CORRESPONDENCE

1. Stoeckius, M. *et al. Nat. Methods* **6**, 745–751 (2009).
2. Fernandez, A.G. & Piano, F. *Curr. Biol.* **16**, 1757–1763 (2006).
3. Galy, V., Askjaer, P., Franz, C., Lopez-Iglesias, C. & Mattaj, I.W. *Curr. Biol.* **16**, 1748–1756 (2006).
4. Rasala, B.A., Orjalo, A.V., Shen, Z., Briggs, S. & Forbes, D.J. *Proc. Natl. Acad. Sci. USA* **103**, 17801–17806 (2006).
5. Franz, C. *et al. EMBO Rep.* **8**, 165–172 (2007).
6. Edgley, M.K., Baillie, D.L., Riddle, D.L. & Rose, A.M. in *WormBook* (The *C. elegans* research community, 2006).
7. Kamath, R.S. *et al. Nature* **421**, 231–237 (2003).

Software for bead-based registration of selective plane illumination microscopy data

To the Editor: Selective plane illumination microscopy (SPIM)[1] allows isotropic, time-lapse, *in toto* imaging of large, living biological specimens by acquiring three-dimensional (3D) images of the same sample from multiple angles (views). However, to realize the potential of multiview SPIM imaging, it is necessary to reconstruct a single 3D image from the individual views[2] (**Fig. 1a**, **Supplementary Fig. 1** and **Supplementary Video 1**).

SPIM multiview registration is complicated by degradation of the signal along the illumination as well as detection axes (**Fig. 1b**), limited overlap between the views, different orientations of the optical sections and development of the specimen during acquisition. We developed a SPIM registration method and implemented it in a plugin for Fiji. The software enables efficient, sample-independent registration of multiview SPIM acquisitions using fluorescent beads in rigid mounting medium as fiduciary markers.

We first detected the beads with subpixel accuracy using a difference of Gaussian[3] filter, reducing the registration problem to the matching of point clouds. To efficiently identify corresponding beads in different views, we developed a translation and rotation invariant local geometric descriptor (**Fig. 1c**) that identifies each bead by the unique constellation of its neighboring beads. This constellation is preserved across views transformed by rotation and translation in three dimensions. For efficient matching, we defined an orthogonal local coordinate system (**Supplementary Fig. 2**) in each descriptor, expressing the 3D constellation of four beads by a vector of six scalar values, achieving translation and rotation invariance. Similar descriptors in different views have a small Euclidean distance in the six-dimensional descriptor space, and for efficient identification of nearest neighbors we presorted the six-dimensional scalar vectors using a hierarchical tree-based algorithm to reduce the matching problem to logarithmic complexity. Constellations of four beads that accidentally look similar (false correspondences) are rejected using the random sample consensus[4] on an affine transformation model followed by a robust regression filter (**Supplementary Methods**).

In the final step of the registration framework, we globally minimized the displacement of all true correspondences identified in all pairs of views using an iterative optimization scheme resulting in an affine transformation model for each view (**Fig. 1d**, **Supplementary Fig. 3** and **Supplementary Videos 2** and **3**). Typically, we identified thousands of corresponding bead descriptors equally distributed around the imaged sample, and the global optimization converged within seconds to a final average displacement of about one pixel

Figure 1 | Bead-based registration framework. (**a**) Several stacks of two-dimensional images of the same specimen acquired from different views have to be registered to obtain a single 3D image. (**b**) Three SPIM sections of *Drosophila* embryo stained with nuclear marker show the deterioration of the fluorescence signal along the illumination and detection axes. (**c**) Four color-coded examples of 3D constellations of four beads (central bead and its three nearest neighbors forming a bead descriptor) used to identify corresponding beads in different views (blue lines show view boundary in three dimensions, and gray circles represent the beads). (**d**) A 3D visualization of the global optimization progress on eight SPIM views of fixed *Caenorhabditis elegans* worm. Displacement of corresponding bead descriptors is color-coded from red (maximum displacement) to green (minimal displacement). The global optimization is initialized with all views on top of each other. Three iterations (0, 10 and 283) are shown along with average displacement across all descriptors. (**e–i**) Sections through living *Drosophila* embryo expressing His-YFP in all cells; imaged and reconstructed from seven SPIM views (bottom) compared to single SPIM view (top). Single-view acquisitions were stopped approximately in the middle of the embryo to avoid optical aberrations resulting from light scattering and to speed up the acquisition. The lateral resolution in the reconstructed multiview image (**e**) is comparable with the axial resolution (**f**) and is superior to the resolution of the single view. *y-z* sections (**g–i**) at positions marked by white arrowheads in **f**. Scale bars, 50 μm.
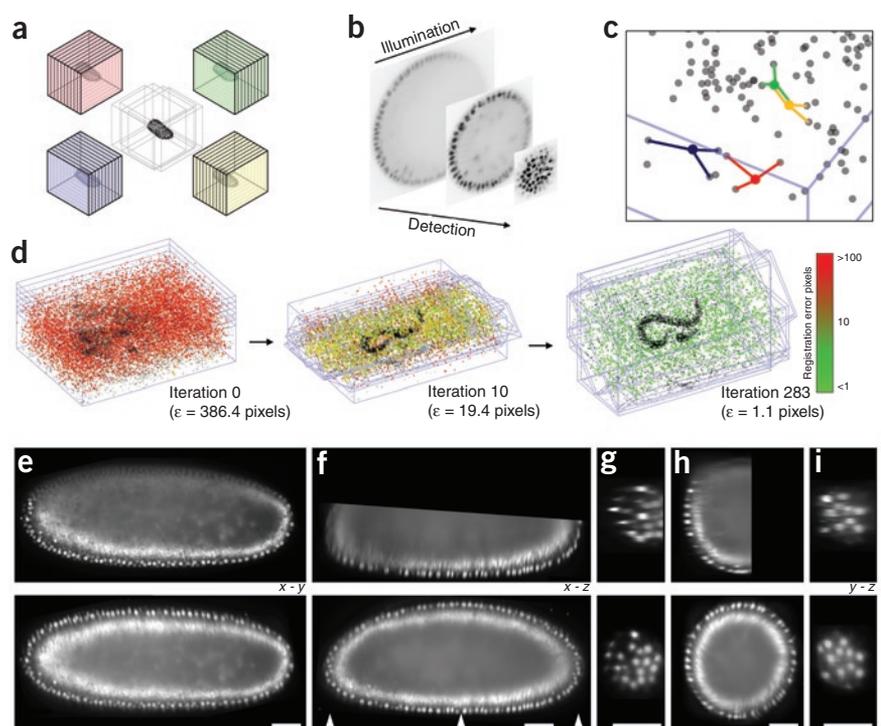
(**Supplementary Table 1**). The average bead displacement and the ratio between correspondence candidates and true correspondences is a quantitative measure of the reconstruction success, which is crucial for automatic validation of registration results in long time-lapse recordings. The beads can be removed optically or computationally from the sample (**Supplementary Methods**).

We applied the bead-based registration framework to SPIM recordings of early *Drosophila melanogaster* embryos, which are very challenging samples for multiview reconstruction owing to the scattering of the yolk that severely limits the overlap between views. We imaged *Drosophila* embryos expressing ubiquitous His-YFP from five and seven views in an extended time-lapse recording covering early embryonic development. We registered each time point separately and then registered all time points to each other compensating for minor drift during image acquisition (**Supplementary Methods**). We combined content-based fusion with nonlinear blending[5] to compensate for brightness differences at boundaries between views (**Supplementary Fig. 4**). The reconstructed multiview acquisition of the specimen showed, in contrast to the single view, comparable lateral and axial resolution (**Fig. 1e,f**). We never imaged the anterior and posterior poles of the embryo with full lateral resolution in this acquisition, and yet the cells were clearly distinguishable, demonstrating the precision of the multiview reconstruction (**Fig. 1g–i**). In the middle of the specimen, the resolution was lower because only some views contributed high-content information whereas other views were blocked by the yolk (**Fig. 1h**). The reconstructed time-lapse recording provided an unprecedented four-dimensional view of *Drosophila* embryogenesis (**Supplementary Videos 4** and **5**).

The bead-based registration framework is sample-independent (**Supplementary Data** and **Supplementary Fig. 5**) and enables fully unguided registration without prior knowledge of the arrangement of the views (**Supplementary Video 4**). The software outperforms intensity-based registration approaches[6,7] in terms of precision and speed, enabling accurate registration of large, multiview acquisitions in minutes (**Supplementary Data**, **Supplementary Fig. 6** and **Supplementary Table 1**). The run time of the bead-based registration framework is comparable to the time it takes to acquire the multiview data, and thus, to our knowledge, it is currently the only solution allowing robust, real-time registration of time-lapse SPIM recordings. Moreover, the bead-based registration framework is applicable to other optical sectioning microscopy techniques (**Supplementary Fig. 7** and **Supplementary Video 6**), considerably expanding the possible applications in biology. We provide our bead-based registration algorithm to the bioimaging community as an open-source plugin for Fiji (**Supplementary Fig. 8** and **Supplementary Software**; http://pacific.mpi-cbg.de/wiki/index.php/SPIM_Registration).

Note: Supplementary information is available on the Nature Methods website.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Stephan Preibisch[1,2], Stephan Saalfeld[1,2], Johannes Schindelin[1] & Pavel Tomancak[1]

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany.
[2]These authors contributed equally to this work.
e-mail: tomancak@mpi-cbg.de

1. Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J. & Stelzer, E.H. *Science* **305**, 1007–1009 (2004).
2. Huisken, J. & Stainier, D.Y. *Development* **136**, 1963–1975 (2009).
3. Lindeberg, T. *J. Appl. Stat.* **21**, 224–270 (1994).
4. Fischler, M.A. & Bolles, R.C. *Commun. ACM* **24**, 381–395 (1981).
5. Preibisch, S., Saalfeld, S. & Tomancak, P. *Bioinformatics* **25**, 1463–1465 (2009).
6. Preibisch, S., Rohlfing, T., Hasak M.P. & Tomancak P. *SPIE Medical Imaging 2008* (eds., Reinhardt, J.M. & Pluim, J.P.W.) **6914**, 69140E-69140E-8 (2008).
7. Swoger, J. *et al. Opt. Express* **15**, 8029–8042 (2007).

Partitioning biological data with transitivity clustering

To the Editor: Clustering is a common computational technique for data analysis in the life sciences. Essentially one defines clustering as a partitioning of arbitrary data objects into groups, such that the objects in each group, or cluster, have common traits, with respect to a similarity function. Ideally, objects from the same cluster are more similar to each other than to objects from different clusters. A density parameter controls the size and the number of resulting clusters.
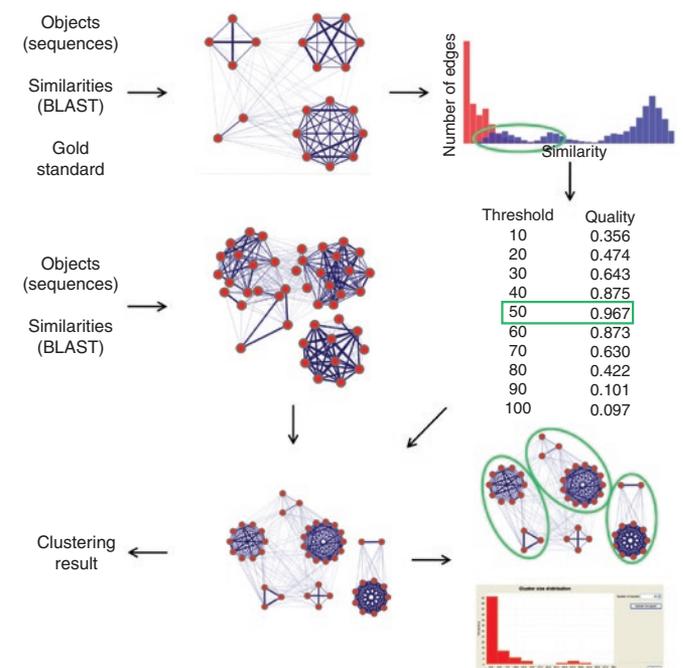
Figure 1 | Data analysis and clustering workflow. After importing and visualizing a subset of the data, a histogram of the distribution of similarities within and between gold standard clusters is created to estimate a promising region for the density parameter. Iterative clustering with varying thresholds and subsequent comparison against the gold standard identifies the 'best threshold'. After importing and visualizing the whole dataset, this threshold is used to detect and report meaningful cluster assignments, which may be subject to subsequent integrated analyses.