BIFROST: a method for registering diverse imaging datasets

Luke E Brezovec[*,4], Andrew B Berger[*,4], Yukun A Hao[*,4], Albert Lin[*,1,2], Osama M Ahmed[*,1,3], Diego A Pacheco[1], Stephan Y Thiberge[1], Mala Murthy[1,✉], and Thomas R Clandinin[4,✉]

[*]Co-First-Author
[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA
[2]Center for the Physics of Biological Function, Princeton University, Princeton, NJ, USA
[3]Department of Psychology, University of Washington
[4]Department of Neurobiology, Stanford University

Abstract

Quantitative comparison of brain-wide neural dynamics across different experimental conditions often requires precise alignment to a common set of anatomical coordinates. While such approaches are routinely applied in functional magnetic resonance imaging (fMRI), registering *in vivo* fluorescence imaging data to *ex vivo*-derived reference atlases is challenging, given the many differences in imaging modality, microscope specification, and sample preparation. Moreover, in many systems, animal to animal variation in brain structure limits registration precision. Using the highly stereotyped architecture of the fruit fly brain as a model, we overcome these challenges by building a reference atlas based directly on *in vivo* multiphoton-imaged brains, called the Functional Drosophila Atlas (FDA). We then develop a novel two-step pipeline, BrIdge For Registering Over Statistical Templates (BIFROST), for transforming neural imaging data into this common space, and for importing *ex vivo* resources, such as connectomes. Using genetically labeled cell types to provide ground truth, we demonstrate that this method allows voxel registration with micron precision. Thus, this method provides a generalizable pipeline for registering neural activity datasets to one another, allowing quantitative comparisons across experiments, microscopes, genotypes, and anatomical atlases, including connectomes.

Drosophila melanogaster **| Whole-brain imaging | image processing**
Correspondence: `mmurthy@princeton.edu`
Correspondence: `trc@stanford.edu`

Main

Understanding brain function requires simultaneous measurements of neural activity across large populations of cells. Such measurements include functional magnetic resonance imaging (fMRI) studies in humans, non-human primates and rodents [1–3], brain-wide calcium imaging in the nematode *Caenorhabditis elegans* [4–8], the larval zebrafish *Danio rerio* [9–13], and the adult fruit fly *Drosophila melanogaster* [14–21], as well as large scale calcium imaging in the mouse [22]. In the context of fMRI, aligning functional imaging data across subjects is routine, and has both enabled quantitative comparisons across experiments and improved reproducibility and statistical power. At the same time, quantitative comparisons of cellular-level calcium imaging datasets have remained less common, constrained by the limitations of registration methods that can accurately import voxel signals into a common brain space. Here we describe a robust, generalizable method for registering *in vivo* calcium imaging datasets in the fruit fly, allowing quantitative comparisons across animals, experiments and labs with micron precision. These data can then be aligned with existing anatomical atlases, and connectomes, allowing functional signals to be mapped to identified cells and circuits.

Atlases have been widely used to compile data within a common space. Multiple atlases for flies, zebrafish, mice and humans have been constructed. These atlases are each built from a single imaging modality, such as fMRI or fluorescence imaging, and can contain different types of functional, anatomical and molecular data [23–26]. However, cross-modal registration that bridges functional and anatomical measures have remained a central challenge, and despite substantial efforts, some problems continue to defy existing methods [27–29]. In particular, bridging *ex vivo* fixed tissue atlases to *in vivo* atlases has largely defied current methods. Here we leverage ground truth datasets in *Drosophila*, an organism with highly stereotyped neuroanatomy, to optimize a new pipeline that achieves precise cross-modal registration.

A wealth of neuroanatomical datasets have been described in the adult fruit fly [30]. These include resources that allow genetic access to specific cell types, as well as two independent connectomes (one partial and one complete) that describe the synaptic connectivity of the brain [31, 32]. Critical to the widespread use of these data has been the registration of these resources into a common set of spatial coordinates allowing, for example, morphologically defined cell types identified in the connectome to be associated with specific genetic tools [33, 34]. However these reference datasets cannot be used for precise, cellular-scale alignment of *in vivo* functional data [14, 17], because *ex vivo* and *in vivo* datasets are derived from very different imaging modalities and preparation conditions. As a result, these datasets have markedly different image statistics, so far preventing cross-modal registration.

Recent work has developed several approaches to measuring neural activity across large portions of the adult fly brain, using different microscopes, distinct genotypes, and different imaging conditions [14, 17, 19–21, 35]. These data have revealed wide-spread sensory responses and movement-related neural activity, probed the relationships between neural activity and metabolism, and have led to the discovery of novel feeding circuits. To compare data within a single experimental paradigm, each of these studies aligned volumetric neural activity data from individual animals onto bespoke templates, an approach that precludes direct comparisons across labs and

imaging modalities.

Our goal was to develop an image registration pipeline, BrIdge For Registering Over Statistical Templates (BIFROST), that would make quantitative comparisons across imaging modalities possible. To do this, we first created a common template constructed entirely from *in vivo* brain volumes, named the Functional Drosophila Atlas (FDA), that can accommodate *in vivo* functional datasets from different experiments and labs. Second, we aligned the FDA with extant *ex vivo* templates, thereby importing atlas labels, powerful tools for neuron identification, neuropil annotation, and connectivity. Using these atlas labels, we demonstrate that our registration pipeline outperforms existing methods [36, 37]. Third, we registered different *in vivo* datasets containing genetically defined cell types to the FDA, thereby validating our method. These ground truth measurements demonstrate that our alignments are precise to the micron scale. Fourth, we demonstrate this pipeline can be used to align partial sub-volumes of the brain into FDA, allowing users the flexibility to focus on regions of interest, but retain the ability to make quantitative comparisons across the brain. Finally, we demonstrate that our pipeline can be used to precisely register functional imaging data to multiple connectomes. Thus, this pipeline creates a common space for *in vivo* neural imaging data, provides easy-to-use tools for accurate co-registration, and enables direct comparisons to *ex vivo* anatomical resources including connectomes [31, 32].

Results

Overview

Functional imaging datasets collected using fluorescence microscopy often comprise two separate channels, with one channel recording neuronal activity using one sensor (such as a calcium indicator), and one channel recording signals associated with a structural marker that broadly labels the brain. In our approach, the structural signals from individual brains in a single experiment are first registered together to form a template. The warp parameters derived from this transformation are then applied to the neuronal activity channel from each brain, thereby bringing these signals into the template space. Next, templates derived from each experiment or laboratory are aligned to the Functional Drosophila Atlas (FDA), allowing all datasets to be quantitatively compared to each other, and to other resources that are contained in FDA.

Developing the Functional Drosophila Atlas

Our goal was to develop an accurate pipeline for registering brain-wide imaging data to a single atlas. In flies, previous work has described atlases that span the entire brain using ex vivo datasets, and as well as atlases that span the central brain in vivo [14, 17, 30, 38]. However, no in vivo atlas spanning the entire brain has been described in either sex. To develop an atlas that best captures the structure of the female fly brain *in vivo*, a widely used model, we sought to suppress both individual and technical variation. To do this, we first imaged each individual brain 100 times at a resolution of 0.6 x 0.6 x 1 µm, capturing expression of a pan-neuronally expressed cell surface marker (myristylated tdTomato), using two photon microscopy of the intact brain inside the head. These 100 volumes were then aligned using linear (affine) and non-linear (Symmetric Normalization (SyN)) transformations, as implemented in Advanced Normalization Tools (ANTs) [36, 37]. These were then averaged to define a single volumetric image of each brain that suppressed technical variation in each collected volume. This process was repeated for 30 individuals, and based on a qualitative assessment, 16 were selected for further image processing. Each of these images were normalized, sharpened, and iteratively aligned using linear and non-linear transformations to construct the FDA (Extended Data Fig. 1, Extended Data Fig. 2, Methods).

We next tried to align *ex vivo* resources, including JRC2018F anatomical labels and genetic tools, the hemibrain connectome and the FlyWire Connectome to the FDA using ANTs [31, 32, 38]. However, many regions of the brain aligned poorly. We therefore adapted SynthMorph, a learned contrast-invariant registration method, and used it in sequence with linear and non-linear SyN transformations to improve registration of the *ex vivo* resources to the FDA [36, 39].

Registering individual datasets to FDA

We collected neural activity (nSyb>GCaMP6s; the dependent channel) and anatomical data (nSyb>myr::tdTomato; the alignment channel) at brain wide scale in different labs using different imaging systems (Fig. 1, Methods). To register these datasets to the FDA, we first generated a dataset template by iteratively aligning the anatomical scan from each animal using linear and non-linear transformations. We next used the combination of linear, non-linear SyN, and SynthMorph to register these anatomical scans to the FDA. The transformations that best align each anatomical scan were then applied to the corresponding neural activity data, thereby registering the functional signals to the FDA.

Quantifying registration performance

Making quantitative measurements of registration accuracy is challenging [40]. To address this challenge, we took two independent approaches. First, we quantified the performance of our method by measuring the overlap of small well-defined anatomical regions that were manually labeled independently in both the *ex vivo* and *in vivo* atlases. Second, we expressed a fluorescent marker in small cell-type specific sub-populations, and quantified their alignment within and across labs, and to connectomes.

BIFROST outperforms existing methods for registration across modalities

We first quantified registration performance by measuring the alignment of neuropils labeled in the FDA space to the corresponding neuropils labeled in an established *ex vivo* anatomical atlas, JRC2018F (Fig. 2)[38]. Alignment accuracy was quantified for each pair of neuropils using the Sørenson-Dice coefficient, which captures the normalized fraction of voxels that overlap across both neuropil masks [41, 42]. As a control, we first used a linear transform to align JRC2018F to the

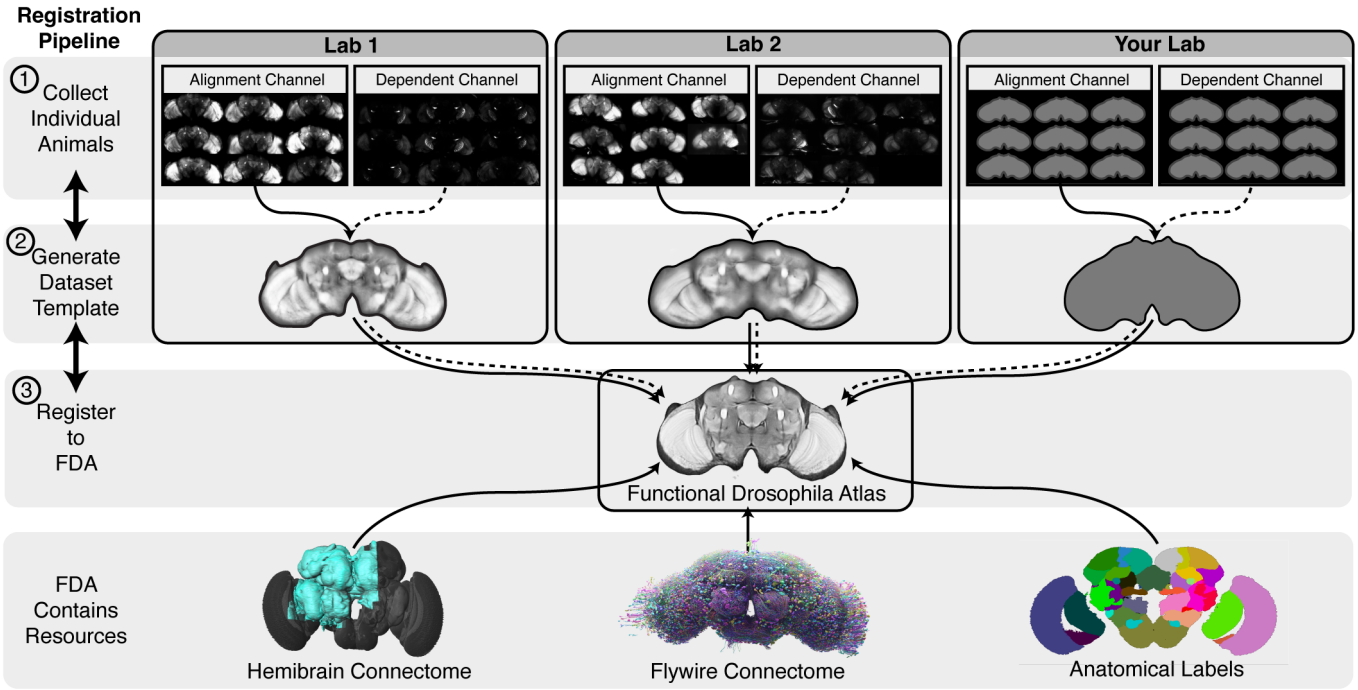

Figure 1. Overview of the BIFROST pipeline. (Step 1) Collect data from multiple animals, labeling both an anatomical alignment channel and a dependent neural activity channel. (Step 2) A dataset template is constructed, unifying individuals animals in the dataset into a common template. The dataset template is constructed from the anatomical channels and the resulting transforms are applied to the neural data to register them into the template space. (Step 3) Dataset templates are aligned to the Functional Drosophila Atlas (FDA), such that all datasets can now be directly compared. Other resources have been registered to this space, including anatomical labels and connectomes.

FDA, and achieved an average Sørenson-Dice score of 0.52 (range: 0.13 to 0.75). Next, we added a non-linear transformation step (SyN), the core non-linear transformation embedded in the widely used registration pipeline ANTs. However, SyN achieved only a modest increase in performance, with an average Sørenson-Dice score of 0.54 (range: 0.18 to 0.77), emphasizing the challenge of cross-modal registration. However, by adding SynthMorph to complete the BIFROST pipeline and perform the same alignment, we achieved an average Sørenson-Dice score of 0.65 (range: 0.45 to 0.84). Thus, BIFROST provides an effective tool for registering signals across the brain.

Quantifying registration accuracy using sparse cell populations

While the Sørenson-Dice coefficient of labeled anatomical ROIs is widely used to estimate the precision of registration, this approach also has significant limitations [40]. The stereotyped architecture of the fly brain, combined with cell-type specific genetic labelling, make possible a quantitative assessment of registration precision, giving access to ground truth measurements that are generally not possible in other experimental systems. We first expressed a fluorescent indicator in a single genetically-identifiable cell type, Lobula Columnar 11 neurons (LC11). We chose the LC11 population because LC11 axons converge onto a single glomerulus, facilitating precise estimation of glomerulus position in 3D (Fig. 3 and Extended Data Fig. 3). This glomerulus lies in the posterior ventral lateral protocerebrum (PVLP) and posterior lateral protocerebrum (PLP), two large neuropil that displayed relatively low contrast in the structural imaging channel. Thus, aligning LC11 within and across laboratories provides a challenging test-case for the BIFROST pipeline. As above, we compared the performance of the BIFROST pipeline to alternative, truncated pipelines that omitted various alignment steps, and included images collected independently in two laboratories (Fig. 3b). Each image was from the same strain, and expressed the neural activity marker GCaMP6s only in LC11 (as the dependent channel), as well as myristylated-td-Tomato in all neurons (as the structural channel). Qualitatively, individual LC11 glomeruli from both laboratories were similar in appearance after registration (Fig. 3c). By projecting the data along each imaging axis, we observed a nearly uniform error distribution, even including along the Z axis (corresponding to the anterior to posterior axis of the brain), the axis that generally suffers most from image distortion (Fig. 3 and Extended Data Fig. 3c).

We quantified alignment precision using the brain-to-brain variation in the position of the glomerulus centroid, independently for both hemispheres. The average pairwise displacement of any two centroids was 5.2 µm in Laboratory 1, 6.1 µm in Laboratory 2, and 7.3 µm across laboratories (Fig. 3e). Notably, the BIFROST pipeline outperformed all truncated variations of the pipeline, particularly for cross-lab comparisons (Fig. 3e and Extended Data Fig. 3d).

To test whether this registration precision could be extended to a different brain region, we repeated this experiment using an additional cell population labelled by doublesex (DSX), a neuronal population that extends throughout many neuropils and comprises only fine <10 µm diameter) processes (Extended Data Fig. 4a,b). Again, BIFROST achieved comparable results, displaying 5.5 µm average pairwise displacements between cen-

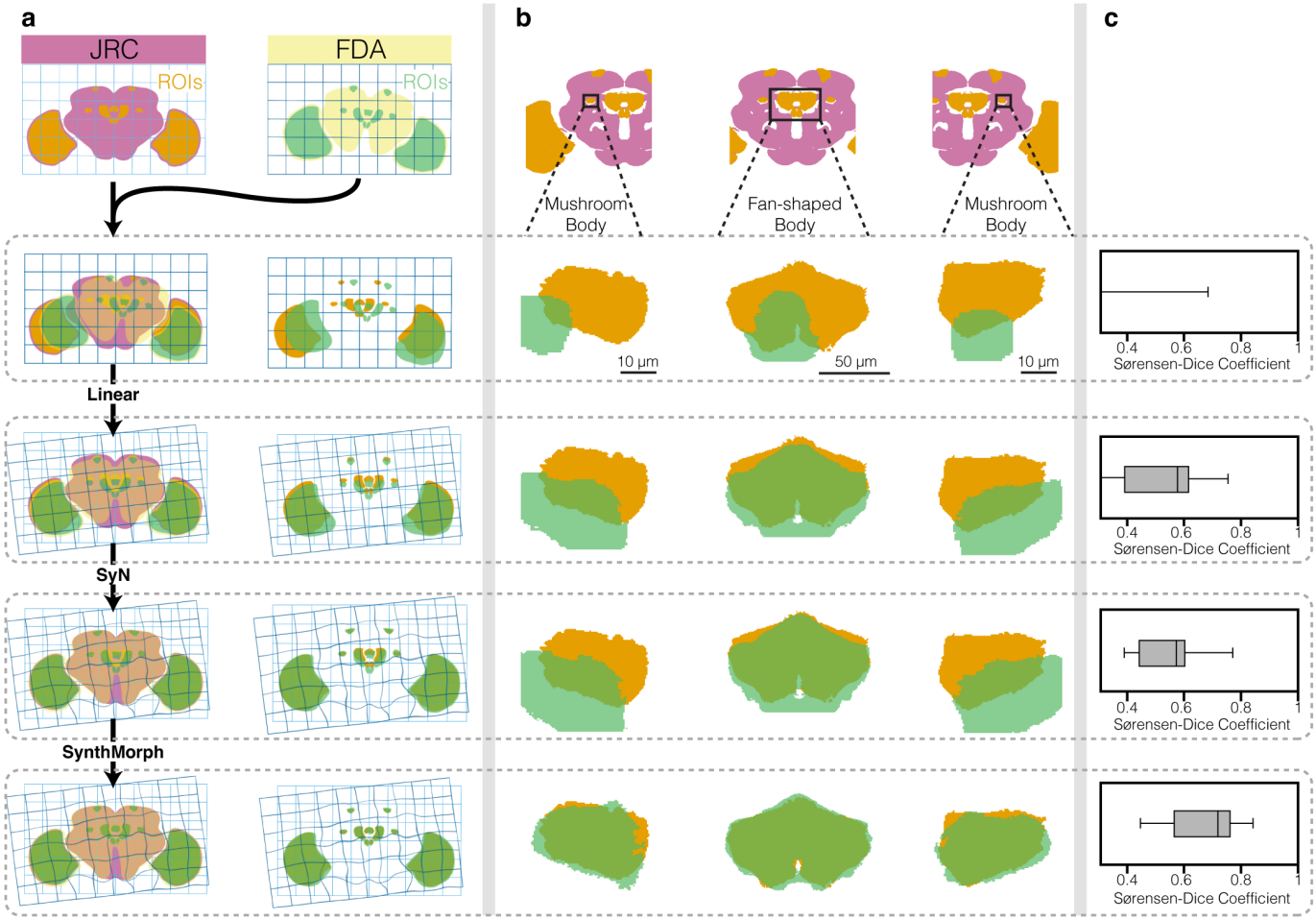

Figure 2. BIFROST registration outperforms classical registration across modalities. a, Schematization of the registration steps of the BIFROST pipeline. (left) The FDA was transformed into the space of JRC2018F through successive applications of one linear and two non-linear SyN and SynthMorph steps. (right) The transformations computed at each step were applied to ROI annotations. Only a subset of the 13 labeled ROIs are illustrated. **b**, Selected ROIs at successive steps of the pipeline. **c**, Quantification of ROI overlap using the Sørenson-Dice coefficient. Box center line indicates median, box limits indicate quartiles, whiskers indicate minimum and maximum.

troids in a readily identifiable structure within the DSX population (Extended Data Fig. 4c,d).

Registration of brain sub-volumes

Many experiments capture neural activity signals from only a sub-region of the brain and would benefit from registration across animals. We therefore adapted the BIFROST pipeline to align sub-volumes into the FDA (see Methods). To test the accuracy of sub-volume alignment, we generated a simulated sub-volume dataset by selecting a 95 x 95 x 38 µm sub-region of one hemisphere from each LC11 brain (Fig. 4). Importantly, this sub-volume was not selected from the LC11 template; rather, it was selected independently for each fly, blind to variation in brain orientation and position. We then constructed a sub-volume template from the individual sub-volumes (Fig. 4b,c). We aligned this sub-volume mean to the FDA using BIFROST (Fig. 4d,e) then assessed the accuracy of alignment as before (Fig. 4f-h). We found good agreement between the two sets of aligned data, with an average pairwise displacement between centroids of approximately 6 µm, demonstrating that the BIFROST pipeline can align partial brain volumes to the FDA.

Registration of the FDA with connectomes

Our goal was to align neuronal skeletons and synapse positions derived from connectomes to the FDA. Prior work had described the transformations that register both the hemibrain connectome and the FlyWire connectome to JRC2018F, an ex vivo atlas that integrates these and other datasets [38]. We therefore sought to align the in vivo FDA with the ex vivo JRC2018F. However, the image statistics associated with these imaging modalities have substantial differences including (1) changes in brain morphology due to physical constraints of the head, (2) distortion created by fixation, and changes in the angle of the imaging axis (3) differences in the spatial distribution of fluorescence signals due to in vivo labeling of cell membranes versus ex-vivo immunohistochemistical labeling of synaptic antigens and (4) differences in SNR characteristics associated with single and two-photon microscopy.

To overcome these challenges, we first registered JRC2018F to the FDA using a combination of linear transformations and SynthMorph, to create a template. Next, we iteratively applied the non-linear transformations in ANTs to align synapse and skeleton positions from both the HemiBrain connectome and FlyWire with this template [32, 43, 44]. We next examined the

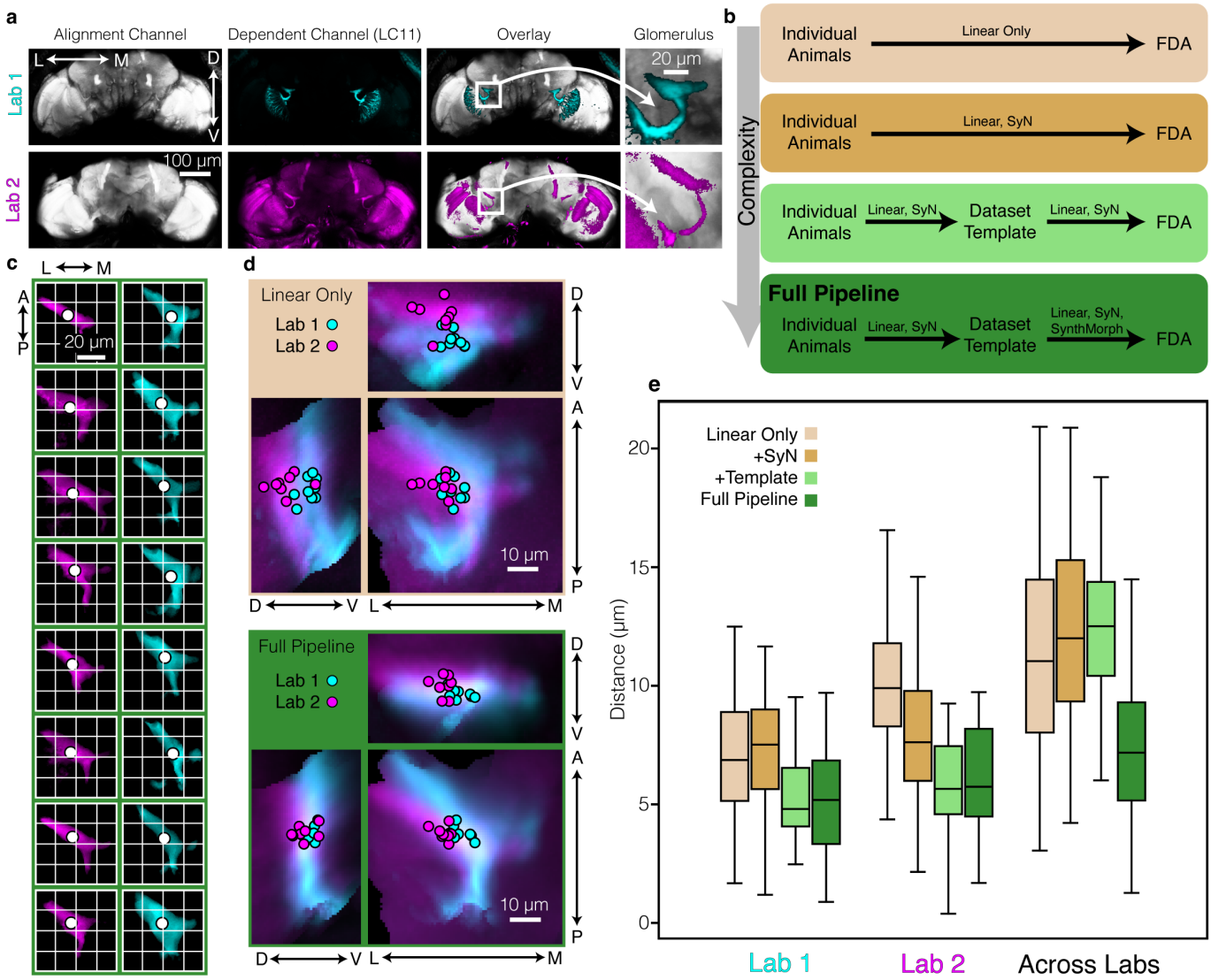

Figure 3. BIFROST accurately registers brains within and across laboratories. a, A single brain from each laboratory showing the alignment channel (pan-neuronally expressed myr-tdtomato) and the dependent channel (LC11-expressed GcAMP6s). The LC11 glomerulus region is highlighted. **b**, Schematic flow chart of the BIFROST pipeline, as well as truncated versions that omit individual steps. **c**, High magnification views of LC11 glomeruli in individual animals from both laboratories after registration into the FDA using BIFROST. Dot denotes centroid of each glomerulus. **d**, As in **c**, but individual glomeruli are overlaid and projections along each axis are shown. For comparison, glomeruli transformed by the linear-only pipeline and the full BIFROST pipeline are overlaid(Lab1 n=9; lab2 n=8). **e**, Quantification of the distribution of pairwise centroid distances within and across laboratories, for each pipeline variant. Box center line indicates median, box limits indicate quartiles, whiskers indicate 1.5x the inter-quartile range.

accuracy of this registration by comparing the positions of the LC11 glomeruli measured in our in vivo datasets to that identified in both the hemibrain and FlyWire (Fig. 5). Remarkably, this cross-modal alignment was as precise as the alignment across in vivo datasets, with a precision of approximately 5 µm (Fig. 5c-f) for LC11, and 7 µm for DSX (Extended Data Fig. 4d). Thus direct comparisons between anatomical wiring diagrams and functional volumetric images are now feasible with high precision.

Discussion

We developed BrIdge For Registering Over Statistical Templates (BIFROST), a pipeline for registering neural activity data across animals, laboratories, and imaging modalities. To facilitate registration, we created the Functional Drosophila Atlas (FDA), an in vivo atlas that provides a common space in which to register neural datasets. Using genetically labeled specific neuron populations as the basis for ground-truth, we demonstrate that BIFROST registers neural data across modalities at the micron scale, comparable to previous fixed-tissue-only alignments [45]. We used BIFROST to register connectomes, anatomical labels, and genetic resources into FDA, allowing researchers to relate functional neural data at either brain-wide scale, or specific sub-volumes, to these resources. Thus, this toolkit will allow quantitative comparisons of neural activity across diverse datasets and experimental designs in this important model system. Moreover, as this pipeline addresses a long-standing challenge in precisely registering signals across imaging modalities, it is readily adaptable to other experimental systems, including vertebrates.

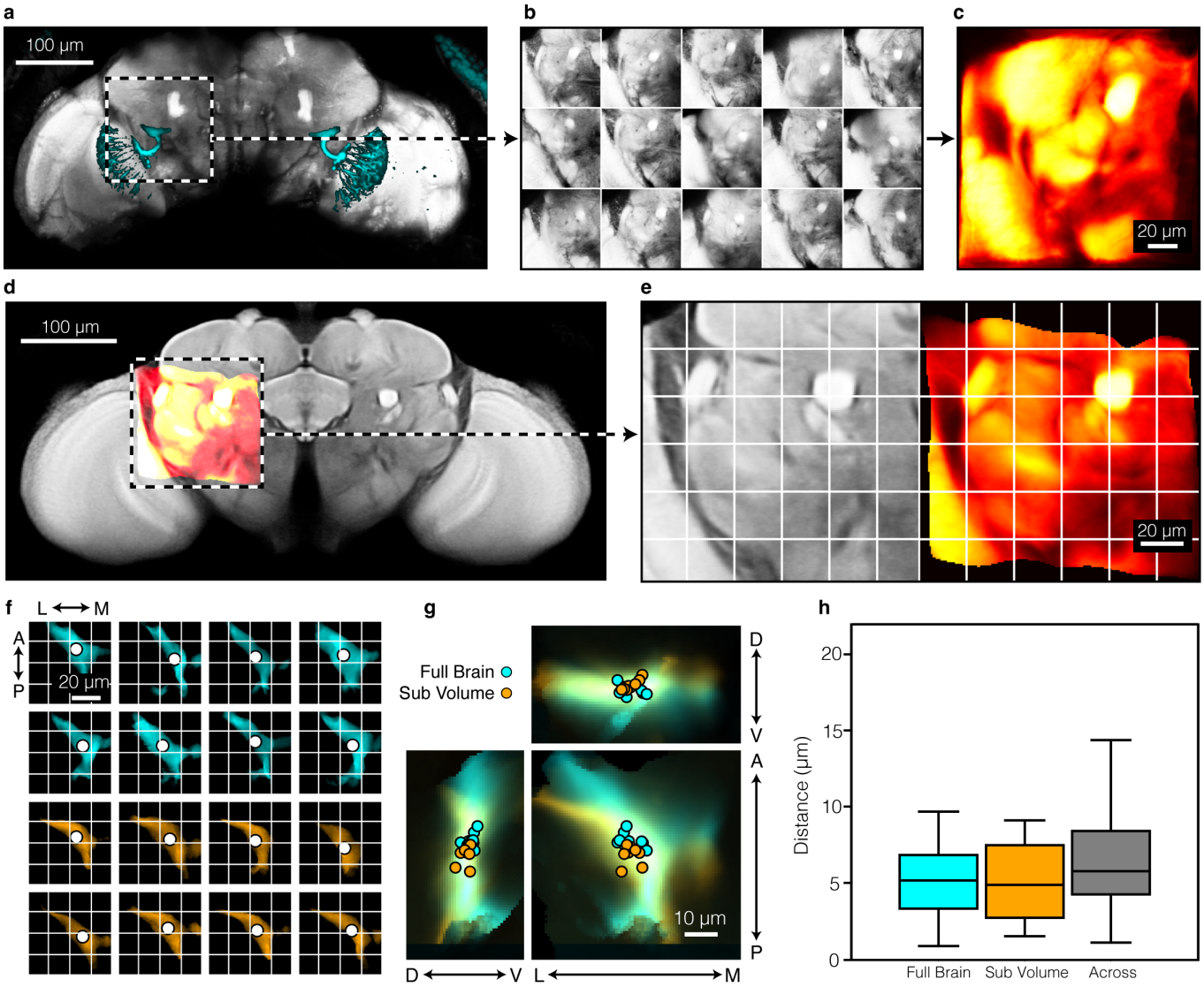

Figure 4. Sub-volumes of the brain can be accurately registered. a, Example fly from laboratory 1 showing the alignment channel (pan-neuronally expressed myr-tdtomato, gray) and the dependent channel (LC11 expressed GcAMP6s, cyan). Dashed white box indicates LC11 glomerulus region. **b**, Each brain sub-volume from each individual animal. **c**, The dataset template generated from sub-volumes. **d**, The template sub-volume (red) was aligned to the FDA (grey) using BIFROST. **e**, High magnification view of the aligned mean sub-volume. FDA (left, grey) and aligned sub-volume (right, red) are shown. **f**, High magnification view of each LC11 glomerulus after registration with BIFROST to FDA. Aligned LC11 glomeruli from either the whole brain (cyan) or subvolume imaging (orange). Dot denotes centroid of each glomerulus. **g**, As in **f**, but overlaid across animals and projected along each axis. **h**, Quantification of pair-wise centroid distances after alignment using either the full brain image or the sub-volume. Box center line indicates median, box limits indicate quartiles, whiskers indicate 1.5x the inter-quartile range.

Quantitatively comparing signals across populations of cells and animals under multiple conditions is critical for studying the neural basis of complex behaviors [4, 9]. For example, large-scale imaging approaches have been used to describe functional correlations, metabolic processes, sensory processing, locomotion and feeding [14, 16, 17, 19, 20, 35, 46]. BIFROST will allow such datasets to be integrated into a single space defined by the FDA, and to be quantitatively compared. Moreover, by achieving precise, cross-modal registration with connectomes, BIFROST allows population-level neural activity signals to be associated with particular candidate neurons, as well as associated genetic tools. In addition, given measurements of single genetically targeted cell types that have been identified in the connectome, investigators could directly obtain information about the population-level activity patterns that include these cells, a capacity that will only grow as additional population-level datasets are incorporated into the FDA using BIFROST. Thus, by enabling smooth transitions between large scale measurements and specific identified circuits and cell types, the approach we describe will enhance the development and testing of both small and large-scale circuit and computational models, including those that link neural activity to sensory input, internal state, and behavior.

Large scale functional imaging approaches, as well anatomical studies using connectomics, are of rapidly expanding interest in many systems, including worms, flies, fish, mice, and primates. Moreover, there is broad interest in using connectomic constraints to inform computational models of neural activity [47–49]. However, these efforts have been substantially hindered by the absence of precise registration tools that

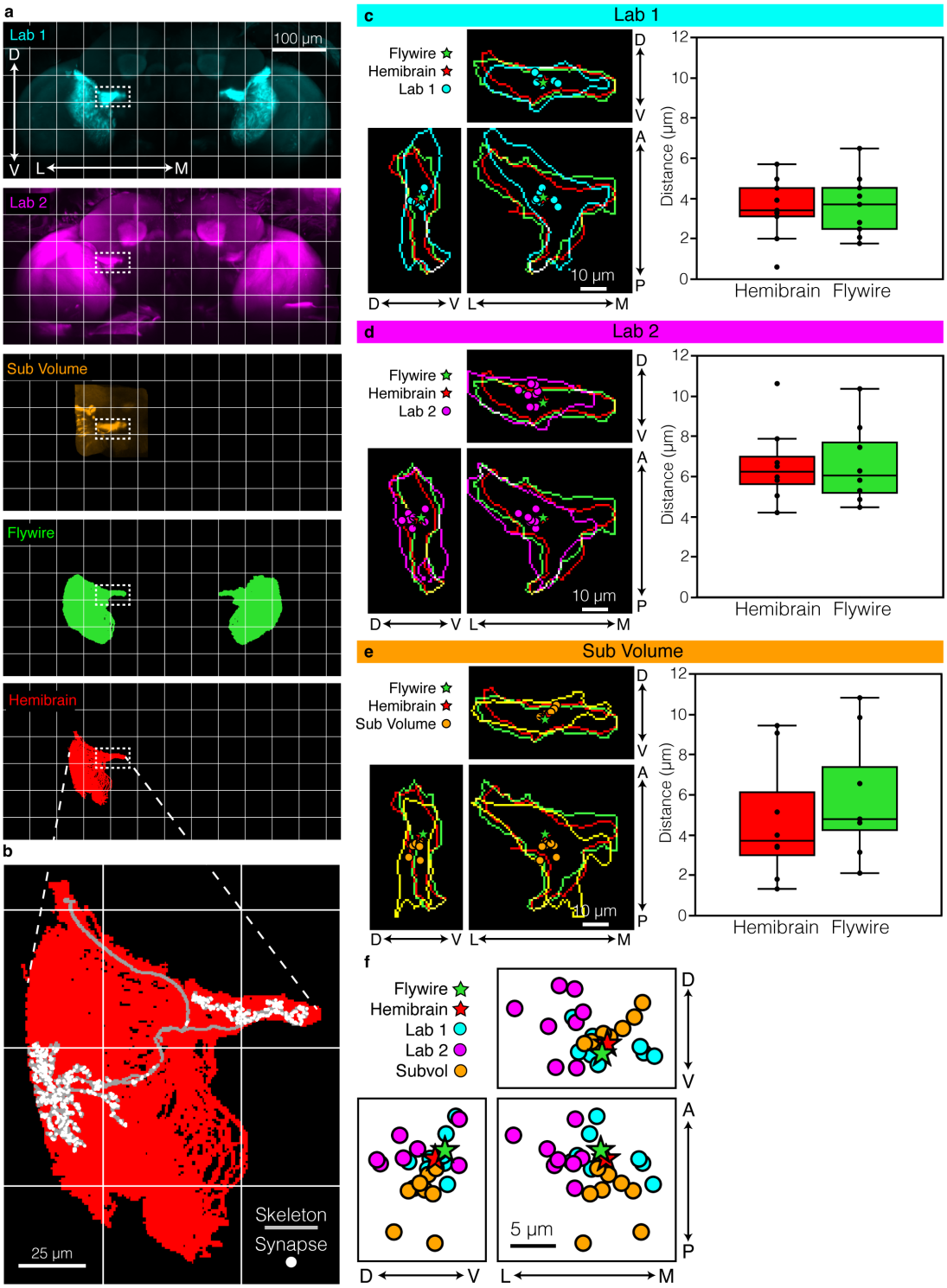

Figure 5. BIFROST enables registration of connectomes to *in vivo* imaging data. a, Maximum projections of brains after alignment to the FDA. Template image of the LC11 channel for laboratory 1 (cyan), laboratory 2 (magenta), brain sub-volume (orange), LC11 skeletons from the flywire connectome (green), and LC11 skeletons from the hemibrain connectome (red), are shown. **b**, Example of a single LC11 skeleton and synapses after being aligned to the FDA. **c**, Comparing alignment accuracy of Laboratory 1 with the hemibrain and flywire connectomes. Left: high magnification view of the LC11 glomerulus; projections along each axis are shown. Cyan contour demarks the edges of the template image of LC11 from laboratory 1, while red and green denote the edge of LC11 skeletons from the hemibrain and flywire connectomes registered into the FDA. Right: quantification of the distribution of pair-wise centroid distances between each individual LC11 glomerulus and the hemibrain and flywire connectomes. Box center line indicates median, box limits indicate quartiles, whiskers indicate 1.5x the inter-quartile range. **d**, As in **c**, but for Laboratory 2. **e**, As in **c**, but using the sub-volume. **f**, Overlay of LC11 centroids from all brains.

can bridge across functional and anatomical imaging modalities. We demonstrate that BIFROST can achieve alignment across these disparate modalities with micron precision in the fruit fly, a capacity that should rapidly accelerate connectomics-informed functional modeling in this system. Finally, while we have optimized BIFROST using the fly brain as a model, this pipeline should directly enable cross-modal image registration in other organs and animals.

Methods

Genotypes

Flies were grown at 25°C on molasses (Clandinin Lab) or cornmeal (Murthy Lab) media, and imaged at 3-5 days post eclosion. The flies used to generate the FDA were *w+/w+;UAS-myr::tdTomato/UAS-GCaMP6f; nSyb-Gal4/+*. The flies used to label LC11 neurons were *w+/w-;nSyb-LexA,LexAop-myr::tdTomato/R22H02-p65ADZp;UAS-GCaMP6s/R20G06-ZpGAL4DBD*. The flies used to label DSX neurons were *w+/w+;brp>STOP>v5-LexA,LexAop-myr::tdTomato/UAS-myr::tdTomato;DSX-FLP,LexAop-GCaMP6s/nSyb-Gal4*.

Mounting and Dissection - Clandinin Lab

Flies were immobilized using a chilled Peltier plate, then fitted into a mount comprising a 3D-printed plastic dish holding a steel shim to secure the head and thorax. UV curable glue was applied to the dorsal part of the head, and on the dorsal thorax. A saline solution was added to the dish for dissection (103 mM NaCl, 3 mM KCl, 5 mMTES, 1 mM NaH2PO4, 4 mM MgCl2, 1.5 mM CaCl2, 10 mM trehalose, 10 mM glucose, 7 mM sucrose, and 26 mM NaHCO3). The posterior head cuticle was cut using a tungsten needle and removed to expose the whole brain. Dissection forceps were used to remove fat and trachea.

Mounting and Dissection - Murthy Lab

Flies were chilled on ice and placed in a Peltier-cooled "sarcophagus" held at 4°C, with the head of the animal restrained in a 3D printed holder. We positioned the head at a 90° angle relative to the thorax and restrained it via UV-cured glue and wax. The holder was then filled with saline, and the cuticle on the posterior side of the head was removed using fine forceps (Dumont 5SF) and a sharp needle. Fat and trachea were removed before imaging.

Two-Photon Imaging - Clandinin Lab

Imaging data was collected using a resonant scanning Bruker Ultima IV system with a piezo drive and a Leica 20x HCX APO 1.0 NA water immersion objective. Either a Chameleon Vision II femtosecond laser (Coherent), or a MaiTai BB (SpectraPhysics) was used to excite GCaMP and tdTomato at 920nm. A 525/50nm filter and a 595/50nm filter were applied to the GCaMP and tdTomato emission photons, respectively. Photons in both channels were collected simultaneously using two GaAsP photomultiplier tubes (Hamamatsu). 100 imaging volumes were collected at 0.6 x 0.6 x 1 µm (1024 x 512 x 241 XYZ voxels).

Two-Photon Imaging - Murthy Lab

Imaging data was collected on a custom-built 2-photon resonant scanning microscope equipped with a Chameleon Ultra II Ti:sapphire laser (Coherent) and a 25x water immersion objective (Olympus XLPLN25XWMP2). Dissected flies were placed below the objective and perfused with saline. The laser was used to excite GCaMP and tdTomato at 920 nm, with a 520/70nm filter (Semrock) applied to the green channel and a 617/73nm filter (Semrock) applied to the red channel. Note, the slightly wider band-pass of this green filter (compared to Clandinin Lab) likely contributed to additional bleed-through of photons from tdTomato, as can be seen in Figure 3. Photons in both channels were simultaneously collected using GASP photomultiplier tubes (Hamamatsu). We recorded 100 whole-brain volumes at a resolution of 0.49 x 0.49 x 1 µm (1024 x 512 x 300 XYZ voxels), to a sample depth of 300 µm. The microscope was controlled by ScanImage.

Creation of FDA

Each anatomical scan was created by first imaging the myr-tdTomato signal 100 times at 0.6 x 0.6 x 1 µm (1024 x 512 x 241 XYZ voxels). These 100 volumes were averaged, then each volume was warped (linear and non-linear) to this mean using ANTs, thereby correcting for motion. These aligned volumes were then averaged, creating the anatomical scan for each brain. Separate scans were then selected for additional processing with an intensity based masking, removal of non-contiguous blobs, and histogram equalization to brighten overly dark areas and darken overly-bright areas. Each brain was mirrored across the Y axis, doubling our effective data to 32 brains. These 32 brains were all linearly aligned to a single seed brain chosen from the 32, and averaged ("linear0"). The 32 brains were linearly aligned to "linear0", and again averaged, producing "linear1". Next, the individual anatomical scans were sharpened (using unsharp masking), and aligned again (linear and non-linear) to "linear", and averaged to produced "SyN0". The last step was repeated two more times to produce the final FDA.

The BIFROST pipeline

The BIFROST pipeline comprises four steps. First, a dataset template is created from structural images from each animal. Second, this dataset template is registered to the FDA. Third, each timepoint from each dependent channel is registered to the dataset template. Finally, these registered data are transformed into FDA space using the transformation calculated on step two.

We provide the BIFROST pipeline as a Snakemake workflow that describes the dependency structure of the whole pipeline [50]. This facilitates parallel execution of independent steps, and as a result BIFROST can be transparently scaled be transparently scaled from local execution on a single machine to thousands of parallel jobs on a cluster. This parallelization is critical, because serial execution of the ANTs dependent steps over such large datasets would take weeks to months to complete. BIFROST can be executed on all common cluster scheduling systems including Slurm, PBS and SGE and on cloud services via Kubernetes and several common cloud APIs [50]. This implementation allows a dataset over any num-

ber of animals with any number of channels, each imaged for an arbitrary number of timepoints and stored following a particular directory structure to be quickly submitted to a cluster for parallel execution with only minimal customization.

The BIFROST pipeline: creation of dataset templates. Dataset templates were constructed from structural images of each animal following a standard procedure [38]. These images were mirrored, doubling the effective sample size, and pre-processed with the scikit-image implementation of contrast limited adaptive histogram equalization (CLAHE) to enhance contrast [51, 52]. More specifically, we used CLAHE to normalize contrast across different imaging planes. Template construction begins with a single image, chosen arbitrarily from the pre-processed images to serve as the initial template. Linear transformations then aligned each pre-processed image to this initial template. Next, the transformed images were averaged to obtain a new template. Following this linear iteration, template construction continued with several (typically four) iterations of non-linear alignment and averaging. In each iteration, individual images were non-linearly transformed to the current template using SyN. Next, the transformed images were averaged to obtain a mean image and the transformations themselves were also averaged. To complete each iteration, the next template is obtained by transforming the mean image through the inverse of the mean transformation [53, 54]. The fourth iteration of this cycle produces the final dataset template.

The BIFROST pipeline: registration with SynthMorph. We wrote a configurable tool for image registration which registers a "moving" image to a "fixed" image with successive linear, non-linear SyN and SynthMorph transforms. If necessary, both "moving" and "fixed" images can be downsampled to reduce memory burden and computation time. As we found that SynthMorph was essential to effective registration in the central brain, but not the optic lobes, we also generated a mask for the SynthMorph transform. Next, we pre-processed the "moving" and "fixed" images by re-scaling intensities to the interval $[0, 1]$ and applied the scikit-image implementation of CLAHE [51, 52]. By default CLAHE was applied to the "moving" image but can be applied to the "fixed" image as well.

After this pre-processing, linear and non-linear (SyN) transforms registering the "moving" image to the "fixed" image were computed and applied in sequence. To save memory, the images were downsampled, typically to 160 x 160 x 192 voxels. Images were transposed prior to downsampling to align the longest axis of the image with that of the inference volume. Next, SynthMorph was run on the downsampled images yielding a warp field. The warp field was mirror symmetrized along the axis that minimizes the root mean square distance between the "moving" image and its mirror. Next, the mask was transformed through the linear and non-linear transformations. The warp field was then up-sampled to the original size of the images and applied to the "moving" image at all locations outside the mask, yielding the final image. All transformations and meta-data needed to apply the full transform were saved in a HDF5 file [55]. All tooling for computing registrations and applying the resulting transforms are provided as part of our Python package.

Manual segmentation of neuropils

We used ITK-SNAP to manually draw regions of interest in an early version of the FDA, and then registered into the final FDA [56, 57]. We used ITK-SNAP's built-in contrast-based segmentation to delineate the boundaries of the whole brain. We then hand-segmented the mushroom bodies (calyx, peduncles, ventral lobes, and medial lobes), central complex (protocerebral bridge), and optic lobes in each z-slice of the volume.

Calculation of Sørenson-Dice coefficients for the cross-modal quantification

Region of interest annotations in the space of JRC2018F were obtained by registering a previous template that was published with regional annotations into the space of JRC2018F [58]. Correspondences between these ROIs and those annotated in the FDA were identified manually. Given sets X and Y, the Sørenson-Dice coefficient is defined as

$$\frac{2|X \cap Y|}{|X| + |Y|}$$

where $|X|$ and $|Y|$ are the cardinalities of the sets. The Sørenson-Dice coefficient was computed for each ROI in a voxelwise manner.

Defining the positions of LC11 and DSX centroids

After alignment, whole-brain volumes were cropped to a region that contained the feature of interest (the terminal glomerulus for LC11, and a specific stalk for DSX). Box size was 95 x 57 x 76 µm for LC11, and 38 x 30 x 23 µm for DSX. In addition, for LC11, fluorescence outside of the PLP and PVLP regions were masked using the anatomical ROIs to avoid expression from LC11 dendrites in the lobula. An intensity threshold was then manually selected for each animal that best removed background fluorescence while maintaining the shape of the glomerulus. The image was then binarized and the center of mass determined

Aligning Connectomes to the FDA

The coordinates of the skeleton and synapses of LC11 were fetched from the online resources for the Hemibrain and FlyWire and were transformed into the space of JRC2018F using the flybrains Python package [59–61]. In the flybrains package, the coordinate systems for the Hemibrain and FlyWire are labeled as "JRCFIB2018Fraw" and "FLYWIRE" respectively. These data were further transformed from JRC2018F into the space of the FDA by application of a bridging transformation. The bridging transformation was obtained by first applying the BIFROST pipeline to transform JRC2018F to the FDA. Due to technical limitations, the BIFROST pipeline does not support coordinate transformations. Therefore, we recapitulated this transform with ANTs by iteratively aligning the original JRC2018F to the JRC2018F that was transformed into FDA space by the BIFROST pipeline. The final transform was obtained after three iterations, which captured the original BIFROST transformation.

AUTHOR CONTRIBUTIONS

T.R.C., M.M., and L.E.B. conceived the project. L.E.B., Y.A.H., A.L., O.M.A., and D.A.P. collected data. L.E.B., A.B.B., Y.A.H., and O.M.A. analyzed the data. A.B.B. built the pipeline. S.Y.T. built the microscope used to collect data in the Murthy Lab. L.E.B., A.B.B., Y.A.H., A.L., O.M.A., and T.R.C wrote the manuscript. M.M. and T.R.C. advised throughout the project.

ACKNOWLEDGEMENTS We would like to thank Megan Wang, Max Aragon, and members of the Clandinin lab for helpful discussions. We would also like to thank Greg Jefferis for sharing his advice, as well as Niyathi Annamaneni for help with figure design. OMA was supported by the Burroughs Welcome Fund Postdoctoral Enrichment Program and the Princeton Presidential Postdoctoral Fellowship. LEB was supported by a National Science Foundation Graduate Research Fellowship. This work was supported by NIH NINDS R35 to MM, R01EY022638 to TRC, NIH NRAIN R01 NS110060 to MM and TRC, Simons Collaboration of the Global Brain award to MM and TRC. TRC is a Chan-Zuckerberg BioHub Investigator.

References

[1] Jonathan D Power, Bradley L Schlaggar, and Steven E Petersen. "Studying brain organization via spontaneous fMRI signal". In: *Neuron* 84.4 (2014), pp. 681–696.

[2] Nikos K Logothetis et al. "Functional imaging of the monkey brain". In: *Nature neuroscience* 2.6 (1999), pp. 555–562.

[3] Elisabeth Jonckers et al. "Functional connectivity fMRI of the rodent brain: comparison of functional connectivity networks in rat and mouse". In: *PloS one* 6.4 (2011), e18876.

[4] Saul Kato et al. "Global Brain Dynamics Embed the Motor Command Sequence of Caenorhabditis elegans". In: *Cell* 163.3 (Oct. 2015), pp. 656–669. ISSN: 10974172. DOI: 10.1016/j.cell.2015.09.034.

[5] Harris S. Kaplan et al. "Nested Neuronal Dynamics Orchestrate a Behavioral Hierarchy across Timescales". In: *Neuron* 105.3 (2020), pp. 562–576. ISSN: 10974199. DOI: 10.1016/j.neuron.2019.10.037. URL: https://doi.org/10.1016/j.neuron.2019.10.037.

[6] Eviatar Yemini et al. "NeuroPAL: A Multicolor Atlas for Whole-Brain Neuronal Identification in C. elegans". In: *Cell* 184 (2021), pp. 272–288. DOI: 10.1016/j.cell.2020.12.012.

[7] Vladislav Susoy et al. "Natural sensory context drives diverse brain-wide activity during C. elegans mating". In: *Cell* 184.20 (2021).

[8] Kelsey M. Hallinen et al. "Decoding locomotion from population neural activity in moving C. elegans". In: *bioRxiv* (2021). DOI: 10.1101/445643. eprint: https://www.biorxiv.org/content/early/2021/01/15/445643.full.pdf. URL: https://www.biorxiv.org/content/early/2021/01/15/445643.

[9] Misha B. Ahrens et al. "Whole-brain functional imaging at cellular resolution using light-sheet microscopy". In: *Nature Methods* 10.5 (2013), pp. 413–420. ISSN: 15487091. DOI: 10.1038/nmeth.2434.

[10] Timothy W Dunn et al. "Brain-wide mapping of neural activity controlling zebrafish exploratory locomotion". In: *eLife* 5.MARCH2016 (2016), pp. 1–29. ISSN: 2050084X. DOI: 10.7554/eLife.12741.

[11] Martin Haesemeyer et al. "A Brain-wide Circuit Model of Heat-Evoked Swimming Behavior in Larval Zebrafish". In: *Neuron* 98.4 (2018), pp. 817–831. ISSN: 10974199. DOI: 10.1016/j.neuron.2018.04.013. URL: https://doi.org/10.1016/j.neuron.2018.04.013.

[12] Gilles C. Vanwalleghem, Misha B. Ahrens, and Ethan K. Scott. "Integrative whole-brain neuroscience in larval zebrafish". In: *Current Opinion in Neurobiology* 50 (2018). ISSN: 18736882. DOI: 10.1016/j.conb.2018.02.004.

[13] Xiuye Chen et al. "Brain-wide Organization of Neuronal Activity and Convergent Sensorimotor Transformations in Larval Zebrafish". In: *Neuron* 100.4 (2018), pp. 876–890. ISSN: 10974199. DOI: 10.1016/j.neuron.2018.09.042.

[14] Kevin Mann, Courtney L Gallen, and Thomas R Clandinin. "Whole-Brain Calcium Imaging Reveals an Intrinsic Functional Network in Drosophila". In: *Current Biology* 27.15 (2017), pp. 2389–2396. ISSN: 09609822. DOI: 10.1016/j.cub.2017.06.076. URL: http://dx.doi.org/10.1016/j.cub.2017.06.076.

[15] Sophie Aimon et al. "Fast near-whole–brain imaging in adult Drosophila during responses to stimuli and behavior". In: *PLoS biology* 17.2 (2019), e2006732.

[16] Kevin Mann et al. "Coupling of activity, metabolism and behaviour across the Drosophila brain". In: *Nature* 593.7858 (2021), pp. 244–248.

[17] Diego A Pacheco et al. "Auditory activity is diverse and widespread throughout the central brain of Drosophila". In: *Nat Neurosci* 24 (2021), pp. 93–104. DOI: https://doi.org/10.1038/s41593-020-00743-y.

[18] Albert Lin et al. "Imaging whole-brain activity to understand behaviour". In: *Nature Reviews Physics* 0123456789 (2022). DOI: 10.1038/s42254-022-00430-w.

[19] Luke E. Brezovec et al. "Mapping the Neural Dynamics of Locomotion across the Drosophila Brain". In: *bioRxiv* (2022). DOI: 10.1101/2022.03.20.485047. eprint: https://www.biorxiv.org/content/early/2022/03/21/2022.03.20.485047.full.pdf. URL: https://www.biorxiv.org/content/early/2022/03/21/2022.03.20.485047.

[20] Sophie Aimon et al. "Global change in brain state during spontaneous and forced walk in Drosophila is composed of combined activity patterns of different neuron classes". In: *Elife* 12 (2023), e85202.

[21] Evan S Schaffer et al. "Flygenvectors: The spatial and temporal structure of neural activity across the fly brain". In: *bioRxiv* (Sept. 2021).

[22] Jeffrey Demas et al. "High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy". In: *Nature Methods* 18.9 (2021), pp. 1103–1111.

[23] Owen Randlett et al. "Whole-brain activity mapping onto a zebrafish brain atlas". In: *Nature methods* 12.11 (2015), pp. 1039–1046.

[24] Michael Kunst et al. "A cellular-resolution atlas of the larval zebrafish brain". In: *Neuron* 103.1 (2019), pp. 21–38.

[25] Seung Wook Oh et al. "A mesoscale connectome of the mouse brain". In: *Nature* 508.7495 (2014), pp. 207–214.

[26] Katrin Amunts et al. "Julich-Brain: A 3D probabilistic atlas of the human brain's cytoarchitecture". In: *Science* 369.6506 (2020), pp. 988–992.

[27] Yiming Xiao et al. "Evaluation of MRI to ultrasound registration methods for brain shift correction: the CuRIOUS2018 challenge". In: *IEEE transactions on medical imaging* 39.3 (2019), pp. 777–786.

[28] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. "Mutual-information-based registration of medical images: a survey". In: *IEEE transactions on medical imaging* 22.8 (2003), pp. 986–1004.

[29] Lei Qu et al. "Cross-modal coherent registration of whole mouse brains". In: *Nature Methods* 19.1 (2022), pp. 111–118.

[30] Ignacio Arganda-Carreras et al. "A Statistically Representative Atlas for Mapping Neuronal Circuits in the Drosophila Adult Brain". In: *Frontiers in Neuroinformatics* 12 (2018). ISSN: 1662-5196. DOI: 10.3389/fninf.2018.00013. URL: https://www.frontiersin.org/articles/10.3389/fninf.2018.00013.

[31] Louis K Scheffer et al. "A connectome and analysis of the adult Drosophila central brain". In: *Elife* 9 (2020).

[32] Sven Dorkenwald et al. "FlyWire: online community for whole-brain connectomics". In: *Nature Methods* 19.1 (2022), pp. 119–128.

[33] Hanchuan Peng et al. "BrainAligner: 3D registration atlases of Drosophila brains". In: *Nature methods* 8.6 (2011), pp. 493–498.

[34] Jody Clements et al. "NeuronBridge: an intuitive web application for neuronal morphology search across large data sets". In: *bioRxiv* (2022), pp. 2022–07.

[35] Daniel Münch, Dennis Goldschmidt, and Carlos Ribeiro. "The neuronal logic of how internal states control food choice". en. In: *Nature* (July 2022), pp. 1–9.

[36] Brian B Avants et al. "A reproducible evaluation of ANTs similarity metric performance in brain image registration". In: *Neuroimage* 54.3 (2011), pp. 2033–2044.

[37] Brian B Avants et al. "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain". In: *Medical image analysis* 12.1 (2008), pp. 26–41.

[38] John A Bogovic et al. "An unbiased template of the Drosophila brain and ventral nerve cord". In: *Plos one* 15.12 (2020), e0236495.

[39] Malte Hoffmann et al. "SynthMorph: learning contrast-invariant registration without acquired images". In: *IEEE transactions on medical imaging* 41.3 (2021), pp. 543–558.

[40] Torsten Rohlfing. "Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable". In: *IEEE transactions on medical imaging* 31.2 (2011), pp. 153–163.

[41] Thorvald A Sørensen. "A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons". In: *Biol. Skar.* 5 (1948), pp. 1–34.

[42] Lee R Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[43] Louis K. Scheffer et al. "A connectome and analysis of the adult drosophila central brain". In: *eLife* 9 (2020), pp. 1–74. ISSN: 2050084X. DOI: 10.7554/ELIFE.57443.

[44] Zhihao Zheng et al. "A Complete Electron Microscopy Volume of the Brain of Adult Drosophila melanogaster". In: *Cell* 174.3 (2018), 730–743.e22. ISSN: 0092-8674. DOI: https://doi.org/10.1016/j.cell.2018.06.019. URL: https://www.sciencedirect.com/science/article/pii/S0092867418307876.

[45] Jefferis G.S. et al. "Comprehensive Maps of Drosophila Higher Olfactory Centers: Spatially Segregated Fruit and Pheromone Representation". In: *Cell* 128.6 (2007), pp. 1187–1203.

[46] Maxwell H Turner et al. "Visual and motor signatures of locomotion dynamically shape a population code for feature detection in Drosophila". In: *Elife* 11 (2022), e82587.

[47] Maxwell H Turner, Kevin Mann, and Thomas R Clandinin. "The connectome predicts resting-state functional connectivity across the Drosophila brain". In: *Current Biology* 31.11 (2021), pp. 2386–2394.

[48] Janne K Lappalainen et al. "Connectome-constrained deep mechanistic networks predict neural responses across the fly visual system at single-neuron resolution". In: *bioRxiv* (2023), pp. 2023–03.

[49] Lu Mi et al. "Connectome-constrained Latent Variable Model of Whole-Brain Neural Activity". In: *International Conference on Learning Representations*.

[50] Felix Mölder et al. "Sustainable data analysis with Snakemake". In: *F1000Research* 10 (2021).

[51] Stéfan van der Walt et al. "scikit-image: image processing in Python". In: *PeerJ* 2 (June 2014), e453. ISSN: 2167-8359. DOI: 10.7717/peerj.453. URL: https://doi.org/10.7717/peerj.453.

[52] Karel Zuiderveld. "Contrast limited adaptive histogram equalization". In: *Graphics gems* (1994), pp. 474–485.

[53] Brian Avants and James C Gee. "Geodesic estimation for large deformation anatomical shape averaging and interpolation". In: *Neuroimage* 23 (2004), S139–S150.

[54] Brian B Avants et al. "The optimal template effect in hippocampus studies of diseased populations". In: *Neuroimage* 49.3 (2010), pp. 2457–2466.

[55] The HDF Group. *Hierarchical Data Format, version 5*. 1997-2023. URL: https://www.hdfgroup.org/HDF5/ (visited on 05/18/2023).

[56] Paul A. Yushkevich et al. "User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability". In: *Neuroimage* 31.3 (2006), pp. 1116–1128.

[57] *ITK-SNAP*. 2006-2023. URL: http://www.itksnap.org (visited on 05/18/2023).

[58] Arnim Jenett et al. "A GAL4-driver line resource for Drosophila neurobiology". In: *Cell reports* 2.4 (2012), pp. 991–1001.

[59] *neuprint*. URL: https://neuprint.janelia.org/ (visited on 05/24/2023).

[60] *codex*. URL: https://codex.flywire.ai/ (visited on 05/24/2023).

[61] *navis flybrains*. URL: https://github.com/navis-org/navis-flybrains (visited on 05/24/2023).

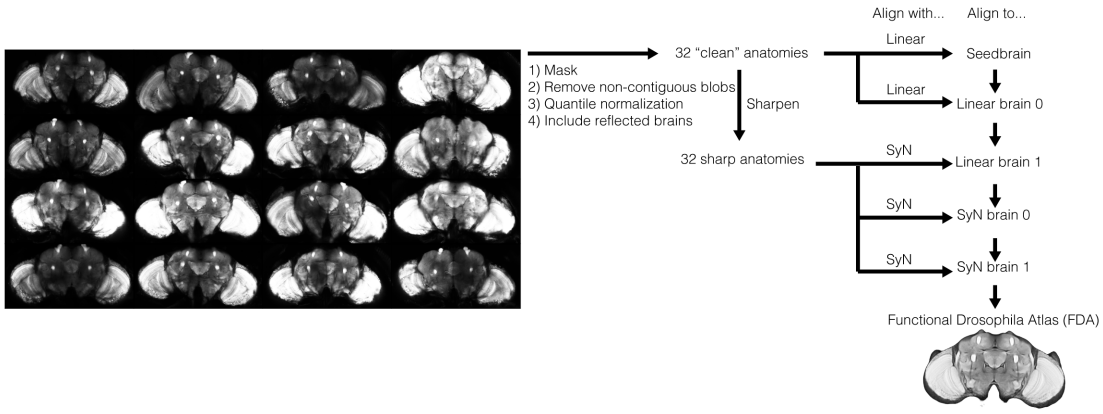

Figure S1. Creation of the Functional Drosophila Atlas (FDA). Full details in methods.

Supplement

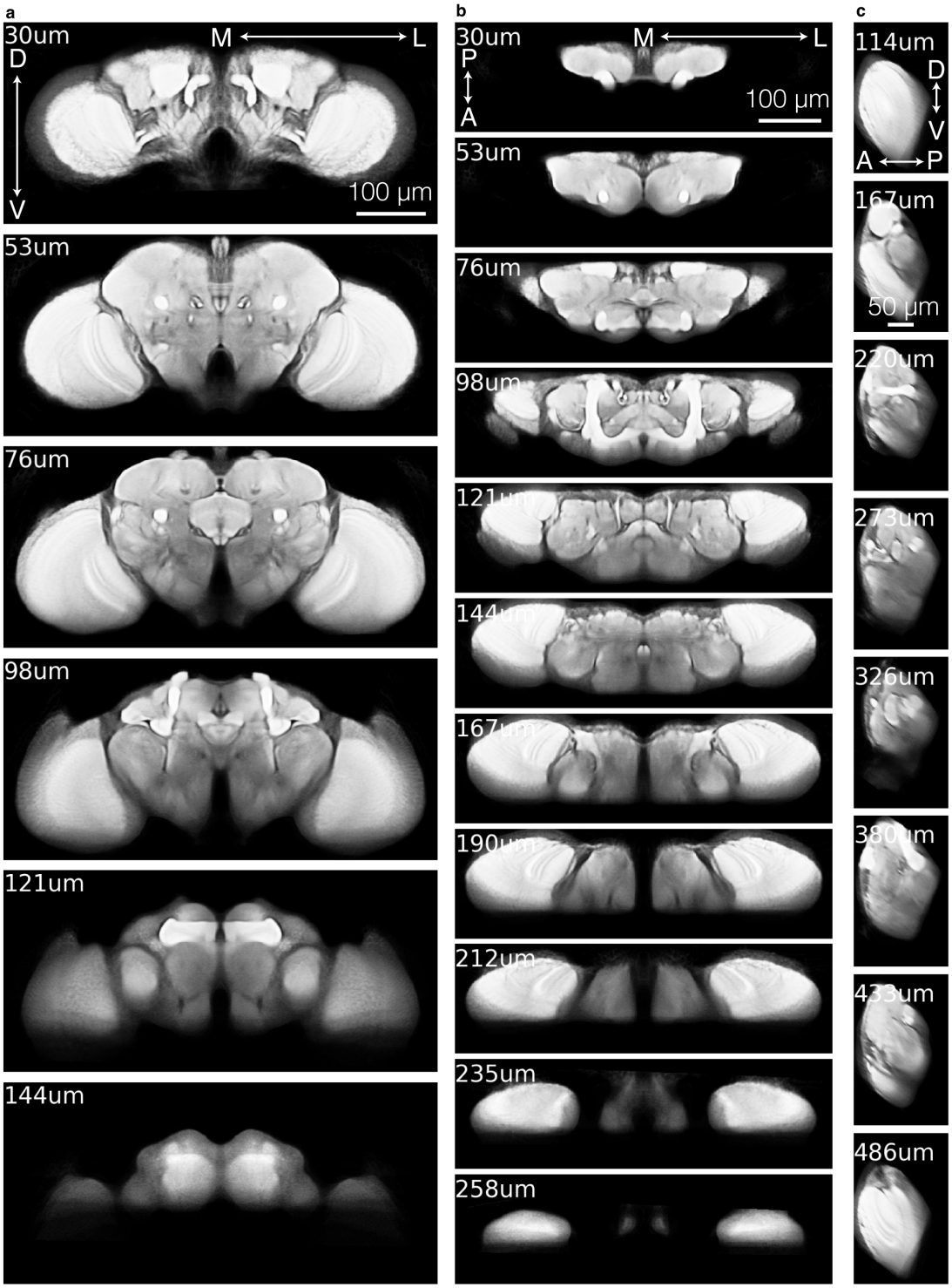

Figure S2. The Functional Drosophila Atlas (FDA). a, Slices through FDA, moving along the anterior-posterior axis. Micron labels indicates depth along the axis. **b**, Same as **a**, except along the dorsal-ventral axis. **c**, Same as **a**, except along the medial-lateral axis.

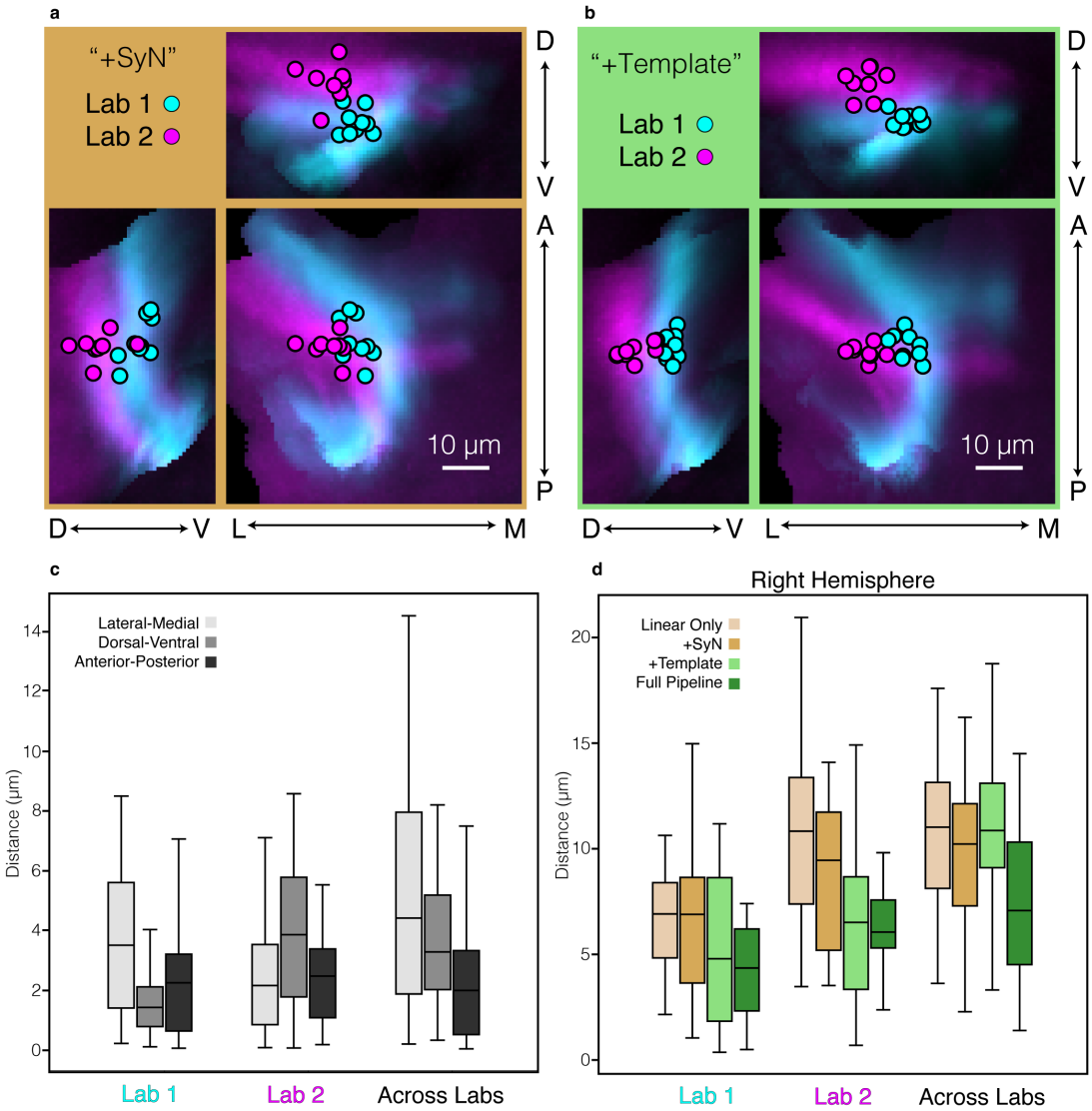

Figure S3. Quantifying LC11 registration accuracy. a, Same as figure 2D, except showing the "SyN" pipeline complexity. **b**, Same as figure 2d, except showing "With Template" complexity. **c**, Full pipeline complexity, showing centroid distances along each orthogonal axis. Box center line indicates median, box limits indicate quartiles, whiskers indicate 1.5x the inter-quartile range. **d**, Same as Figure 2e, except showing results from the other hemisphere.

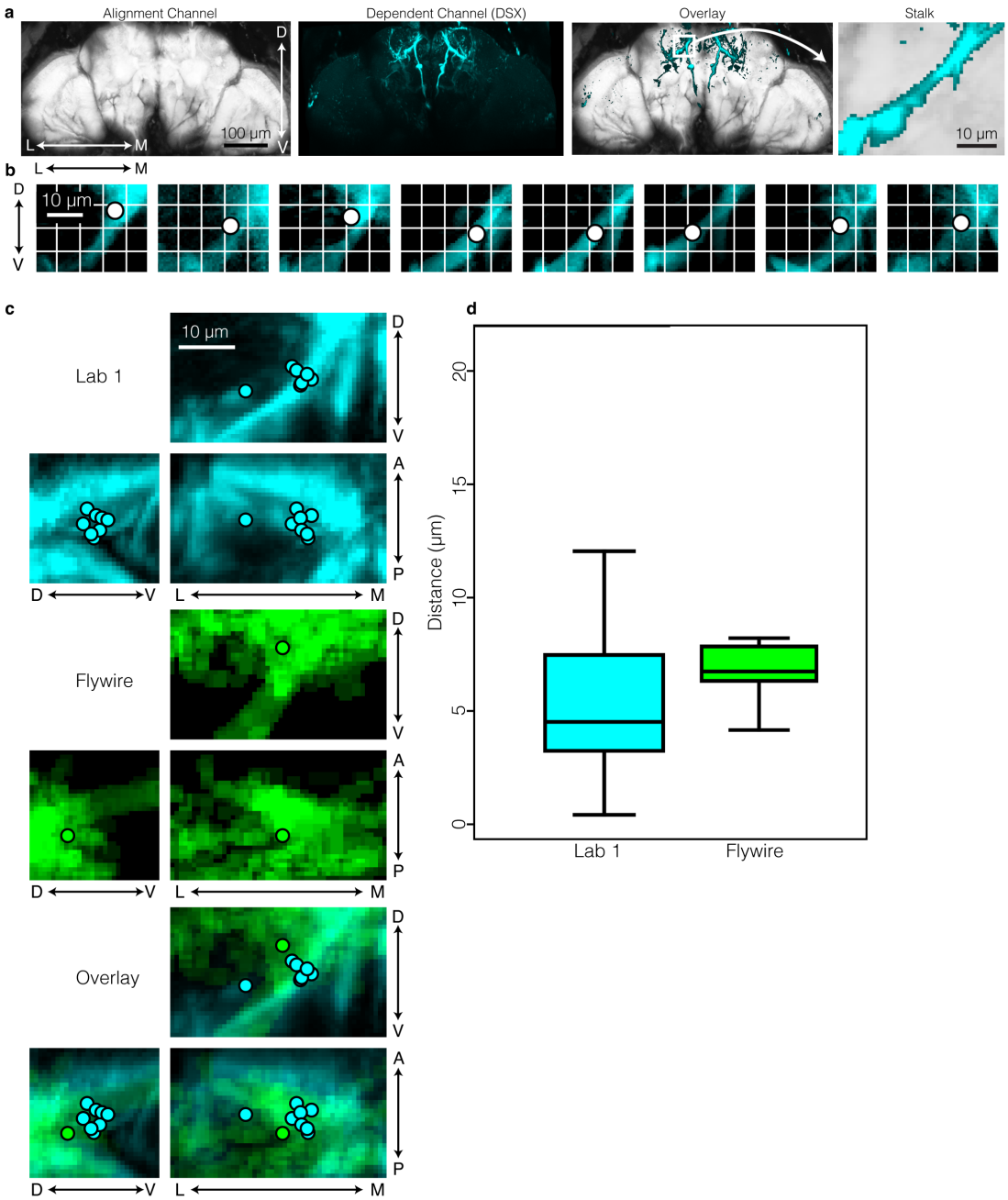

Figure S4. Quantifying DSX registration accuracy. a, Example DSX fly showing the alignment channel (myr-tdtomato) and dependent channel (DSX). Zoom shows DSX stalk region that will be quantified. **b**, Zoom in of DSX stalk region in FDA showing DSX expression of individual animals after the full pipeline was applied. Dot indicates centroid of each stalk, which will be used to quantify distance. **c**, Same as in B, but animals are overlayed and projections along each axis are shown. **d**, Quantification of pair-wise centroid distances. Box center line indicates median, box limits indicate quartiles, whiskers indicate 1.5x the inter-quartile range.