the late wave, blue nuclei at the leading edge moving towards the vegetal pole and pink nuclei close to the animal pole moving towards the vegetal pole. Internalized cells move towards the

animal pole on the dorsal hemisphere. In contrast, on the ventral hemisphere, internalized cells move only briefly towards the animal pole and then change their migration direction back towards the vegetal pole. A comprehensive morphogenetic model of hypoblast formation is provided in [Figure 6](#).

Supplementary References

- S1. Kubitschek, U., Kuckmann, O., Kues, T. and Peters, R.: Imaging and tracking of single GFP molecules in solution (2000). *Biophys J*, 78:2170-9.
- S2. Stelzer, E. H. K. et al.: Nonlinear absorption extends confocal fluorescence microscopy into the ultra-violet regime and confines the observation volume (1994). *Opt Comm*, 104:223-8.
- S3. Stelzer, E. H. K.: Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: fundamental limits to resolution in fluorescence light microscopy (1998). *J Microscopy*, 189:15-24.
- S4. Engelbrecht, C. J. and Stelzer, E. H. K.: Resolution enhancement in a light sheet-based microscope (SPIM) (2006). *Opt Lett*, 31:1477-9.
- S5. Verveer, P. J., Swoger, J., Pampaloni, F., Greger, K., Marcelllo, M. and Stelzer, E. H. K.: High-resolution three-dimensional imaging of large specimens with light sheet-based microscopy (2007). *Nat Methods*, 4:311-3.
- S6. Hale, G. M. and Querry, M. R.: Optical constants of water in the 200-nm to 200- μm wavelength region (1973). *Appl Opt*, 12:555-63.
- S7. Gong, Y., Mo, C., and Fraser, S.E.: Planar cell polarity signalling controls cell division orientation during zebrafish gastrulation (2004). *Nature*, 430:689-93.
- S8. Rembold, M., Loosli, F., Adams, R.J., and Wittbrodt, J.: Individual cell migration serves as the driving force for optic vesicle evagination (2006). *Science*, 313:1130-4.