**BigStitcher: Reconstructing high-resolution image datasets of cleared and expanded samples**

**David Hörl1,2,#, Fabio Rojas Rusak3,#, Friedrich Preusser2, Paul Tillberg4, Nadine Randel4, Raghav K. Chhetri4, Albert Cardona4,5, Philipp J. Keller4, Hartmann Harz1, Heinrich Leonhardt1, Mathias Treier3,6, Stephan Preibisch2,\***

1Department of Biology II, Ludwig-Maximilians-Universität München, 82152 Planegg-Martinsried, Germany

2Imaging Lab, Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany

3Cardiovascular and Metabolic Sciences, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany

4Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

5Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

6Charité-Universitätsmedizin Berlin, 10117 Berlin, Germany

#Equal contribution

**\***Correspondence should be addressed to S.P. ([stephan.preibisch@mdc-berlin.de](mailto:stephan.preibisch@mdc-berlin.de))

**Abstract** (70 words)

**New methods for clearing and expansion of biological objects create large, transparent samples that can be rapidly imaged using light-sheet microscopy. Resulting terabyte-sized image acquisitions consist of many large, unaligned image tiles that suffer from optical aberrations. We developed the BigStitcher software that efficiently handles and reconstructs large multi-tile, multi-view acquisitions compensating all major geometric transformations introduced by optical effects, thereby making single-cell resolved whole-organ datasets amenable to biological studies.**

**Main Text** (1500 words incl. abstract)

Sample clearing1 and expansion microscopy (ExM)2 are powerful protocols that create large, transparent volumes of whole tissues and organisms (**Fig. 1a,b**). Using light-sheet microscopy (**Fig. 1c**) 3-5, these samples can be imaged with subcellular resolution in their entirety within a few hours6. These acquisitions have the potential to be powerful tools for whole-tissue and whole-organism studies since they preserve endogenous fluorescent proteins and are compatible with most staining methods (**Supplementary Fig. 1**).

However, raw data acquired by the microscope is not directly suitable for visualization and analysis. Many large, overlapping three-dimensional (3d) image tiles are collected that amount to many terabytes in size and require image alignment (**Fig. 1d-n**). Due to sample-induced refraction and scattering of the light-sheet in the direction of illumination (**Fig. new**)7, 3d image tiles are typically acquired twice with alternating illumination from opposing directions to achieve full coverage (**Fig. 1d** and **Supplementary Fig. 2**). Similarly, emitted light is distorted by the sample, effectively limiting maximal imaging depth at which useful data can be collected (**Fig. 1n**). Additionally, sample-induced light refractions cause depth- and wavelength-dependent aberrations in the acquired images (**Fig. 1j,k**). To reconstruct and make these complex datasets easily accessible to biologists and computer scientists we developed the BigStitcher software. It enables interactive visualization using BigDataViewer8, fast and precise alignment, quality estimation, real-time fusion, deconvolution, as well as support for alignment of multi-tile acquisition taken from different physical orientations, so-called multi-tile *views*, thereby effectively doubling the size of specimens that can be imaged (**Fig. 1n**), while further orthogonal views can render the resolution isotropic.

Bigstitcher features a new, user-friendly importer for a multitude of vendor-specific and custom formats and accesses image data through memory-cached, virtual loading10, optionally combined with virtual flatfield correction (**Supplementary Fig. 4** and **Supplementary Note 2**). Performance is optimal when images are stored using a multiresolution, blocked, compressed format enabling interactive visualization, processing and interaction with terabyte-sized image datasets. The importer therefore supports resaving single-block images into the BigDataViewer HDF5 format8.

Accurate reconstruction of these large, complex datasets requires the compensation for different types of image and intensity transformations introduced by the acquisition process. We therefore developed an interactive, stepwise process that compensates for all relevant transformations while providing spatially localized feedback on the quality of the acquired image data.

Firstly, overlapping 3d image tiles are acquired to cover the entire sample at each acquisition angle. Although approximate tile locations are typically known, translation stages usually show significant inaccuracies (**suppl fig ??**). To compute locations for every image tile we developed an image stitching algorithm optimized for very large datasets that can deal with acquisitions arranged in non-regular grids (**Fig. 2a**) containing empty images and multiple independent samples (**Supplementary Fig. 5**). Since acquisitions often consist of hundreds of gigabyte-sized image tiles, each showing very different information content (**Fig. 2**), we initially compute each shift between all pairs of overlapping tiles (*links*) using the parameter-free Phase Correlation (PC) method11,13,14 on downsampled images. It computes all possible shifts between two images and intensity peaks in the resulting PC matrix correspond to shifts with high correlation that we localize with subpixel precision (**Supplementary Note 3** and **Supplementary Fig. 6**). Using simulations, we illustrate that our new pairwise stitching method achieves errors below 1 pixel while reducing computation times 100-fold (**Supplementary Fig. 7-10**). Since correlation-based approaches can fail for image pairs characterized by repetitive patterns, noise or low information content (**Supplementary Fig. 11**), computing final image tile locations requires global optimization. Our new implementation extends the concept of identifying tile positions by minimizing the distance between all links11,12, which, compared to computing a minimum spanning tree15, averages out normally distributed link errors (**Supplementary Fig. 10**) during optimization as tiles are typically linked to many neighbors (**Fig. 2a**). Incorrect links are filtered by quality and by iteratively removing the link disagreeing most with the global optimization result[stitching] using a new compound metric. Currently, unconnected tiles (e.g. empty images) and multiple independent objects in an acquisition are handled by ignoring them11,13, or assuming regular, 2d translational grids14. Here, we present a generic solution to this problem by introducing the concept of *strong* and *weak* links (**Supplementary Fig. 5**) that is independent of the original tile arrangement and is not limited to translations. Strong links are equal to confirmed links, while weak links are derived from current transformations (e.g. approximately known tile positions). Optimizing both link types in an acquisition yields accurate registration results within strongly linked regions and as-good-as-possible alignments for weakly linked groups of tiles (**Fig. 2a** and **Supplementary Fig. 5** and **Supplementary Note 4**). However, correct tile placement (i.e. solving the classical stitching problem) represents the first step and is usually not sufficient to properly align such datasets.

Secondly, microscopy images suffer from spherical and chromatic aberrations [???]. When using high-end optics, these aberrations can be within 1-2 pixels range and single affine transformations per tile are sufficient for alignment (**Suppl. Figs ??, ??**). For these cases we implemented an easy-to-use interest-point-based alignment step that automatically extracts interest points and applies a variation of the iterative closest point algorithm (ICP)16 on affine transformations combined with our new global optimization (**Suppl Fig. ??**) to compensate small rigid or affine distortions that arise from spherical, and also chromatic aberrations if the same autofluorescent structures are visible across channels (**Fig. 1i,j** and **Supplementary Fig. 12 & ?? and Online Methods/Limitations**).

Thirdly, although samples are highly transparent (**Fig. 1b**) light scattering is an issue when imaging centimeters deep into fixed tissue. Although improved designs were recently proposed [???], using dual-sided light-sheet illumination (**Fig. 1c**) to image each 3d image tile twice from either direction (**Fig. 1d**) remains the most prominent method to significantly increase the sample size for which high resolution image data can be collected laterally. Prior to alignment, we automatically suggest the best illumination direction for each tile by estimating image sharpness at the lowest pre-computed resolution level since most tiles only hold useful information from either direction (**Fig. 1d** and **Supplementary Fig. 2 and Online Methods**).

To better understand transformations introduced by sample-induced light refraction we performed simulations of light propagation in tissue (**Fig?? and Online Methods**). Although emitted light of both illumination directions travels through the same tissue, refractions within the illumination light path can lead to different parts of the sample being illuminated, which we confirmed in acquired lightsheet datasets (**Fig ??**). Therefore, as long as the lightsheet remains within focus of the detection objective, non-rigid, sample dependent transformations can be introduced, however, once the lightsheet is out-of-focus, blurred image data is acquired that cannot be reconstructed using BigStitcher. To compensate for these transformations we implemented a virtual, non-rigid alignment method based on identified corresponding interest points (**Online Methods & Fig. ??**). Depending on the magnitude of refraction, faster affine transformations may be sufficient for precise alignment. To allow an intermediate between affine and non-rigid transformations, we added the possibility to virtually split image tiles into smaller blocks thereby enabling piece-wise affine transformations (**Online Methods & Fig. ??**).

Finally, since emitted light is distorted by the sample, maximum imaging depth is limited. To overcome this problem, we acquire samples from opposing directions by rotation (**Fig. 1c**) or by simultaneous acquisition with two objectives4. We therefore implemented a method for registration of large multi-tile *views*, where each *view* consists of a set of aligned image tiles from one physical orientation. We segment interest points in virtually fused, downsampled images of each multi-tile view and identify corresponding interest points using adaptations of geometric hashing17. This significantly improves matching performance and robustly aligns large volumes using affine transformations, effectively doubling the imaging depth of any sample (**Fig. 1c** and **Tab. 1**). Similar to dual-sided illumination, precise *multiview* alignment needs to account for light refraction in the illumination and additionally the excitation light path. Using example data, we illustrate and quantify that our piece-wise affine and non-rigid alignment are able to compensate for such aberrations as long as the image quality allows extraction of interest points in all overlapping multi-tile *views* (**Fig. ??**).

Since image quality is not constant across the (reconstructed) sample due to light aberrations, it is crucial to efficiently estimate it in order to ensure a high-quality acquisition. However, manual inspection at the highest resolution for the entire sample is impossible due to its size. We therefore developed a relative Fourier Ring Correlation (rFRC) that is able to automatically and rapidly estimate image quality throughout terabyte-sized acquisitions even in the presence of common sCMOS camera patterns (**Suppl. Fig 1 & Online Methods**).

For downstream analysis datasets can be fused or directly analyzed using BigDataViewer plugins. We implemented an algorithm for real-time fusion by multithreaded processing of the currently visible plane in virtual images using blockwise multi-resolution loading, which can optionally be performed downsampled and on regions of the sample (**Supplementary Fig. 13**) with support for brightness-equalization. It enables fusion of terabyte-sized images on machines with little memory (**Supplementary Fig. 14**), while increased memory and compute power leads to very fast processing (**Tab. 1**).

Deconvolution is an established method to increase contrast and resolution in light microscopy acquisitions and required point spread functions (PSF) are typically estimated using fluorescent beads embedded with the sample18. To handle multi-tile, multi-view acquisitions, we extended deconvolution code19 allowing BigStitcher to deconvolve selected regions and significantly improve image quality (**Fig. 2b-e**). To be able to apply this approach to cleared samples we developed a protocol for embedding fluorescent beads in polymerization solution enabling robust PSF measurement (**Fig. 2b,e**). We furthermore combine ExM with IsoView light-sheet microscopy4 allowing acquisition of multi-view, multi-tile datasets of expanded tissues enabling reconstruction of entire *Drosophila* larval nervous systems with spatially isotropic subcellular resolution (**Fig. 2c,e**).

BigStitcher is a powerful software package that enables efficient and automatic processing of terabyte-sized datasets. It addresses major unsolved issues such as easy import, managing of large images, datasets acquired in a non-regular grid, globally optimal alignment of sparse datasets, illumination selection, rigid and non-rigid multi-view alignment of multi-tile acquisitions, PSF extraction, and interactive fusion. The aligned dataset and all intermediate steps are interactively displayed. The user has the option to verify and interact with the alignment process at any time to confirm and potentially guide proper alignment of complicated datasets (**Supplementary Fig. 3,11,15,16**). Automatic reconstruction of even large datasets can be achieved within tens of minutes and BigStitcher clearly outperforms existing software in terms of performance, functionality, and user-interaction (**Tab. 1**)11,14,15. BigStitcher supports cleared samples (**Fig. 2a,b)**, ExM samples (**Fig. 2c,e** and **Supplementary Fig. 17**), standard 2D and 3D confocal and widefield acquisitions, as well as tiled, multi-view light-sheet acquisitions (**Fig. 2f**). BigStitcher is implemented in ImgLib210, open-source and provided as a Fiji20 plugin with comprehensive documentation (<http://imagej.net/BigStitcher>). It is compatible with the ImageJ Macro language for most of its functionality and can thus easily be automated. These properties make BigStitcher a powerful and scalable tool for the handling and reconstruction of tiled, high resolution image datasets acquired by new light microscopy technologies.

**Online Methods**

*Animals*

For clearing we used a previously generated BsxH2BeGFP mouse line21, where the exon 1 of the bsx gene is replaced starting at the ATG with the coding sequence for histone2B eGFP. Brains from 10-week old female BsxH2BeGFP/+ mice were used for tissue clearing and imaging. *C. elegans* dauer larvae expressing tagRFP fused to a nuclear localizing sequence under the pan-neuronal rab-3 promotor in all neuron nuclei22 were obtained by selecting dauer larvae in 1% SDS for 30 minutes23. Dauer larvae were fixed with 4% PFA for 30 minutes on ice, placed in 70% Ethanol overnight at 4°C and subsequently stained with DAPI. Experiments were conducted according to institutional guidelines of the Max Delbrück Center for Molecular Medicine in the Helmholtz Association after approval from the Berlin State Office for Health and Social Affairs (LAGeSo, Landesamt für Gesundheit und Soziales, Berlin, Germany). *Drosophila* larva used for ExM were obtained from the strain w;;attp2, carrying an empty attp2 landing site24.

*Clearing and Expansion*

Tissue clearing was performed using the CLARITY protocol1. Mice were deeply anesthetized by intraperitoneal injection of 100 mg/kg Ketamine and 15 mg/kg Xylazine. Mice were exsanguinated by transcardial perfusion with 25 ml cold PBS followed by whole body perfusion with 25 ml cold monomer solution (4 % v/v acrylamide, 4 % w/v Paraformaldehyde (PFA), 0.25 % w/v VA-044 in PBS). The brains were collected and fixed in monomer solution for 2 more days. Next, the whole brains were placed in fresh monomer solution and oxygen was removed from the tubes by vacuum and flushing the tube with nitrogen gas for 15 minutes. The samples were then polymerized by placing the tubes in a 37 °C water bath under gentle shaking for 2 hours. Polymerized brains were placed in clearing solution (4% SDS in 200 mM Boric acid). Brains were incubated in clearing solution for 1 week at 37 °C with daily solution change. Then, the brains were actively cleared using the X-Clarity setup from Logos Bioscience for 24 hours with a current of 1 A at 37 °C. Cleared brains were washed twice overnight with 0.1 % v/v Triton X-100 in PBS and once with PBS. Before imaging, brains were placed overnight in FocusClear for refractive index matching.

For ExM, the nervous system of a 1st instar *Drosophila* larva of was extracted, fixed in 4% PFA/1xPBS/0.1%Triton for 1 hour and washed 2x10 min in 1xPBS/0.1% Triton. Before each antibody usage, the nervous system and the antibodies were blocked in 5% goat serum/1xPBS/0.1%Triton for one hour. Following the blocking, the nervous system was incubated overnight at 4°C in 1:500 monoclonal Anti-α-Tubulin antibody produced in mouse (Sigma Aldrich T6199 1mg/ml). After 5x10 min washing (1xPBS/0.1%Triton), the secondary antibody 1:250 Anti-Mouse CF™488A antibody produced in goat (Sigma Aldrich AB4600387 2mg/ml) was added overnight at 4°C. The stained brain was washed in 1x PBS and then processed using an Expansion Microscopy (ExM) method in which the gel recipe and procedure was modified to achieve 7,5-fold expansion in each dimension (unpublished, BioArxiv in preparation). Briefly, the specimen was treated with acryloyl-X as in standard ExM and embedded using a gel recipe modified from the original method2. The modified recipe uses a reduced cross-linker concentration to achieve greater expansion. After digestion with proteinase K as in the original method, a re-embedding step toughens up the gel, which would otherwise have poor mechanical properties.

*Imaging*

3D images of cleared mouse brains were imaged using the Zeiss Lightsheet Z.1 microscope. Each sample was attached to the sample holder using a cyanoacrylate-based glue. The mounted sample was placed in the FocusClear pre-filled imaging chamber. Images were acquired using the EC Plan-NEOFLUAR 5×/NA 0.16 objective together with the LSFM 5x/NA 0.1 illumination objectives on a Zeiss Light-sheet Z.1. The data was acquired using dual side illumination and from different angles. Images were collected with two 1920 X 1920 pixels sCMOS cameras and stored in the Zeiss CZI file format. Fixed *C. elegans* dauer larvae were embedded in 1.2% agarose containing fluorescent beads and imaged using the same microscope in a water-filled sample chamber. Imaging was performed using the 20x/ NA 1.0 objective with additional 2x zoom. 3D images from a cleared and expanded central nervous system of a *Drosophila* 1st instar larva were acquired using an IsoView light-sheet microscope4 that has been modified for multi-tile acquisition. To prepare the sample for imaging, excess gel surrounding the expanded sample was removed using a scalpel, leaving four flat and smooth gel surfaces for imaging. Some extra gel was left underneath the sample for mounting, and the sample was affixed to a cylindrical post using a cyanoacrylate-based glue. Mounted sample was placed in the imaging chamber filled with deionized water. Orthogonal views for each tile of the sample were acquired sequentially by switching the illumination and detection orders in IsoView. Images were acquired using SpecialOptics 16x/NA 0.71 objectives and recorded using full frames (2048 x 2048 pixels, pixel pitch of 0.4125 μm in sample space) of Orca Flash 4.0 v2 sCMOS cameras. The sample was held stationary during multi-view acquisition of each tile, and depth-sectioned images were acquired every 0.8125 μm by translating the detection piezos over a range of 750 μm. A tile for each view thus covered a field of 832 μm (X), 832 μm (Y), and 750 μm (Z). Automated tiling across the entire sample was achieved by moving the sample in predefined steps of 600 μm in X, Y, and Z until full coverage was achieved. Bi-directional light-sheet illumination was achieved using opposing SpecialOptics objectives and the illumination NA was chosen to be 0.0315 for a confocal parameter of approximately 416 μm. The light-sheets from opposing arms were shifted approximately by their Rayleigh length (208 μm) toward the illumination objectives so that each light-sheet provided uniform coverage of the respective half in the acquired image.

*PSF Measurement and PSF Extraction*

In light-sheet microscopy measured PSFs often differ significantly from simulated ones due to variable precision of light-sheet alignment in every experiment. Therefore, light-sheet deconvolution usually relies on the extraction of PSFs from the actual experiment18,19. To be able to perform PSF extraction in cleared tissue we developed a new protocol. Estapor Fluorescent Microspheres (F-XC 030) were diluted 1:20000 in monomer solution containing bis-acrylamide (0,05 % v/v bis-acrylamide, 4 % v/v acrylamide, 4 % w/v Paraformaldehyde (PFA), 0.25 % w/v VA-044 in PBS). The monomer solution was polymerized under constant vacuum and shaking at 37 °C for 2 hours.  The formed hydrogel was incubated in FocusClear overnight and imaged using the Zeiss Lightsheet Z.1 microscope with the same experimental settings used to acquire previous samples. For *C. elegans* dauer imaging fixed larvae were embedded in 1.2% agarose together with Estapor Fluorescent Microspheres (F-Z 030), diluted 1:2000. For ExM data acquired on the IsoView microscope depth-sectioned images (0.4125 μm step) of fluorescent beads (200nm diameter) embedded in 0.6% low-melting-temperature agarose were imaged using the same experimental settings as for sample imaging. For all samples, PSFs were extracted by detecting interest points in the acquired bead images. Potential bead aggregates were excluded by manual removal on the maximum intensity projection using the BigStitcher module “Manage Interest Points > Remove Interactively”.

*Data Processing Pipeline*

All data shown in this paper was processed using the BigStitcher Fiji plugin. Zeiss CZI files and TIFF files exported by custom microscopes were imported using the AutoLoader and subsequently converted to the HDF5 format. For Zeiss CZI files, approximate tile positions and rotation angles were imported automatically, for other files they were specified by hand using BigStitcher tools (**Supplementary Fig. 3, 16**). For each tile, the best illumination was selected. Tiles were aligned using the phase correlation method together with two-round global optimization, followed by ICP16 refinement. Interest point detection for each multi-tile view was performed. The fast descriptor-based rotation-invariant algorithm or the descriptor-based translation-invariant algorithm after manual rotation application were used to register the interest points of each angle, followed by another round of ICP16 refinement of all image tiles of the acquisition. Fused and deconvolved images were exported as TIFF files.

*Quality Estimation based on Fourier Ring Correlation*

To estimate image quality across entire terabyte sized acquisitions we use a derivative of the Fourier Ring Correlation (FRC) that is robust as well as insensitive to camera noise. *FRCr1,r2(f)* constitutes the per-spatial-frequency (*f*) correlation between two independent realizations *r1*, *r2* of the same signal [https://www.nature.com/articles/nmeth.2448]. In super-resolution, point clouds are therefore typically split into two sets of independent pixels. Here, we use consecutive slices instead and take advantage of the fact that they are nearly identical due to the axial extent of the PSF. Since the result *FRCr1,r2(f)* is an entire correlation spectrum for each *z* plane, we compute a single quality value *Q(z)* by integration over all frequencies *f*

A smoother, symmetric result can be obtained by averaging the FRC spectra obtained using z planes above and below the measured plane

This results in a precise estimation of image quality, except if patterned noise (e.g. sCMOS camera noise) is the dominant source of signal (**Suppl. Fig. 1**). To overcome this instability, we compute a relative FRC (rFRC), which subtracts a *loess*-smoothed (locally estimated scatterplot smoothing) baseline FRC of z-planes spaced by *m* slices that are beyond the axial extent of the PSF.

This effectively measures which additional frequencies the central planes *z*, *z+1* and *z-1*, *z* have in common compared to the planes *z-m*, *z+m* that are beyond the size of the PSF. The resulting values robustly measure image quality in the sample (**Suppl. Fig. 1g**). Since image content can change drastically within a slice we support computation using a manually defined blocksize (e.g. 512x512). To estimate the quality metric for entire acquisitions, we compute for defined points in each image stack. Over all input stacks, these measurements are held as sparse representations using ImgLib2 *RealPointSampleList* that can be rendered virtually and overlaid onto entire fused volumes (**Suppl. Video ??**).

*Illumination Selection*

When imaging large samples using sequential dual-sided illumination, typically only illumination from one direction provides good image quality (**Supplementary Fig. 2**). We therefore implemented *illumination selection* functionality in BigStitcher. We first *combine* all (selected) images by their illumination attribute and thereby group all images that share other attributes besides illumination direction. In each of the resulting groups we select the best image. As quality criteria we offer either rFRC on full resolution images or fast approximation using mean intensity or mean gradient magnitude at the lowest resolution level. While rFRC provides highest distictive power (**Suppl. Fig ??**), both fast approximations are typically sufficient for robust estimation of the higher quality illumination direction (**Fig. 1d**). The image with the highest score is kept, while all other images are marked as *missing* in the dataset, which will lead to them being ignored in subsequent processing steps. Optional resaving of the dataset after this step potentially decreases storage requirement two-fold. In any case, prior to applying automatic illumination estimation, the user has the option to verify and potentially change the result.

*Light Simulation*

We observed non-rigid deformations occuring in areas where image data from opposing lightsheets is recorded by the same camera (**Fig 2?a?**). While it is obvious from existing simulations of light propagation[biobeam] that imaging of the same lightsheet from opposing objectives leads to non-rigid deformations due to different aberrations in the detection light path, we wanted to understand if and how different lightsheet directions can introduce such aberrations if

Since

While it is clear that detection of emitted light from different

In order to model and thereby understand the geometric transformation introduced by light aberrations in tissues we performed simulations of light propagation using raytracing. Here, we specifically focus on understanding the transformations introduced by

*Pairwise Stitching using Fourier-based Phase Correlation*

We calculate pairwise translational shifts using our ImgLib210 implementation of the Fourier-based *phase correlation* algorithm25. In noiseless images, the method produces a phase correlation matrix (PCM) *Q* containing a single δ-impulse at the location corresponding to the shift between the two images. Real images might contain multiple peaks (**Supplementary Fig. 6**) and we localize the *n* highest peaks in *Q* by detecting peaks with subpixel accuracy using a *n*-dimensional implementation of a quadratic fit26.Aside from allowing subpixel-accurate registration, we can use the precision obtained from the subpixel accuracy of the phase correlation to counteract the effects of downsampling, allowing us to achieve registration of similar quality to full-resolution with significant performance gains (**Supplementary Fig. 8-10**). Due to the periodic nature of the Fourier shift theorem, each peak in the PCM actually correspond to 2n possible shifts in *n* dimensions. We therefore test each of these candidate shifts by calculating the cross-correlation between the images *I1* and *I2*, optionally with interpolation in the case of sub-pixel shifts. We choose the shift vector *t* corresponding to the highest cross correlation as the final result after applying downsampling correction, if necessary.

It is often necessary to not only align two single images but groups of images, e.g. all channels of a tile. We therefore implemented a flexible framework for the registration of grouped images (**Supplementary Note 3**). The two images *I1* and *I2* can have arbitrary affine pre-registrations such as sample rotation, correction of axial scaling or already performed registration steps. If pre-registrations of *I1* and *I2* are identical or are only based on different translations or axis-aligned scalings we run the phase correlation on (downsampled) raw input images, otherwise on virtually fused images (**Supplementary Note 3**).

*Downsampling and Simulations*  
To assess the effect of downsampling on the pairwise stitching we use simulations of spheroid-like objects at different signal-to-noise ratios (SNRs) as ground truth. We create realistic images by mimicking image creation in light-sheet microscopy including optical sectioning, 3-fold anisotropy, light attenuation, convolution, and pixel intensity generation using Poisson processes19. Importantly, pairs of overlapping images that we use for benchmarking the subpixel phase correlation method are created using different Poisson processes and are additionally rendered with half a pixel offset of the full resolution images to avoid nearly identical overlaps at high SNRs due to the simulation process (**Supplementary Fig. 7**). We simulate 500 pairwise overlaps, each at SNRs ranging from 1 to 32, and lateral downsamplings ranging from 1x to 8x where axial downsampling is matched as good as possible to achieve near-isotropic resolution as in the actual software. We illustrate that across SNRs downsampled images yield a constant registration quality, which even exceeds that of registration at full resolution for low SNRs. This is achieved through a combination of the smoothing effect during downsampling (**Supplementary Fig. 7**) and precise subpixel-localization (**Supplementary Fig. 8-10**). Due to the smoothing effect registration quality therefore initially increases at 2-fold and 4-fold downsampling (**Supplementary Fig. 8-10**), while when using larger downsamplings the loss of pixel resolution outweighs the effect of smoothing hence the quality drops. Registrations with a constant quality of an average error of below one pixel can be computed at a fraction of the computing time compared to full resolution, typically 4 - 120 times faster. Existing outliers are filtered during global optimization and overall registration quality can further be improved during the ICP16 refinement step.

*Global Optimization*

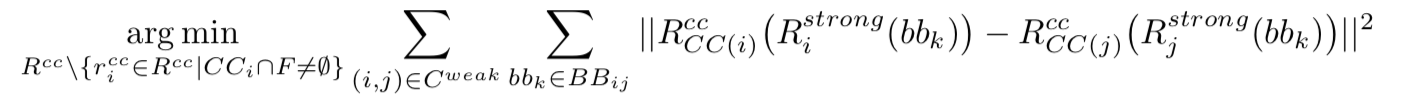
To calculate the final transformations of each image tile we extend the concept of globally optimal registration by iterative minimization of square displacement of point correspondences (**Supplementary Note 4**)11,12. We express pairwise shifts as point correspondences between the bounding box vertices of the overlap region of the images and the same points transformed by the inverse pairwise shift. We globally optimize the registrations *R* in all connected components *CC* of the link graph of images *V* and (strong) links *C*, given point matches (corresponding points) *PM* and fixed images *F* by minimizing:



(1)

In some cases, erroneous pairwise links might not have been filtered out, e.g. due to medium cross-correlation, repetitive patterns, or a low number of correspondences in the ICP refinement. This leads to persistently high registration errors after global optimization, which manifests in a large distance error, i.e. the difference between the individually computed distance between of images (link) and the actual distance between them after global optimization. Iterative removal of the link with the highest distance error from the link graph and repeating the global optimization leads to convergence to user-defined thresholds11. We extend this concept from Preibisch et al.11 to affine transformations, introduce a new heuristic that additionally incorporates link quality and implement it in an extendable framework required for the two-round global optimization (**Supplementary Note 4**).

If the dataset contains empty tiles or multiple disconnected objects with image tiles that do not have links between them, the final transformations will not be propagated between them (**Supplementary Fig. 5**). We therefore developed a novel two-round global optimization that is capable of aligning independent connected components of the link graph using weak links defined by the current transformations of each tile (e.g. approximate locations from metadata or manual alignments). For that purpose, we use the corners of the bounding box of their overlap region, transformed using the results Rstrong of the first round (eq. 1), as point correspondences. The between-component transformations can then be determined by minimizing:



(2)

The final transformations are given by concatenating the in-component and between-component registrations. Using the two-round strategy, registrations are propagated between connected components and distances between neighboring objects are preserved as-well-as-possible (**Supplementary Note 6** and **Supplementary Fig. 5**).

Our global optimization is agnostic to the nature of the point correspondences and transformation model, which allows us to use the same algorithm for translation-based alignment of for example tiled datasets using phase correlation, as well as affine registrations of multi-tile multi-view datasets based on ICP16 refinement or geometric descriptor matching.

*Iterative Closest Point Refinement*  
Although the Phase Correlation-based image stitching produces relatively high-quality alignments, smaller errors can remain due to its inaccuracy (**Supplementary Fig. 8-10**). Furthermore, it is not able to correct for non-translational effects such as chromatic aberration or sample-induced light refraction (**Supp Fig ??**). These effects can be better approximated using affine transformations. We therefore automatically detect interest points and run an Iterative Closest Point algorithm16 for each overlapping pair of images, where the assignment of correspondences is limited by a distance threshold. We use the identified corresponding points of all pairwise links and compute a globally optimal affine transformation for each tile using our new global optimization algorithm. To avoid scaling of datasets, we regularize the affine transformation using a rigid transformation12. The resulting alignment usually improves the alignment quality and the same strategy can be applied to multichannel alignment if the same autofluorescent structures are visible in multipe channels (**Supplementary Fig. 12**).

*Geometric Local Descriptor Matching*

To identify corresponding interest points in between two point clouds, geometric local descriptor matching has been proven to be a powerful technique17,27. The basic idea to express each interest point as a geometric constellation using its *n* (typically three) nearest neighboring interest points. The vector difference between two descriptors then describes how similar the local area of two points is. A geometric local descriptor (GLD) is assumed to be a correspondence candidate if it is at least *m* (typically one to ten) times more similar than the second most similar GLD26. True corresponding interest points between two point clouds are finally identified using the random sample consensus algorithm28 on a regularized affine transformation model. However, fast GLD matching using the rotation-invariant technique based on geometric hashing17 requires relatively randomly distributed points to robustly identify correspondences, while the non-accelerated, redundant, translation-invariant counterpart27 identifies correspondences reliably in non-rotated point clouds of only up to a few thousand points in reasonable time. Here, we developed a new matching procedure by extending both techniques to better suit the requirements when attempting to identify corresponding interest point in between point clouds of prior unknow size derived from imaged structures that are potentially rotated relative to each other.

Redundancy is a powerful mechanism for GLD matching. It uses additional nearest neighbors but excludes some of them sequentially during matching making it more robust to potentially mis-detected interest points27. We therefore extend the fast rotation-invariant technique based on geometric hashing17 with the capability for redundancy. This significantly increases the chance of being able to align randomly oriented point clouds very fast, albeit at low inlier ratios (ratio of true correspondences to total number of correspondence candidates). Rotation invariance is not desired if both point clouds are known to be approximately in same orientation, for example if the rotation of the sample performed by the microscope was known and has been applied to the dataset. Checking for potential rotations simply increases the chance for wrong correspondence candidates. We therefore implemented a fast translation-invariant GLD based on geometric hashing that supports redundancy. All four versions of GLD are available in BigStitcher to enable robust multi-view alignment.

*Virtual Re-blocking*

To allow piece-wise affine transformations or refined illumination selection we implemented a virtual re-blocking of all 3d image stacks of an acquisition (**Suppl. Fig. ??**). The implementation distributes 3d blocks onto each input image stack using a defined overlap and thereby defined a new set of input image stacks for the acquisition. The new, virtual image stacks are computed on the fly from the original image data. Any kind of ImgLoader is supported, including multi-resolution image stacks. Resaving of the dataset as HDF5 or TIFF transforms the virtual image stacks into physical representations.

*Non-Rigid Transformation*

To be able compensate for potentially strong aberrations introduced by light refraction due to the sample we implemented a non-rigid alignment method that we based on the concept of moving least squares[???], which has been show to perform well in biomedical applications[saalfeld]. The underlying principle is to non-rigidly transform images using a set of corresponding points. Therefore, a local transformation is computed for each pixel using a distance-weighted fit of all corresponding points ensuring smoothness. In BigStitcher, corresponding points are a direct result of all interest point-based registration algorithms. To provide a sufficient amount of corresponding interest points, it is yet most useful to derive corresponding interest points using ICP. Regularization is achieved since corresponding interest points are identified on a regularized affine transformation model either using RANSAC or ICP, which both specify a maximum error. This ensures that corresponding points cannot diverge more than this specified error from the regularized affine transformation of each image tile. In combination with virtual re-blocking, this error can be limited to smaller regions than physical tiles.

When computing local transformations for each image, it is necessary to ensure smoothness across *n* overlapping images by defining appropriate point correspondences. The registration identifies pairwise correspondences in between pairs of images, from which we first identify all unique interest points across all images (**Suppl. Fig ??**). For each unique point we determine its location by averaging the locations of all contributing interest points after applying their respective affine transformations. Thereby, the non-rigid transformation only accounts for the remaining error after affine alignment. For each image, corresponding points for moving least squares are subsequently defined between the unique point and the corresponding interest point of the transformed image.

Moving least squares requires to compute a transformation for each pixel, which is computationally expensive and usually cannot be cached due to the image sizes. We therefore implemented a virtual, cached layer that only computes transformation every *m*-th pixel and interpolates affine transformations in between. Since BigDataViewer currently only supports affine transformations, we additionally implemented a multi-resolution preview based on virtually fused non-rigid volumes that can be interactively displayed in an extra BigDataViewer window. We support hybrid fusion of non-rigidly and affine transformed image tiles as non-rigid registration requires significantly increased computational effort.

*Quantification of Registration Results*

To quantify final registration quality, we use a small cleared section of an adult mouse brain. We manually identify corresponding interest points across tiles, illuminations and acquisition angles. Despite very limited aberrations in this sample compared to entire mouse brains we illustrate that using only translation as transformation model is only sufficient for tiled acquisitions, but not across illuminations or acquisition angles. …

Since all other tools for large-scale image registration are

*Virtual Image Fusion*

A set of overlapping, transformed image tiles are fused into one output image using a per-pixel weighted average that minimizes boundary artefacts and can increase contrast by incorporating entropy estimation (**Supplementary Note 6**)17. To correct for unequal brightness and contrast in adjacent images, we optionally perform adjustment of the pixel intensities using a linear transformation per image. An optimal adjustment can be estimated using the same optimization framework used for image registration29 (**Supplementary Note 7**). The memory requirements for the fusion of large volumes can easily exceed the available RAM on a machine due to the size of the output and the combined size of the input images. We therefore developed a novel framework based on ImgLib2 *RandomAcessibleIntervals*10, intensity transformations and coordinate transformations that virtually fuses all pixels of a defined bounding box using all input images and their associated weights. Since the input images are provided through virtual image loading, the size of a virtually fused image is close to zero, irrespective of the size of input and output images. Ideally, input images are available in blocks so that affine transformations that slice input images in arbitrary orientations do not require to load the entire image8. The output image can now be rendered on a pixel-by-pixel basis with minimal memory requirements. Additional caching of the input image and the output images allows an efficient multithreaded fusion for as-fast-as-possible processing given the available memory. Therefore, more RAM will effectively speed up the fusion process (**Tab. 1**), but even machines with very low RAM will be able to fuse terabyte-sized volumes (**Supplementary Fig. 14**). Fused images can be saved by choosing cached or virtual fusion and subsequently saving the ImageJ virtual stack using “Save as image sequence…”.

Downsampling of the output can easily be incorporated by scaling the bounding box and pre-concatenation of the downsampling transformation with each image transformation. If the input files are multi-resolution, we automatically compute the optimal resolution level at which the input needs to be loaded. To optionally further reduce the image size of the fused image, the GUI offers to conserve the original anisotropy between lateral and axial of the acquired sample, which is a sensible choice if the dataset contains a single or opposing (e.g. 0 and 180 degrees) multi-tile views.

*Deconvolution*

In addition to real-time image fusion, we offer deconvolution of bounding-box-defined volumes using a multi-view formulation of the iterative Richardson-Lucy deconvolution algorithm30,31 with Tikhonov regularization32 and various optimizations19. The PSFs required for deconvolution can be extracted from interest points detected in the images (e.g. when subdiffraction fluorescent beads were incorporated with the sample) or supplied as TIFF stacks with odd dimensions by the user. BigStitcher offers GPU acceleration of the deconvolution on CUDA-capable Nvidia GPUs.

To allow deconvolution of multi-tile views, we extended the original deconvolution19 to be based on the *virtual fusion*. Thereby, any number of input image tiles are virtually fused and serve as one of input views for the multi-view deconvolution. Proper multi-view deconvolution of partly overlapping samples requires sophisticated weight normalization in between views19, which we implemented to be computed virtually. Since also the input views are also virtually loaded, the memory requirement of the deconvolution solely depends on the output image size and shows a significantly increased memory-efficiency. All virtual inputs and weights are additionally cached, ensuring highest-possible processing performance for systems with large amounts of RAM.

*Macro Automation and Headless Operation*

In addition to the graphical user interface (GUI), we offer standalone Fiji plugins for most of the individual steps, such as data import, illumination selection, pairwise shift calculation, link filtering, multi-view alignment, global optimization and image fusion/deconvolution. In macro mode results will not be displayed interactively but are instead saved to the XML project file or output files immediately. The individual steps can be recorded as ImageJ *macros* and easily combined into a script for headless batch processing33.

*Limitations of the Framework and other Software Solutions*

BigStitcher is designed for the reconstruction of large multi-tile, dual-illumination, multi-view datasets. To the best of our knowledge it is the only software that supports such type of acquisitions, which includes affine and non-rigid registrations to solve the alignment process for terabyte-sized image data. Several solutions based on image correlation support multi-tile-only data, such as Terastitcher [??], XUVTools [??], ImageJ Stitching [??], which however, only support translation models making them only somewhat applicable to these datasets. The solution developed by the Saalfeld lab can handle even larger datasets (100s of terabytes) and supports affine transformations based on local cross correlation. It is limited to multi-tile acquisitions, requires a cluster to run and has no user interface to access its functionality [??]. Currently, BigStitcher scales well until about 1000 large 3d image tiles per timepoint and sizes of about 10s of terabytes per timepoint. This is, however, mostly due to a limit in rendering capacity of BigDataViewer. Future optimizations of BDV and/or BigStitcher will be able to further increase this limit.

*Example Data*

Example datasets are explained and linked for anonymous download in Supplementary Note 9.

**Acknowledgments**

We thank T. Pietzsch and S. Saalfeld for insightful discussions and BigDataViewer & ImgLib2 support, C. Rueden for Fiji support and maintenance, and the Caenorhabditis Genetics Center at the University of Minnesota for providing C. elegans strains. S.P. and F.P., M.T. were funded by MDC Berlin, F.R.R. and M.T. were funded by the Helmholtz Alliances ICEMED - *Imaging and Curing Environmental Metabolic Diseases and AMPeo’- Aging and Metabolic Programming,* D.H., H.H., H.L. were funded by LMU Munich and the Nanosystems Initiative Munich (NIM), P.T, N.R., R.K.C., A.C., P.J.K. were funded by HHMI Janelia.

**Author contributions**

S.P. conceived the idea in discussions with H.H., H.L, and M.T. D.H. and S.P. developed the algorithms and implemented the software. F.R.R. performed all clearing experiments, reconstructions and benchmarks. F.P. imaged and reconstructed the C. elegans. P.T. and N.R performed ExM sample preparation. R.K.C. and P.J.K. developed the ExM-optimized IsoView microscope and imaged the sample. S.P. reconstructed the ExM sample. S.P., M.T., H.L, H.H., P.J.K, and A.C. supported and supervised the project. S.P., D.H., and F.R.R. wrote the manuscript with input from the coauthors.

**Competing Financial Interests**

The authors declare no competing financial interests.

**References**

1. Chung, K. et al. *Nature*, 497(7449), 332–337 (2013)

2. Chen, F. et al. *Science*, *347*(6221), 543–548 (2015)

3. Huisken, J. et al. *Science*, *305*(5686), 1007–1009 (2004)

4. Chhetri, R. K. et al. *Nature Methods*, *12*(12), 1171–1178 (2015)

5. Truong, T. V. et al. *Nature Methods*, *8*(9), 757–760 (2011)

6. Tomer, R. et al. *Nature Protocols*, *9*(7), 1682–1697 (2014)

7. Richardson, D. S. et al *Cell*, *162*(2), 246–257 (2015)

8. Pietzsch, T. et al. *Nature Methods*, 12(6), 481–483 (2015)

9. Linkert, M. et al. *The Journal of Cell Biology*, 189(5), 777–782 (2010)

10. Pietzsch, T. et al. *Bioinformatics* 28(22), 3009–3011 (2012)

11. Preibisch, S. et al. *Bioinformatics*, 25(11), 1463–1465 (2009)

12. Saalfeld, S. et al. *Nature Methods*, 9(7), 717–720 (2012)

13. Emmenlauer, M. et al. *Journal of Microscopy*, 233(1), 42–60 (2009)

14. Chalfoun, J. et al. *Scientific Reports*, 7(1), 122–10 (2017)

15. Bria, A. et al. *BMC Bioinformatics*, 13(1), 316 (2012)

16. Besl, P. J. et al. *IEEE PAMI*, 14(2), 239–256 (1992)

17. Preibisch, S. et al. *Nature Methods*, 7(6), 418–419 (2010)

18. Verveer, P. J. et al. *Nature Methods*, 4(4), 311–313 (2007)

19. Preibisch, S. et al. *Nature Methods*, 11(6), 645–648 (2014)

20. Schindelin, J. et al. *Nature Methods*, 9(7), 676–682 (2012)

**References (Online Methods)**

21. Sakkou, M. et al. *Cell Metabolism*, 5(6), 450–463 (2007)

22. Nguyen, J. P. et al. *PNAS*, 113(8), E1074–E1081 (2016)

23. Karp, X., *WormBook*, doi/10.1895/wormbook.1.7.1 (2016)

24. Pfeiffer, D. B. et al. *PNAS* 105(28), 9715–20 (2008)

25. Kuglin, C.D. et al. *ICCS*, 163–165 (1975)

26. Lowe, D. G. *International Journal of Computer Vision*, 60(2), 91–110 (2004)

27. Smith, C. S. et al. *The Journal of Cell Biology*, 209(4), 609–619 (2015).

28. Fischler, M. A. et al. *Communications of the ACM*, 24(6), 381–395 (1981)

29. Blasse, C. et al. *Bioinformatics*, 33(16), 2563–2569 (2017)

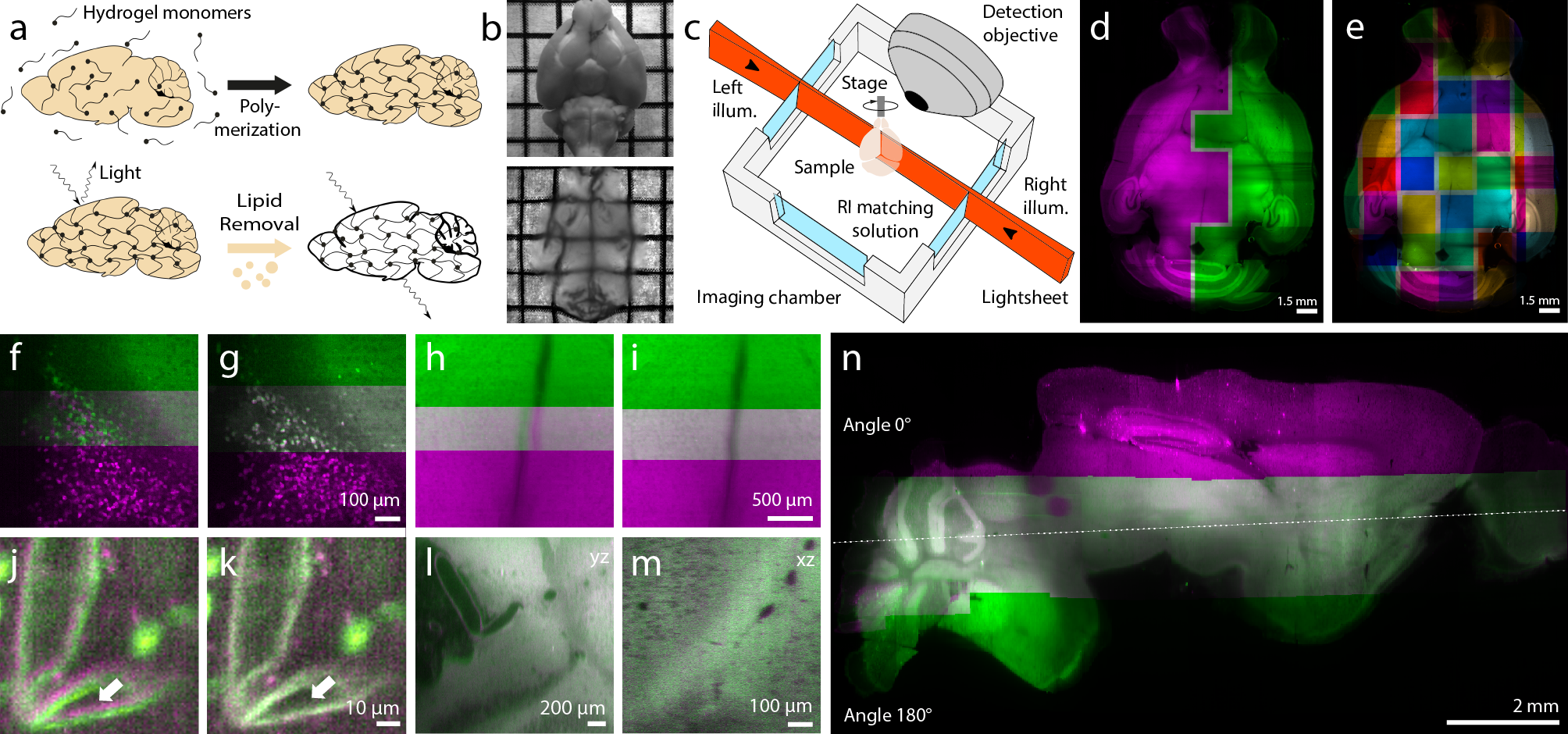
30. Richardson, W.H. J. Opt. Soc. Am. 62, 55–59 (1972).

31. Lucy, L.B. Astron. J. 79, 745–754 (1974).

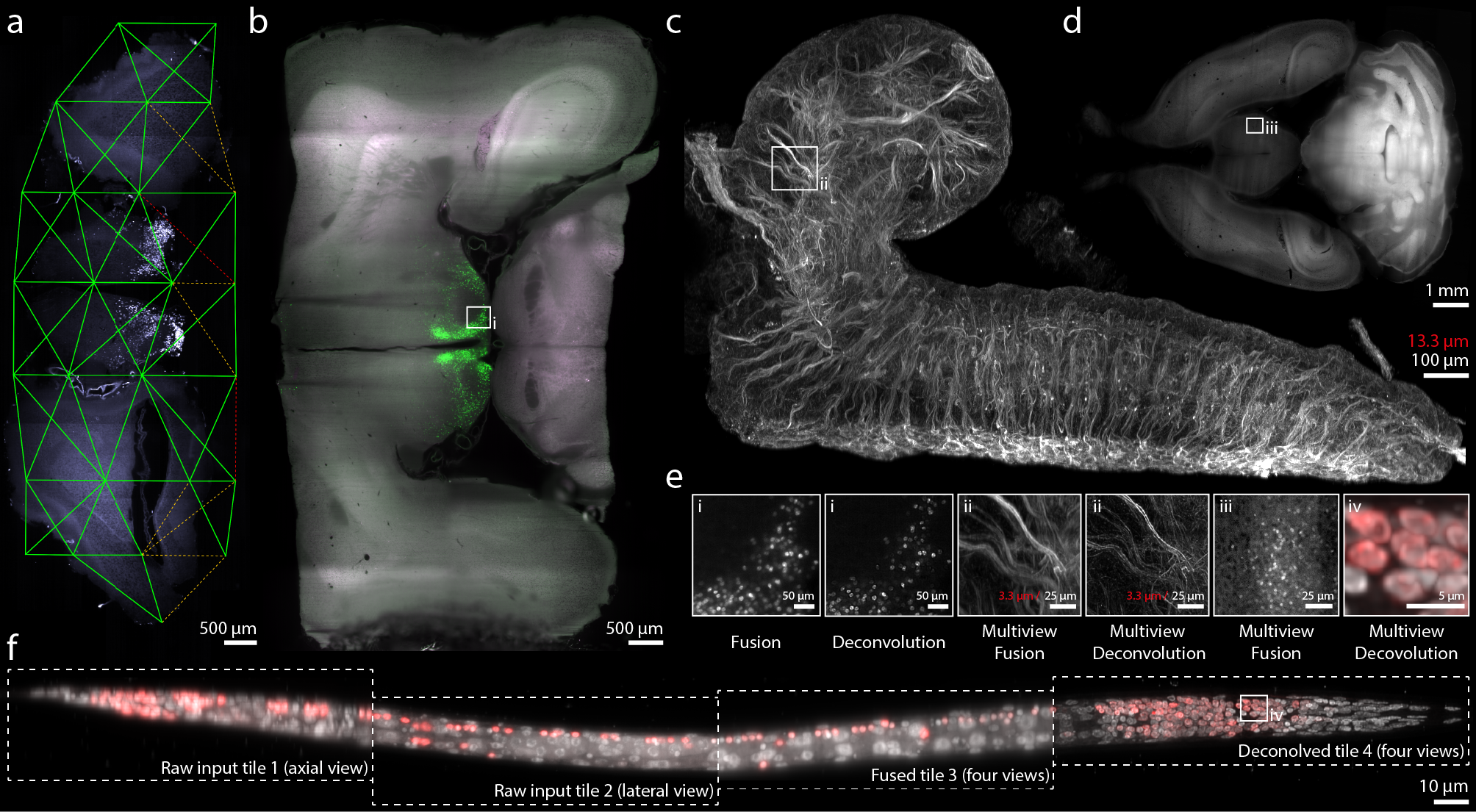
32. Tikhonov, A.N. & Arsenin, V.Y. Solutions of Ill-Posed Problems (Winston, 1977).

33. Schmied, C. et al. *Bioinformatics*, 32(7), 1112–1114 (2016)

**Figures**



**Figure 1**: BigStitcher Principles. **(a)** schematic of the CLARITY sample clearing process. **(b)** adult mouse brain before (up) and after (down) clearing. **(c)** general layout of the type of light-sheet microscope used for most acquisitions17 **(d)** single slice through an entire adult mouse brain acquired with dual-sided illumination. Pink (left illum.) and green (right illum.) image tiles highlight the illumination direction that was automatically selected for each tile. **(e)** overview of an entire section of an acquired adult mouse brain, different colors highlight individual image tiles (each 1920x1920x770 pixels). **(f-i)** illustration of the result of the image stitching from a BsxH2BeGFP brain using phase correlation before (f,h) and after stitching (g,i). **(j,k)** the effect of ICP refinement on two different channels with sufficient autofluorescence visible in both channels. Arrows highlight significant difference before (j) and after (k) refinement. **(l,m)** illustrates the high quality of the multi-view reconstruction for two overlapping multi-tile views at 0 (magenta) and 180 (green) degrees for an axial vs. axial (l) and lateral vs. axial (m) view. **(n)** one slice through an entire adult mouse brain (2.24TB raw data), both views are shown in axial orientation looking along the rotation axis of the microscope. The dotted line illustrates the middle of the section.

****

**Figure 2**: Reconstructed samples. **(a)** One slice through an acquisition of an adult mouse BsxH2BGFP coronal slice encompassing the hypothalamus. Green lines indicate strong links between overlapping image tiles, dotted orange lines refer to links rejected because of low correlation, and red lines illustrate links that were determined to be unconcise. **(b)** One slice through an adult mouse brain expressing an H2B-eGFP lineage tracing marker in BSX-expressing neurons, the box is highlighted in (e). **(c)** Maximum intensity projection of a central part of a 7.5-times expanded central nervous system of a *Drosophila* 1st instar larva with immunostaining for tubulin (Alexa-488) and imaged with multi-tile IsoView light-sheet microscopy, boxes are magnified in (e). (**d**) One slice through a whole multi-tile, multi-view reconstructed adult BsxH2B-GFP/+ mouse brain, the inset is highlighted in (e). **(e)** Zoom-ins to specific areas of (b,c,d,f) illustrating   
(sub-)cellular resolution and the advantage of (multi-view) deconvolution over (multi-view) fusion. **(f)** Fixed *C. elegans* dauer acquired in four tiles with four views each expressing tagRFP in all neuron nuclei, co-stained with DAPI. Boxes highlight the quality of axial and lateral raw input data, multi-view fusion, and multi-view deconvolution.

**Tables**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 1: Performance comparison** | | | | | | |
| Data Size | 130 Mb / 63 Gb / 300 Gb | 130 Mb / 63 Gb / 300 Gb | 130 Mb / 63 Gb / 300 Gb | 130 Mb / 63 Gb / 300 Gb | 130 Mb / 63 Gb / 300 Gb | 1.67 Tb |
| Software | Illumination Selection | Stitching  (depends on Illumination Selection) | ICP refine | Fusion**#** @ 8x down-sampling / full resolution  (depends on Illumination Selection) | Virtual Fusion  @ full res. | Multi-view Registration |
| BigStitcher | n.a. / 9 s /  14 s | 4 s / 1 min 28 s /  7 min 15 s | 9 s / 57 s /  2 min 52 s | 1 s / 9 s / 44 s  8 s / 1 h 19 min /  8 h 43 min | 1 s / 15 s / 19 s | 6 min 30 s |
| TeraStitcher | **X** | 16 s / 34 min 43 s / 2 h 29 min | **X** | 1 min 15 s / 14 min 24 s /  1 h 4  2 min 29 s / 26 min 41 s /  2 h 35 min | **X** | **X** |
| ImageJ Stitcher | **X** | 10 s / 61 min /  5 h 35 min | **X** | **X** / **X** / **X**  16 s / 10 h 36 min /  44 h 40 min | **X** | **X** |
| Xuv Tools | **X** | 2 s / **X** / **X** | **X** | **X** / **X** / **X**  9 s / **X** / **X** | **X** | **X** |
| Arivis | **X** | 27 s / 47 h 15 min / 7 d 4 h\* | **X** | **X** / **X** / **X**  2 s / 5 min / 51 min | **X** | **X** |
| **BigStitcher benchmark for processing a terabyte-sized multi-view dataset** | | | | | | |
| Data Size | 1.67 Tb | | | | | |
| Software | Illumination Selection | Stitching | ICP refine | Fusion @ 4x downsampling | Virtual Fusion @ full res. (display / save) | Multi-view Registration |
| BigStitcher | 59 s | 1h 14 min | 7 min 28 s | 1 h 1 min | 56 s / 23 h 50 min\* | 6 min 30 s |

**Table 1:** Comparison of BigSticher features and performance with other available stitching programs. Benchmarks were performed on a HP Z840 workstation running Windows 10 Pro with two Intel Xenon E5-2667v4 CPUs and 512 GB memory. The benchmarks for the Arivis software were performed on a HIVE system running Windows Server 2012 R2 with two Intel Xenon E5-2640v3 CPUs and 256 GB memory. The latest stable version of each stitching program was used. BigStitcher datasets were stitched using 4x times (x,y) downsampling, and fusion was performed at stated downsampling levels and saved as TIFF stacks. Correctness of the stitching could only be confirmed in the BigStitcher in full detail due to the flexibility of interactive inspection. Processing of multi-view, dual-illumination datasets as well as ICP refinement and virtual fusion is only possible in BigStitcher. Note that fusion in BigStitcher supports intensity adjustment. All displayed values are averaged from three independent runs of each respective software except those marked with a star. **#**Note that image fusion results are not comparable since all applications except BigStitcher fuse translation-only datasets, which is a significantly simpler problem. A cross indicates that the functionality is not supported by the software, “n.a.” indicates that the dataset did not require this feature.