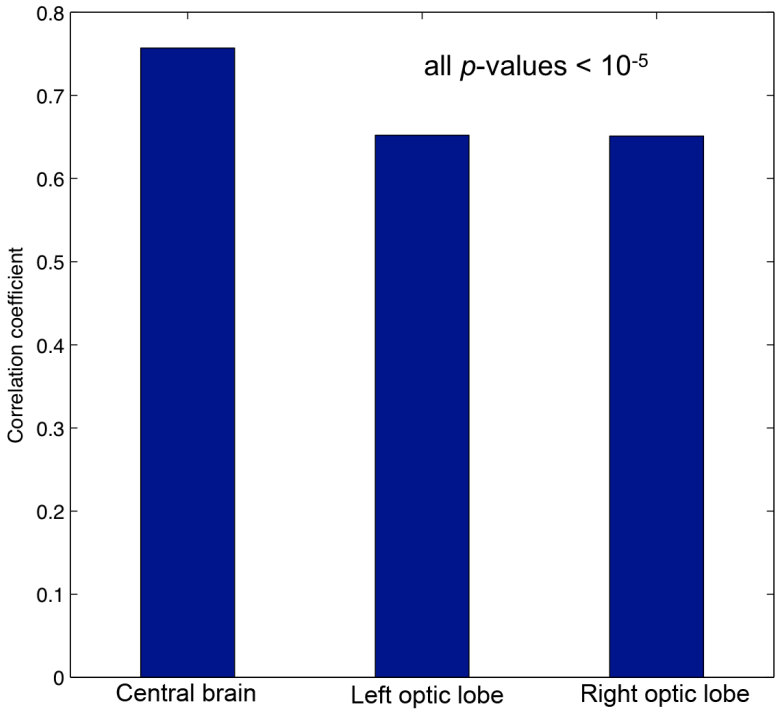
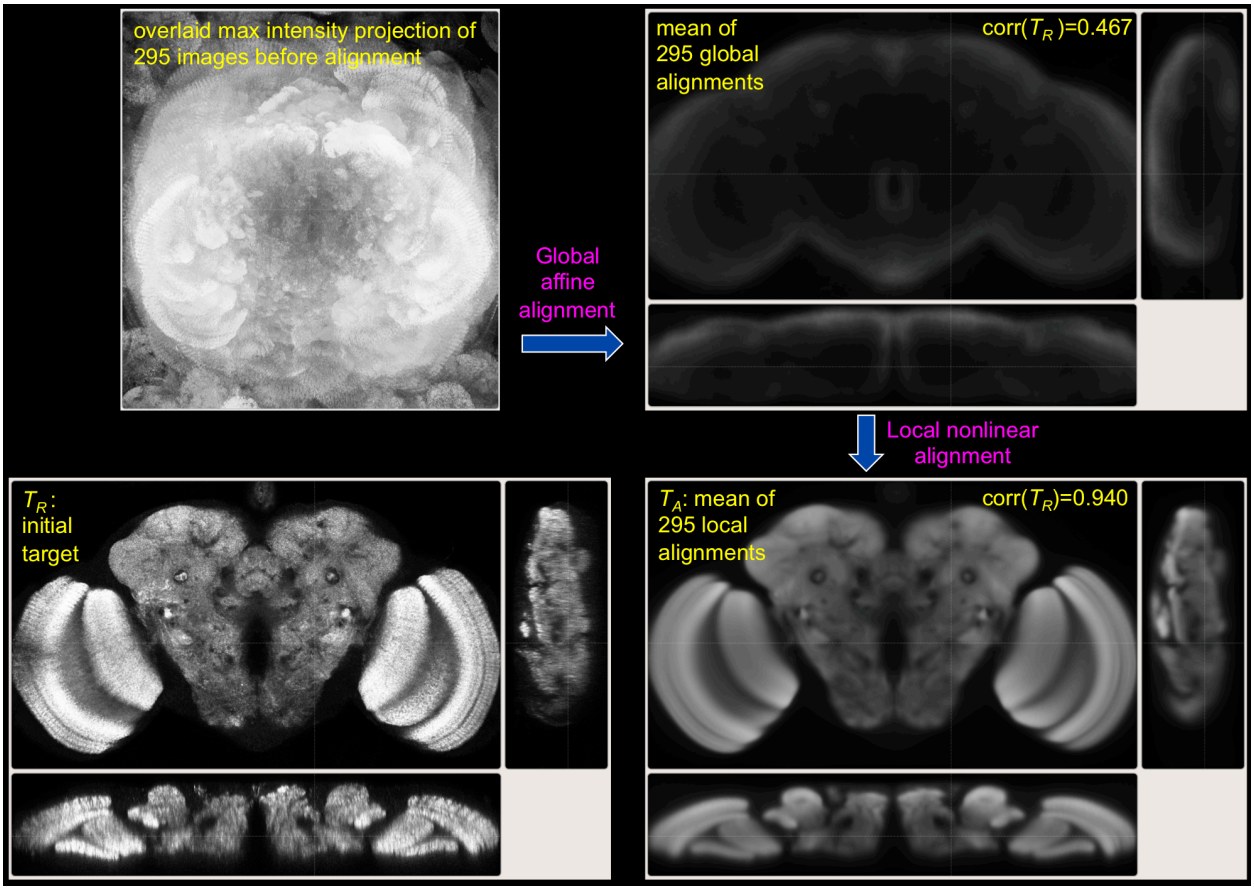
Supplementary Figure 1



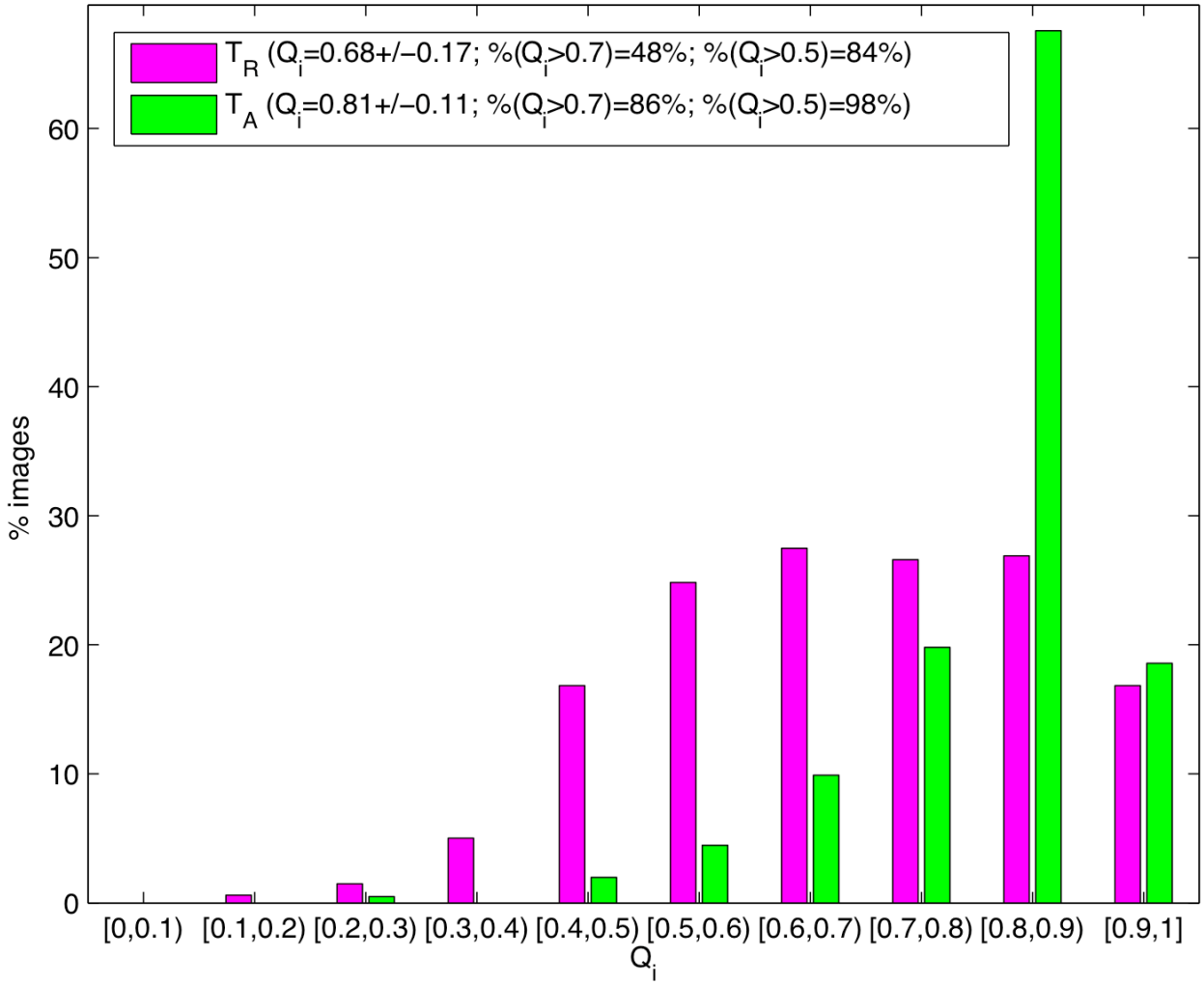
Supplementary Figure 1. Correlation between the automatically and manually generated alignment quality scores Q_i and Q_v of 805 alignments. The central brain, left and right optic lobes were scored separately to best assess the consistency of the two alignment quality scores in different brain areas.

Supplementary Figure 2



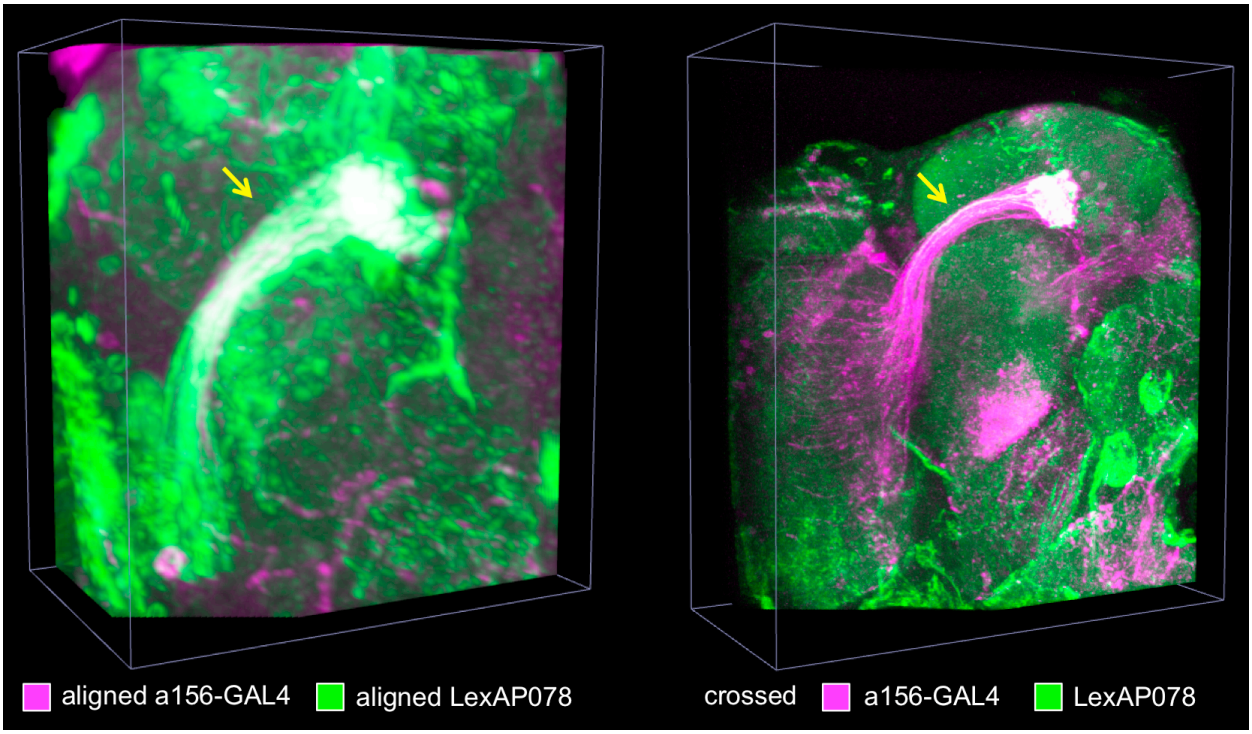
Supplementary Figure 2. Choosing the target brain. 295 images were aligned to an initial target brain T_R. The mean image of all aligned brains, T_A, was chosen as the new target brain. Initially all brains had different orientations and sizes, and thus overlapping them did not yield a meaningful pattern. Through the global alignment these brains were transformed to have the same center of mass, similar orientations and sizes. However, overlapping them did not yet produce a bright consensus and sharp border in the average brain. Once the local alignment was performed, we obtained an average brain (T_A) that was much brighter and less fuzzy than that of the globally registered brains. The correlation (= 0.940) between T_A and T_R was substantially higher than that (= 0.467) between the mean of the globally aligned brains and T_R, because of the good local alignment.

Supplementary Figure 3



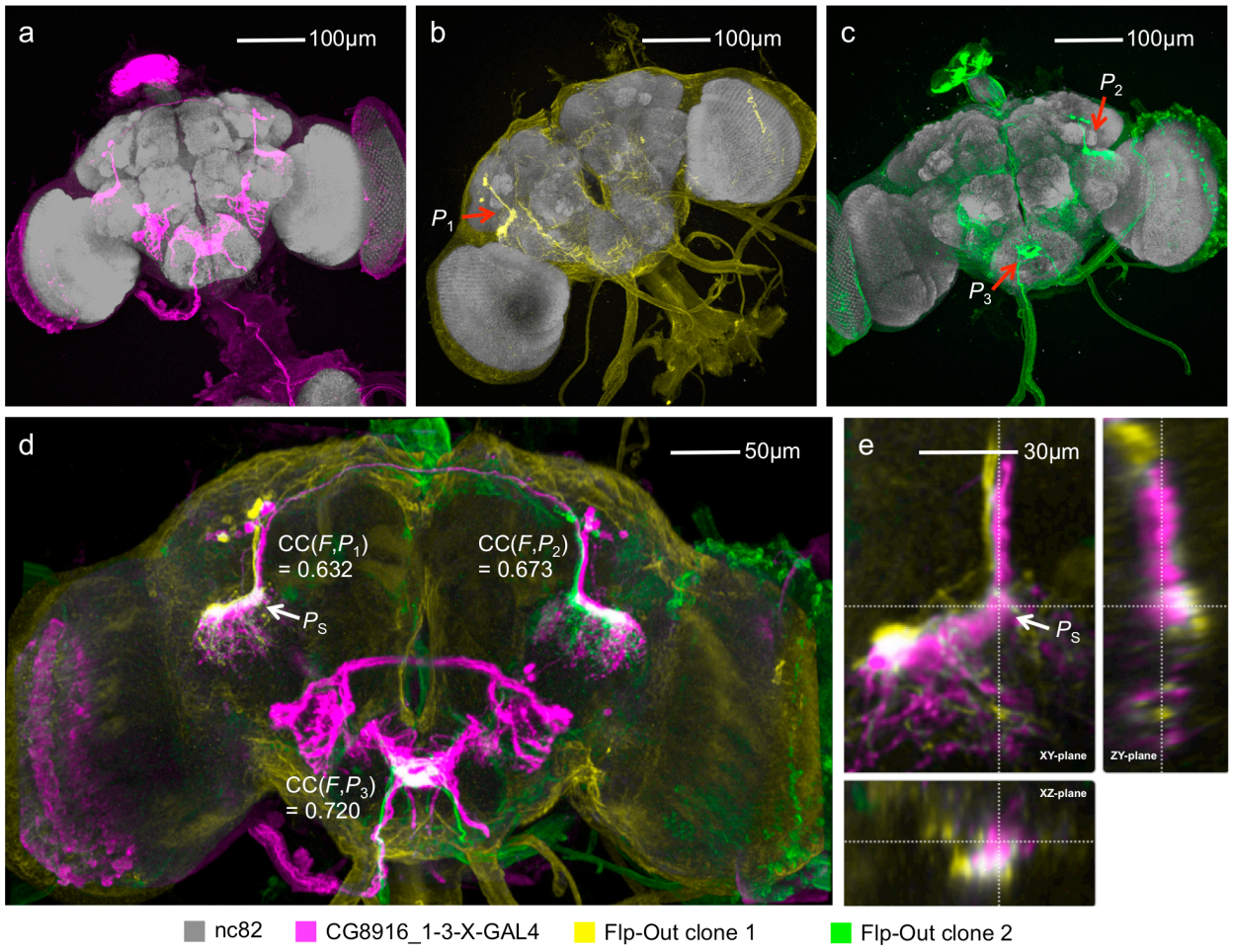
Supplementary Figure 3. Improvement of alignment of 496 brain images by replacing T_R using T_A.

Supplementary Figure 4.



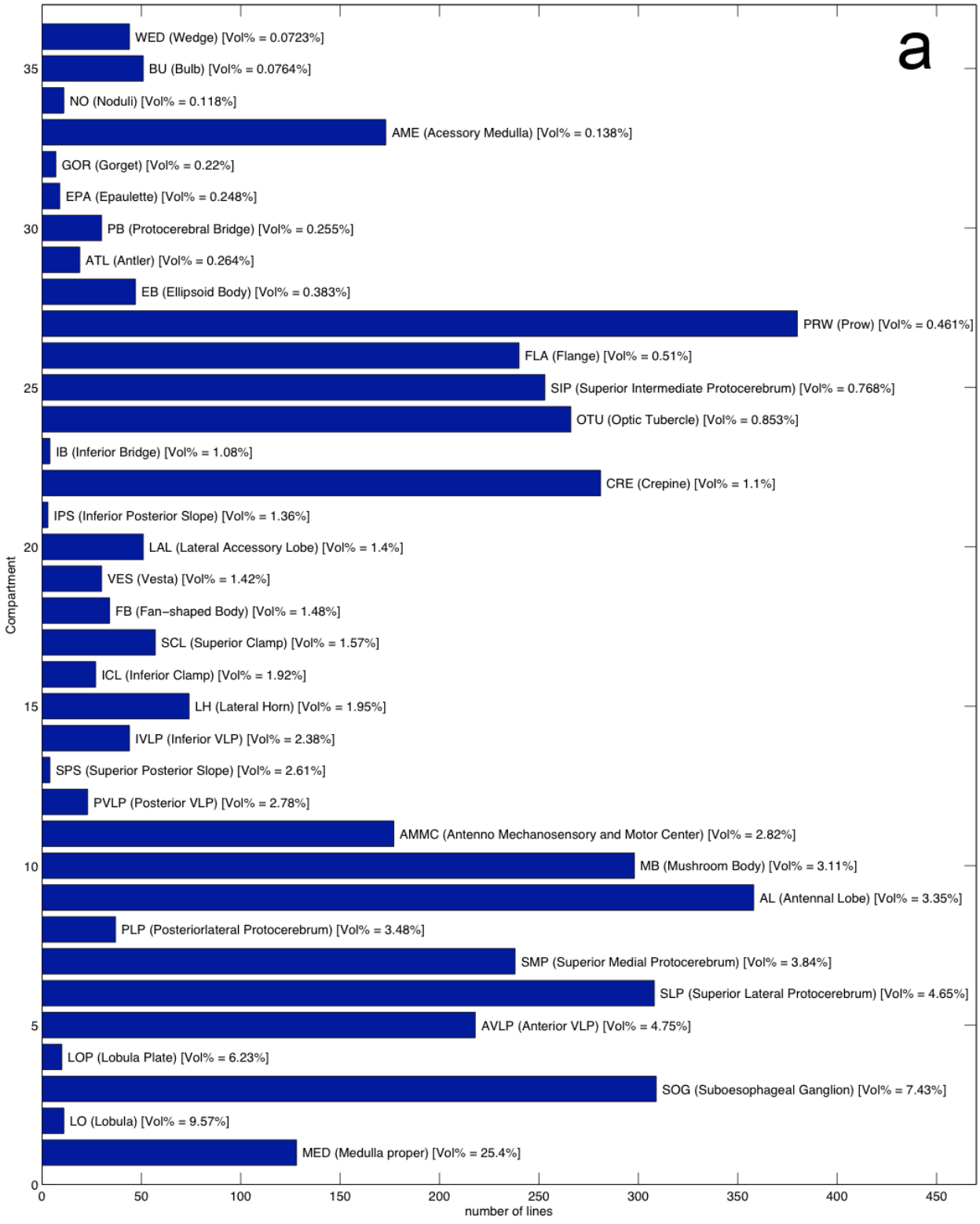
Supplementary Figure 4. Comparison of computational and biological methods to determine expression pattern overlap of the aligned and co-expressed patterns of *a156-GAL4; UAS-mCD8-GFP* and *LexAP078; lexAop-CD2-GFP*. Yellow arrows, the most overlapping regions.

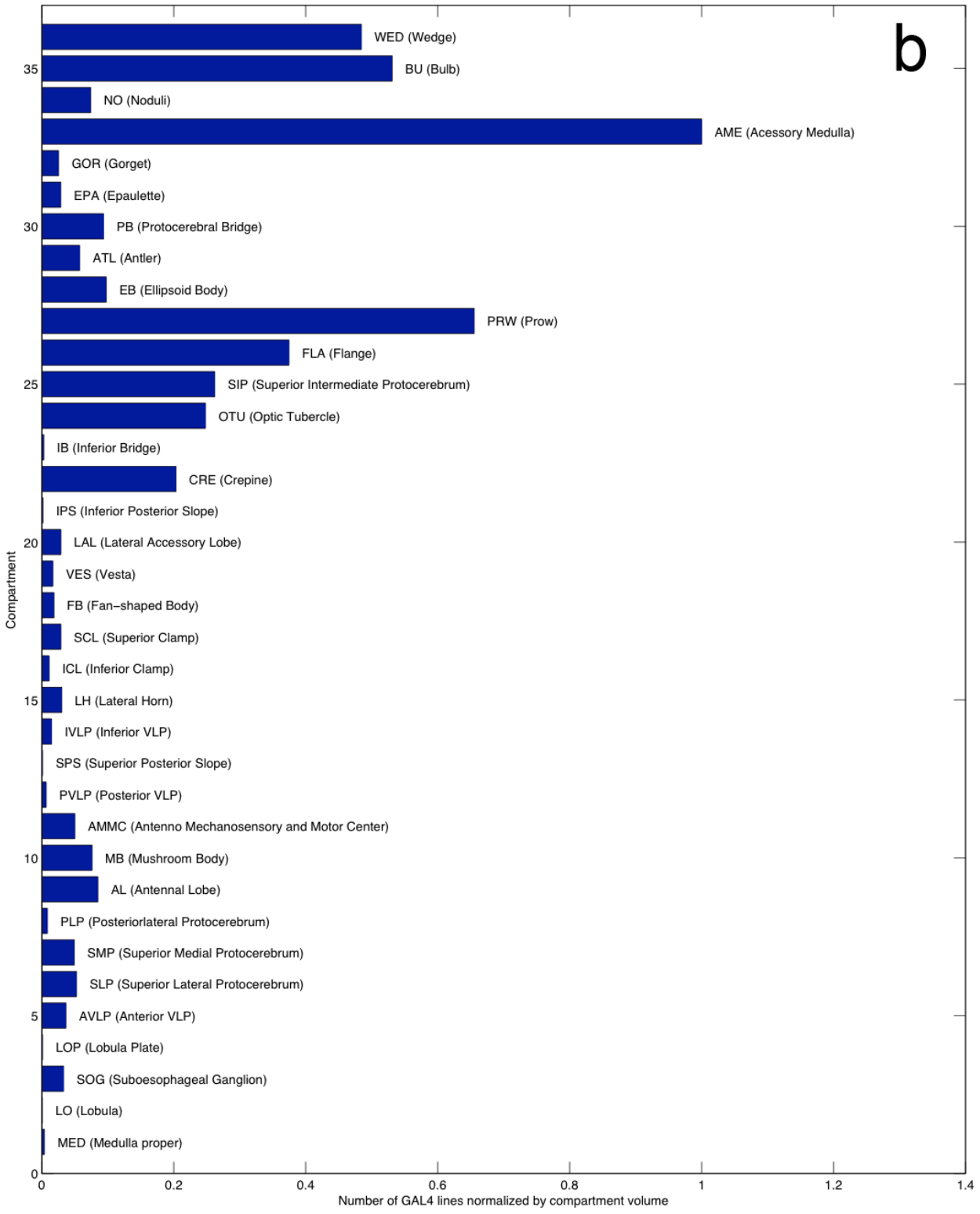
Supplementary Figure 5.

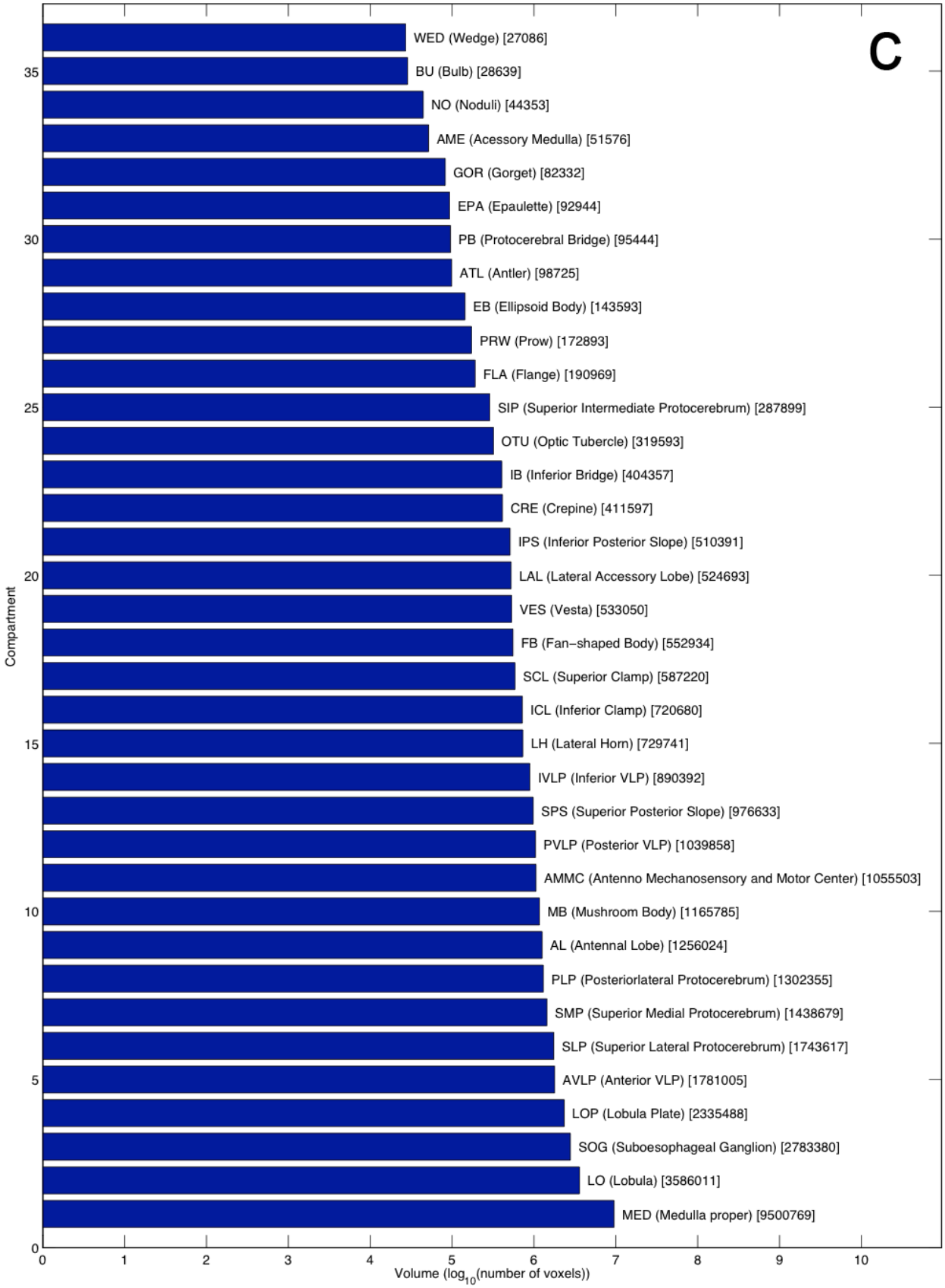


Supplementary Figure 5. Comparison of obligate subsets with parent expression patterns. (**a**) Original GAL4 pattern (magenta) of *CG8916_1-3-X-GAL4; UAS-mCD8-GFP*. (**b-c**) Two example Flp-Out clones (yellow and green) of this GAL4 line. Red arrows and P_1, P_2, P_3: three obligate subset-patterns. (**d**) Aligned original pattern and obligate Flp-Out subsets show substantial overlap. *F*: the parent pattern. CC: the regional correlation coefficients of patterns P_1, P_2, and P_3 and the parent pattern. All *p*-values of calculated correlation statistics are less than 10^{-5}. P_S: a location for zoom-in inspection. (**e**) Cross-sectional view the region around P_S in (d) shows co-localization of the obligate subset and the parent GAL4 pattern.

Supplementary Figure 6.




b
WED (Wedge)
BU (Bulb)
NO (Noduli)
AME (Acessory Medulla)
GOR (Gorget)
EPA (Epaulette)
PB (Protocerebral Bridge)
ATL (Antler)
EB (Ellipsoid Body)
PRW (Prow)
FLA (Flange)
SIP (Superior Intermediate Protocerebrum)
OTU (Optic Tubercle)
IB (Inferior Bridge)
CRE (Crepine)
IPS (Inferior Posterior Slope)
LAL (Lateral Accessory Lobe)
VES (Vesta)
FB (Fan–shaped Body)
SCL (Superior Clamp)
ICL (Inferior Clamp)
LH (Lateral Horn)
IVLP (Inferior VLP)
SPS (Superior Posterior Slope)
PVLP (Posterior VLP)
AMMC (Antenno Mechanosensory and Motor Center)
MB (Mushroom Body)
AL (Antennal Lobe)
PLP (Posteriorlateral Protocerebrum)
SMP (Superior Medial Protocerebrum)
SLP (Superior Lateral Protocerebrum)
AVLP (Anterior VLP)
LOP (Lobula Plate)
SOG (Suboesophageal Ganglion)
LO (Lobula)
MED (Medulla proper)
Compartment
0
5
10
15
20
25
30
35
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Number of GAL4 lines normalized by compartment volume


C
WED (Wedge) [27086]
BU (Bulb) [28639]
NO (Noduli) [44353]
AME (Acessory Medulla) [51576]
GOR (Gorget) [82332]
EPA (Epaulette) [92944]
PB (Protocerebral Bridge) [95444]
ATL (Antler) [98725]
EB (Ellipsoid Body) [143593]
PRW (Prow) [172893]
FLA (Flange) [190969]
SIP (Superior Intermediate Protocerebrum) [287899]
OTU (Optic Tubercle) [319593]
IB (Inferior Bridge) [404357]
CRE (Crepine) [411597]
IPS (Inferior Posterior Slope) [510391]
LAL (Lateral Accessory Lobe) [524693]
VES (Vesta) [533050]
FB (Fan−shaped Body) [552934]
SCL (Superior Clamp) [587220]
ICL (Inferior Clamp) [720680]
LH (Lateral Horn) [729741]
IVLP (Inferior VLP) [890392]
SPS (Superior Posterior Slope) [976633]
PVLP (Posterior VLP) [1039858]
AMMC (Antenno Mechanosensory and Motor Center) [1055503]
MB (Mushroom Body) [1165785]
AL (Antennal Lobe) [1256024]
PLP (Posteriorlateral Protocerebrum) [1302355]
SMP (Superior Medial Protocerebrum) [1438679]
SLP (Superior Lateral Protocerebrum) [1743617]
AVLP (Anterior VLP) [1781005]
LOP (Lobula Plate) [2335488]
SOG (Suboesophageal Ganglion) [2783380]
LO (Lobula) [3586011]
MED (Medulla proper) [9500769]
Compartment
0
5
10
15
20
25
30
35
0
1
2
3
4
5
6
7
8
9
10
Volume (log10(number of voxels))

Supplementary Figure 6. Analysis of a 3D neuronal expression pattern atlas of a large collection of aligned GAL4 lines. (**a**) Distribution of GAL4 expressions in different brain compartments. The volume percentages of compartments in the target brain are shown in the brackets following the compartment names. (**b**) Compartment-volume normalized distribution of GAL4 expressions. The statistic was rescaled to have a range of [0, 1], for better visualization. (**c**) The volume of different compartments in our target brain. In the bracket the number of image voxels in a compartment is also displayed.

Supplementary Note.

"Comparison of several different aligners on LSM images of *Drosophila* brains"

In the following note, we compare the performance of a few established 3D image registration methods and BrainAligner for laser scanning microscope (LSM) images of *Drosophila* brains. We note that different alignment methods are optimized for different problems or datasets. They have their own strength in different applications. This supplementary piece only intends to show why some of the existing approaches are not suitable for our application, and how their key ideas have been integrated in our system.

We evaluate the available methods for alignment accuracy on our data type, speed, and error reporting or quality assessment. Wherever possible we used our target brain and sample data for fair comparison. Some algorithms required us to down-sample the data, and we indicate below where this was done. Based on this comparison, we tried to incorporate the best features of previous approaches with our new methods for optimum performance on our data.

Methods considered

Most image registration methods deal with either 2D images or 3D images. Compared to 3D registration methods, 2D methods often optimize in a much more constrained search space. BrainAligner is a 3D registration program. Thus most 2D tools are not suitable for a meaningful comparison for real applications.

We compare 9 existing automatic methods to BrainAligner. They include 5 widely used 3D

image registration approaches for general biomedical images, which are

1. Demons (Thirion, 1998),
2. Diffeomorphic Demons (Vercauteren et al 2008),
3. Free-form deformation (FFD) registration (Rueckert, et al 1999),
4. Hierarchical local affine registration (Periaswamy and Farid, 2006),
5. HAMMER (Shen et al 2002),

and 4 existing automatic registration software packages that can be directly used to register *Drosophila* brain LSM images, which are

6. CMTK: normalized mutual information (NMI) matching, mean of squared difference (MSD) and B-Spline interpolation (Jefferis, 2007; Rohlfing, et al, 2003),
7. K-FFD: sum of squared difference (SSD) registration using quasi-newton (L-BFGS) optimizer to solve FFD (Kroon, 2008),
8. Q-FFD: SSD registration using gradient descent to solve FFD (Qu and Peng, 2010a),
9. Principal skeleton warping (Qu and Peng, 2010b).

We also compared the pure manual method to determine the corresponding 3D landmarks for registration.

10. Manual landmark determination.

As a comprehensive study, we also investigated BrainAligner's ability to align partial brains.

Demons and diffeomorphic demons (methods #1 and #2)

Demons registration (Thirion, 1998) considers a diffusion process to best match the boundary of

a target image to that of a subject image. “demons”, or attractors, are defined in the target image, typically on its contour points, to attract the deformation of a subject image for best alignment. The Diffeomorphic Demons registration (Vercauteren, et al, 2008) is an “upgraded” version of the Demons method and optimizes the algorithm over a diffeomorphic space. Both Demons and Diffeomorphic Demons algorithms were tested using their “standard” ITK (www.itk.org) implementations. One result is shown in Fig. SN1, from which it becomes clear that the methods can align the contours of the subject image, but fail to produce the smoothly warped interior region of the image.

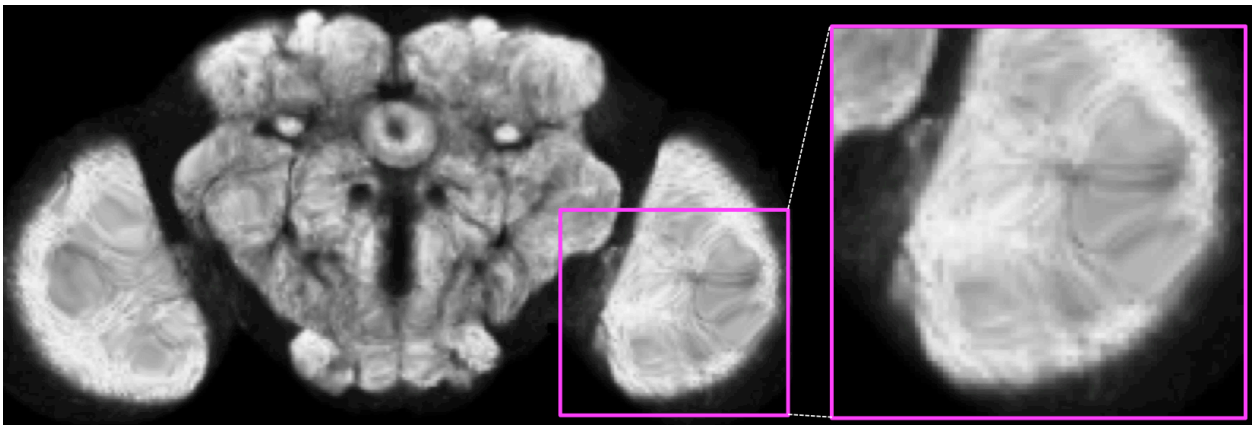


Figure SN1. A typical warped image produced by the Diffeomorphic Demons method (this particular result was produced with the help of Guorong Wu in Dinggang Shen lab, UNC Chapel Hill). During the alignment the interior of the image (e.g. the zoom-in box) was irregularly stretched.

BrainAligner is similar to Demons in that it also uses landmarks established at high curvature points, but Demons randomly assigns landmarks, while BrainAligner uses direct landmark matching of multiple criteria, consensus of these matching results, and direct thin-plate-spline warping. BrainAligner does not have the iterative process in Demons methods, which could be time-consuming. Also different from Demons algorithms, the warping of BrainAligner is smooth, given the faithful corresponding landmarks produced by BrainAligner’s reliable

landmark matching algorithm.

Free-form deformation (FFD) registration (methods #3, #6, #7, and #8)

FFD registration (Rueckert, et al 1999) uses B-spline for localized warping of a subject image to a target. A normalized mutual information score was often used to determine the best corresponding feature points defined in a grid in the target image and points from the subject image. The overall goal is to maximize the similarity of these two images after warping as well as maintaining the best possible smoothness of the warping field. In our implementation and testing, this method was able to detect many feature points that have been defined on a regular grid of the image. However, typically it was difficult to choose the parameters including the grid size, etc. It was quite easy to produce irregular stretching in the warped image, as shown in Fig. SN2. BrainAligner also uses mutual information as one of its methods to detect the corresponding landmarks. However, BrainAligner aggregates the results from multiple feature detectors instead of just one, making the detected landmarks more reliable. Another difference is that BrainAligner uses thin-plate-spline (TPS), instead of B-spline, for generating the warping field. TPS is unique given the corresponding landmarks. Thus BrainAligner has no difficulty in choosing the appropriate parameters for warping-field generation.

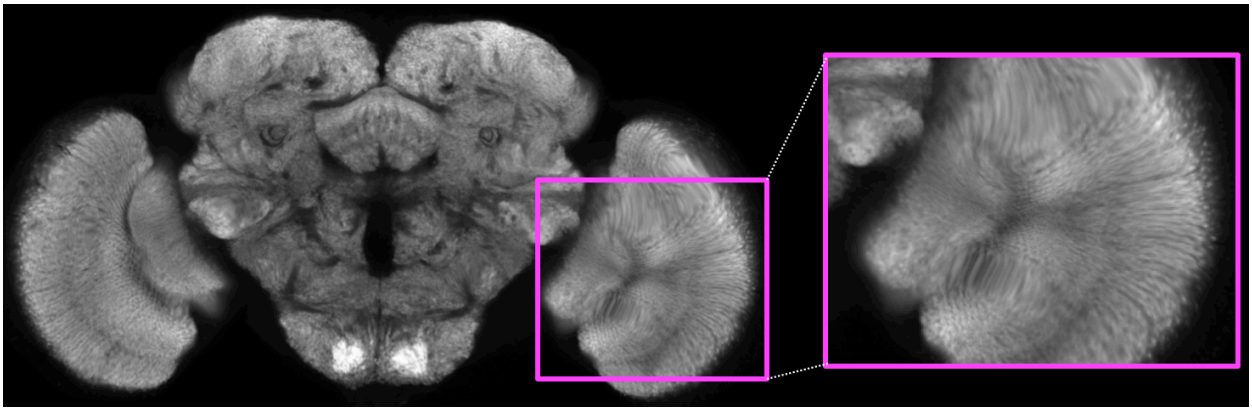
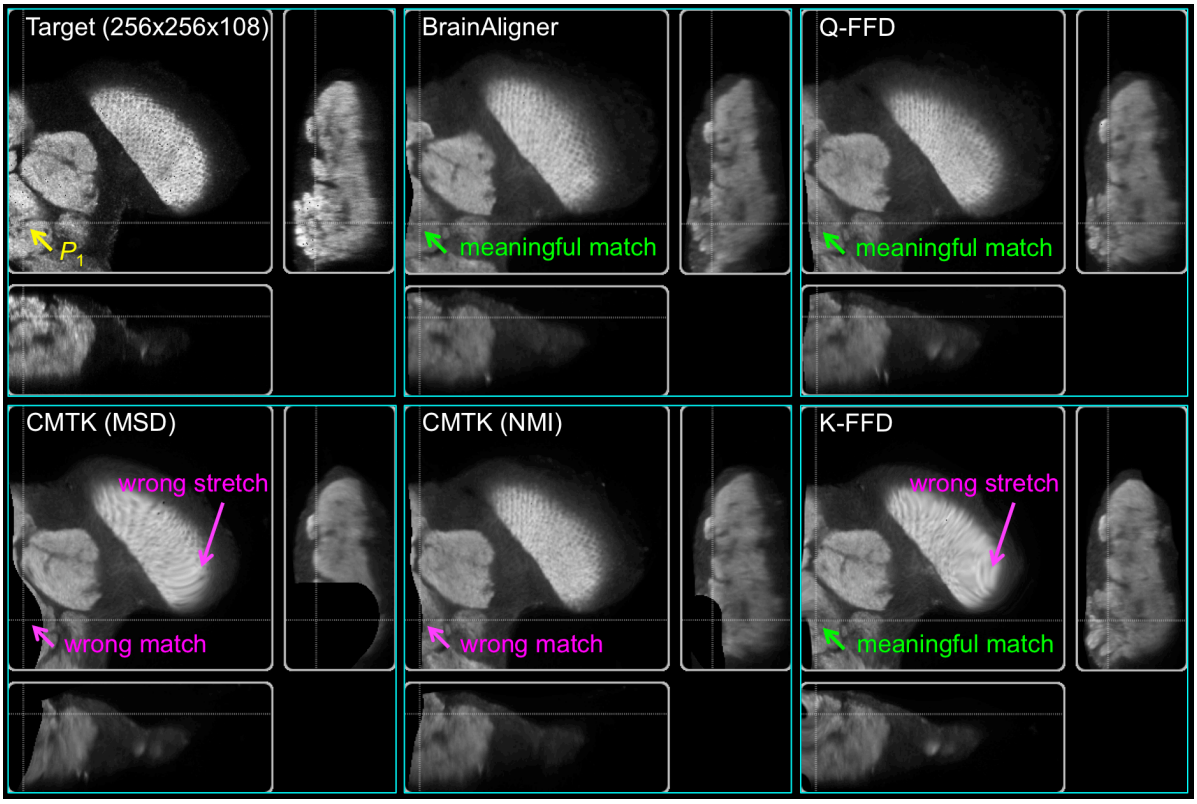
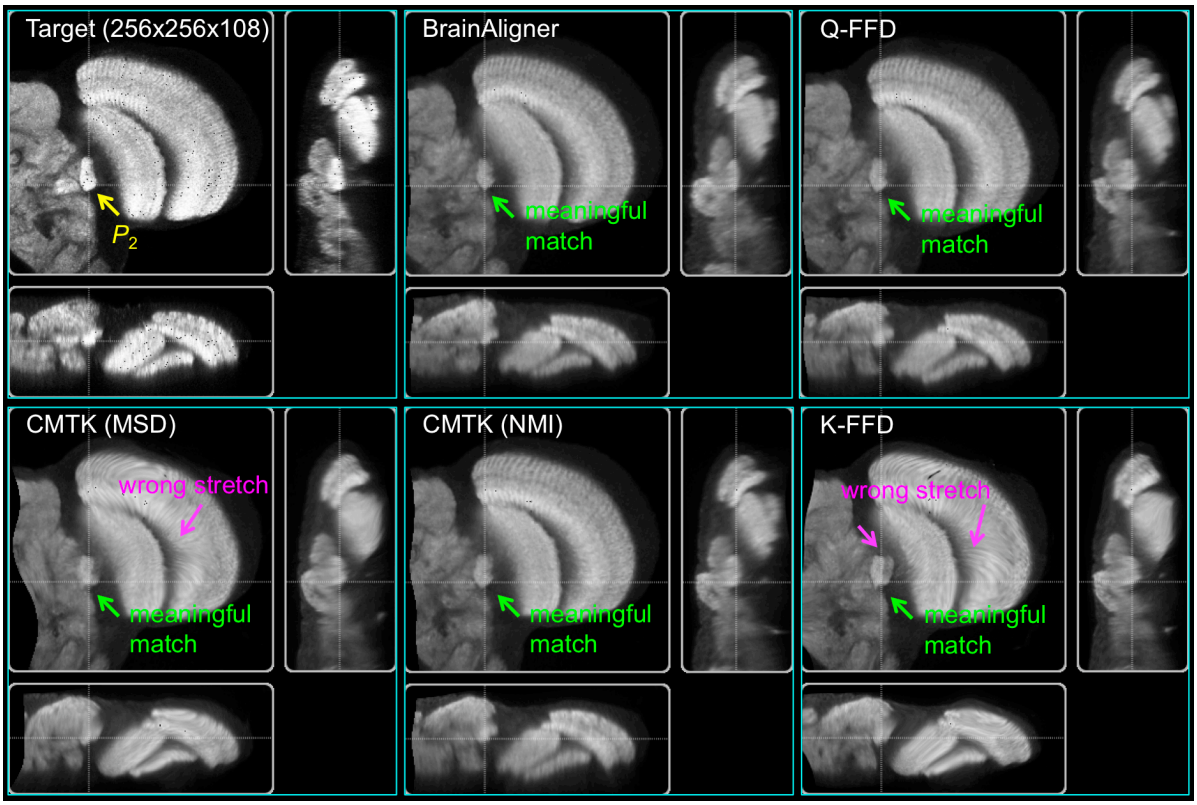


Figure SN2. A typical warped image produced by the free-form deformation using B-spline registration (based on our Q-FFD implementation in Qu and Peng, 2010a). During the alignment the interior of the image (e.g. the zoom-in box) was often irregularly stretched. It was typically hard to set appropriate parameters.

Recently, the Jefferis group has used an FFD variant, called CMTK (Rohlfing, et al, 2003), to register images of *Drosophila* brains automatically (Jefferis, et al, 2007; Cachero, et al, 2010). We bench-tested CMTK using its two settings (NMI, which is "normalized mutual information" matching and the default setting of the software, and MSD, which is "mean of squared difference" matching). For a comparison, we also tested two other FFD-variants, K-FFD (Kroon, 2008) and Q-FFD (Qu and Peng, 2010a). Fig. SN3 shows that CMTK and K-FFD produced irregular stretching or wrong matching of the brain images at different locations. CMTK (NMI) result was meaningful in many locations we checked (e.g. location P_2 in Fig. SN3 (b)), but still had errors (e.g. location P_1 in Fig. SN3 (a)). Q-FFD produced meaningful alignment for this case (but it could also produce wrong stretching in different cases such as Fig. SN2). BrainAligner successfully aligned the subject image to the target in this test. The wrong stretch of FFD is a natural result of the fact that FFD does not have a mechanism to prevent folding (self-intersection) of the warping field.



(a)



(b)

Figure SN3. Comparison of the aligned images produced by BrainAligner and several FFD variants. (a) and (b)

show two typical locations we checked the aligned images. Images tested have 256x256x108 voxels. Compared to the target image, BrainAligner was able to produce smoothly aligned subject image. Different FFD methods produced different results. The irregular stretching of the warped image is one the main problems of the FFD alignments (the other one is their speed, see Fig. SN4). Irregular stretching (also see the zoom-in view of Fig. SN2) is due to wrong matching at various locations.

Of note, all the three FFD variants ran very slowly on our data sets. When we tried to align a pair of typical LSM image stacks that have 1024x1024x*N* voxels (*N* around 160 in raw images; after our image-resampling to produce an isotropic voxel XYZ size which is 0.58μm, *N* is around 230), FFD methods did not produce meaningful alignments within days on a high-performance Linux machine with 16 cores and 48G memory. Even with downsampled images that have 512x512x108 voxels, the fastest FFD methods in our test still took more than 10 hours to produce some result (which were still unsatisfactory) while BrainAligner just took about 4 minutes to produce the meaningful alignment. Thus for the bench test purpose we tried images with 256x256x108 voxels. Fig. SN4 shows a comparison of running speed of all these methods and BrainAligner. Q-FFD was the fastest method (in terms of the total CPU time) among the three FFD variants. However, it was still about 380 times slower than BrainAligner. This is because BrainAligner optimizes a much smaller amount of parameters. The slow speed of FFD methods makes it difficult to use them in large-scale high-throughput experiments that need to align thousands of large LSM images.

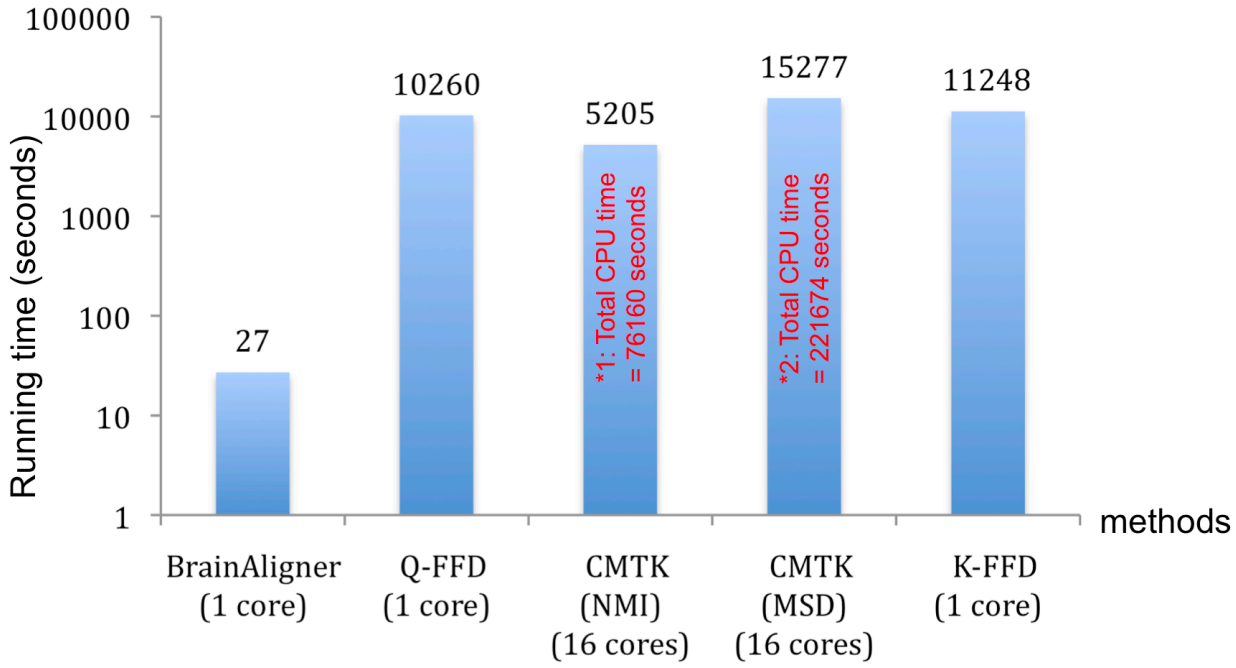


Figure SN4. Comparison of the running speed of BrainAligner and several FFD variants. In the target-subject image pairs, each image has 256x256x108 voxels. The test-computer had 16 cores (2 Intel Xeon 3.33GHz CPUs, each with 8 cores) and 48G memory. The total CPU time (sum of all cores) used by CMTK (NMI) and CMTK (MSD) is also shown.

Hierarchical local affine registration (method #4)

Hierarchical local affine registration (Periaswamy and Farid, 2006) builds on a powerful idea that a smooth nonlinear warping can be approximated by a series of local affine and linear transformations that are further aggregated in a hierarchical, multi-scale manner. For each affine transform, this method estimates all 12 affine transformation parameters plus intensity rescaling and offset factors simultaneously. The expectation maximization algorithm is used for optimization. However, this method is typically slow, and does not produce a robust estimation when there are more than ±30 degrees rotations of the corresponding image patches.

HAMMER (method #5)

Hierarchical attribute matching mechanism for elastic registration, or HAMMER (Shen et al 2002), is a powerful 3D image registration method that hierarchically searches for matching landmarks based on invariant 3D moments features of image patches. It is a good representative of a whole class of feature (attribute) vector based registration methods, where features that are invariant to rotation, shape rescaling, intensity rescaling, etc. are often extracted. However, because 3D searching of the best matching landmarks needs a huge amount of computation, and is also prone to noise, HAMMAR registration is typically done for images whose contours have been pre-extracted. Due to the large data volume this is impracticable for the LSM images in our study. We were unable to use HAMMER to produce promising alignments for our images. However, BrainAligner allows using the invariant moments features as one option to search for reliable landmarks.

Principal skeleton registration (method #9)

Principal skeleton registration (Qu, et al 2010) is an automatic method developed in our group. This method first detects the major curved skeletons of a shape and then warps the entire shape to best align these skeletons as well as some additional control points around these skeletons. This method was shown to be effective to deal with articulated image objects, a class of problems that were not handled by the above approaches. This method can often provide a very good initial alignment, which can be further improved by BrainAligner at an even finer scale.

Manual registration (method #10)

One of the widely used methods to register Drosophila brain images is to find the corresponding landmarks manually, followed by a warping (typically the thin-plate spline warp). Many tools, such as commercially available Amira (Visage Imaging), and free software package V3D (Peng, et al, 2010), may be used for this manual process, We experimentally compared BrainAligner to the pure manual determination of corresponding landmarks. We used V3D for this manual process because it has a powerful tool called Landmark Manager that allows us to browse, jump through, and link the 3D views of landmarks very efficiently.

For comparison, we considered two *a278*-GAL4 images as used in Figure 3. We denote them as S_a and S_b. We first used BrainAligner to automatically register them to our standard target brain T_A, and thus traced the neurite tracts. In T_A, we have 172 landmarks. Therefore, next we requested four annotators to manually, and independently, determine the 172 corresponding landmarks for each of S_a and S_b. Denote these landmark sets as L_{ij}, $i \in \{a, b\}$ and $j \in \{1, 2, 3, 4\}$ (for four annotators).

We found that even with the efficient manual tool, it is still very time-consuming to determine corresponding landmarks manually (Table SN1), which is currently about 4 ~ 6 times slower than the automatic landmark detection algorithm in BrainAligner. In addition, there is a substantial amount of variation of the landmark locations produced by different annotators (Table SN2). For the two images, the average distance between all corresponding landmarks produced by all annotators is 14.1±11.4 voxels, or 8.1±6.6 μm.

Table SN1. Manual time (minutes) spent for determining the subject-image landmarks that correspond to the 172 predefined landmarks in our standard target image. For the same setting, currently BrainAligner typically needs

about 25~30 minutes to automatically determines these corresponding landmarks.

Annotator	Image S_a	Image S_b
1	190 min	105 min
2	180 min	127 min
3	171 min	141 min
4	176 min	140 min
Average	179 min	128 min

Table SN2. Average distances and the respective standard deviations of the corresponding landmarks' spatial locations. For each of the test images S_a and S_b, the pair-wise distance statistics were calculated and shown (since each statistic is symmetrical, we only show half of them for simplicity). "vx" means voxels (each has a size 0.58μm×0.58μm×0.58μm).

Image S_a				Image S_b			
Annotator	2	3	4	Annotator	2	3	4
1	15.1±14.9vx	17.3±14.9vx	15.1±15.9vx	1	11.9±8.7vx	12.7±8.5vx	14.4±10.5vx
2	-	14.6±12.6vx	12.9±10.7vx	2	-	13.2±8.3vx	14.1±9.9vx
3	-	-	14.2±13.0vx	3	-	-	13.3±9.2vx
Overall ave±std	14.8±13.6vx (8.6±7.9μm)			Overall ave±std	13.3±9.2vx (7.7±5.3μm)		

Most importantly, it is extremely difficult for the annotators to visually check the consistency of the relative spatial locations of multiple landmarks. As a result, we found all the warped images produced via the manual landmarking have obvious irregular stretch or inaccurate match in many brain regions (Fig. SN5). In these warped images, accidentally the region around the major neurite tract of interest has less irregular stretch. Therefore, we also traced the neurite tracts from respective warped images and overlaid them on top of each other for comparison (Fig. SN6). Of note, for the same original image (S_a or S_b), in the ideal situation where there were no between-annotator error, the spatial variation of respectively reconstructed neurite tracts of different warped images should be 0. However, in the real manual results, such spatial variation is substantial (Fig. SN6 (a) and (b)). For the two warped images of the same annotator, the average spatial divergence of these neurite tracts is 7.9 μm, which is about twice of the one produced by

BrainAligner (Table SN3).

Based on all these analyses, we conclude that manual alignment has a worse performance than BrainAligner in accuracy, reliability, and speed.

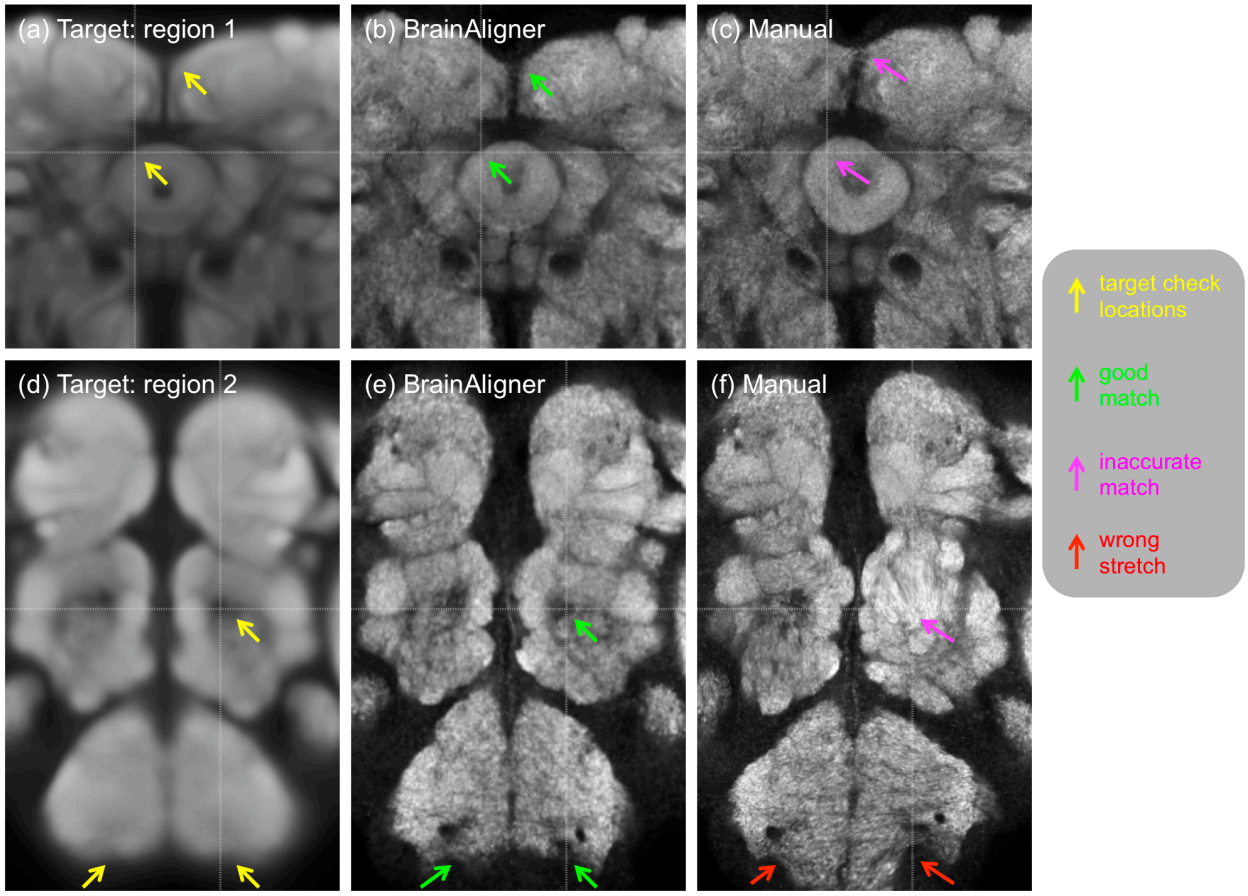


Figure SN5. Comparison of the aligned images produced by BrainAligner and manual landmarking. (a) (d) Two typical brain regions and several locations (yellow arrows) in our standard target image (T_A) where we checked the alignment accuracy. (b) (e) BrainAligner-aligned regions in the subject image S_a. Green arrows indicate the correctly aligned locations. (c) (f) Manually aligned regions in the same subject image (S_a). Magenta and red arrows indicate inaccurate match and wrong stretch, respectively.

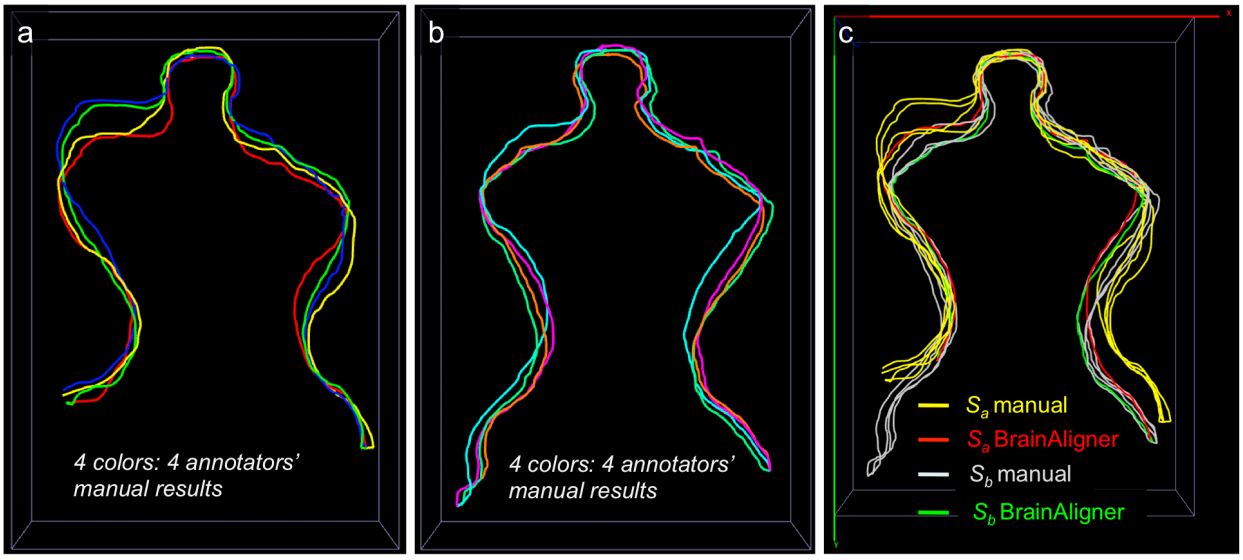


Figure SN6. Comparison of the 3D traced neurite tracts from aligned images produced using both manual alignment and BrainAligner. (a) Overlaid tracts extracted from 4 annotators' manual alignments of the subject image S_a. (b) Overlaid tracts extracted from 4 annotators' manual alignments of the subject image S_b. (c) Neurite tracts extracted from manually aligned brains and BrainAligner aligned brains. The results for both brain images S_a and S_b are shown.

Table SN3. Spatial divergence of the neurite tracts reconstructed from aligned brain-image pairs (S_a versus S_b) produced by each annotator and the BrainAligner.

Spatial divergence	Annotator				BrainAligner
	1	2	3	4	
Voxels	11.9	14.0	13.2	15.5	6.7
Microns	6.9	8.1	7.7	9.0	3.9
Average (microns)	7.9				

Partial brain registration

In some applications, partial brain images may be present. We tested BrainAligner's ability to in aligning the left, center, and right portions of a fruit fly brain. For the subject image S_b used in the previous sub-section, we randomly cropped the three sub-regions from the original complete

image. We ran BrainAligner to produce the local alignments. In the results (Figs. SN 7, 8, and 9) we found that despite of the initially un-matched feature points in the global alignments, BrainAligner is able to produce reasonable alignment for these partial images.

Of note, in some cases the brains may need further pre-processing to make them more globally align-able. One such situation is that the optic lobes may form a flexible articulation to the center brain. We have previously proposed an automatic optic lobe segmentation method (Lam, et al, 2010) to handle such case. Now it is part of our FlyLight project image registration pipeline to align tens of thousands of brain images.

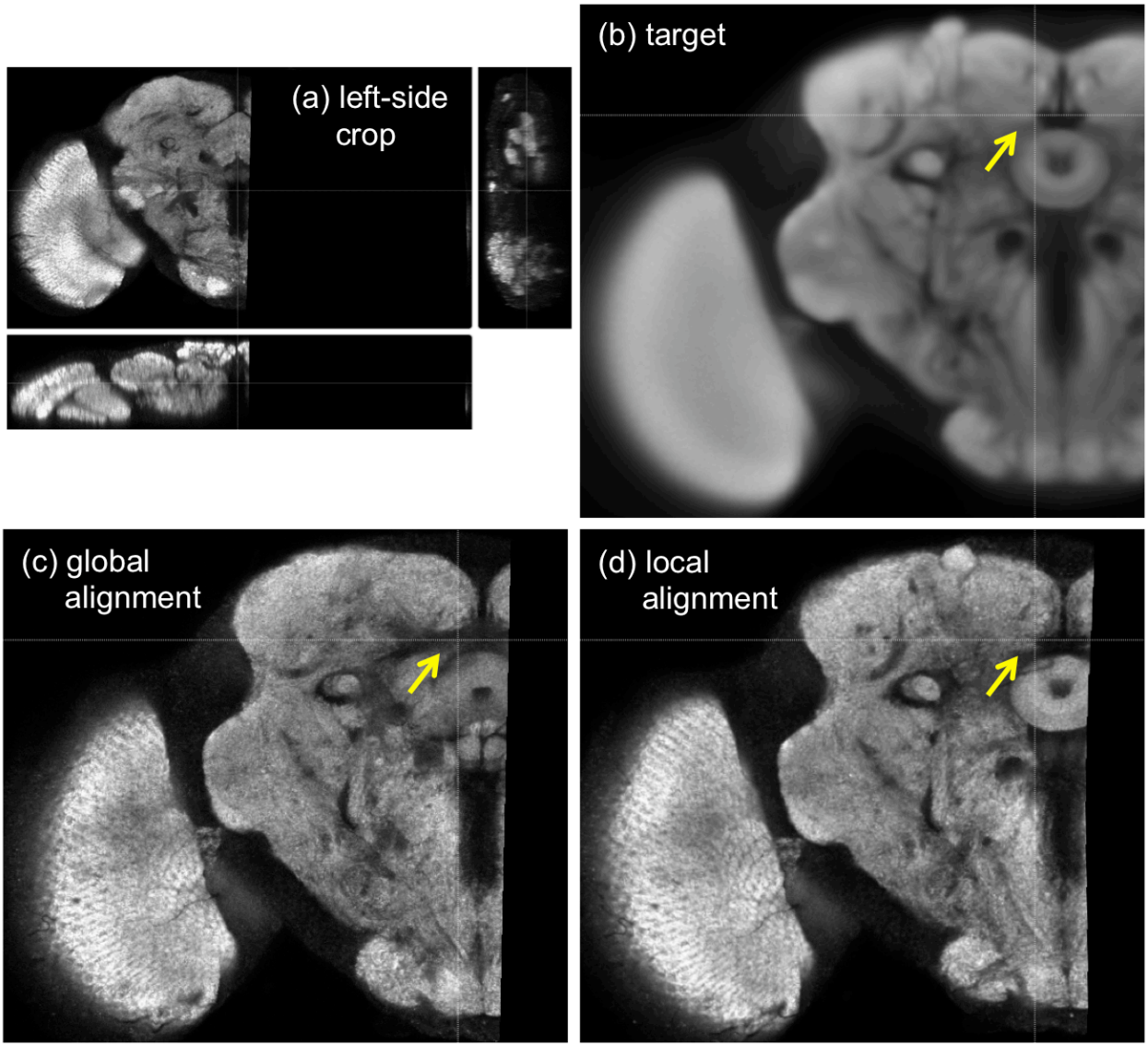


Figure SN7. Aligning the left portion of a brain to our standard target brain. (a) The initial tri-view of the left-side cropped brain image. (b) (c) (d) A check-point at the respective target, global alignment, local alignment brain regions. Yellow arrow: location where there is initial mismatch in the global alignment and correct match in the local alignment.

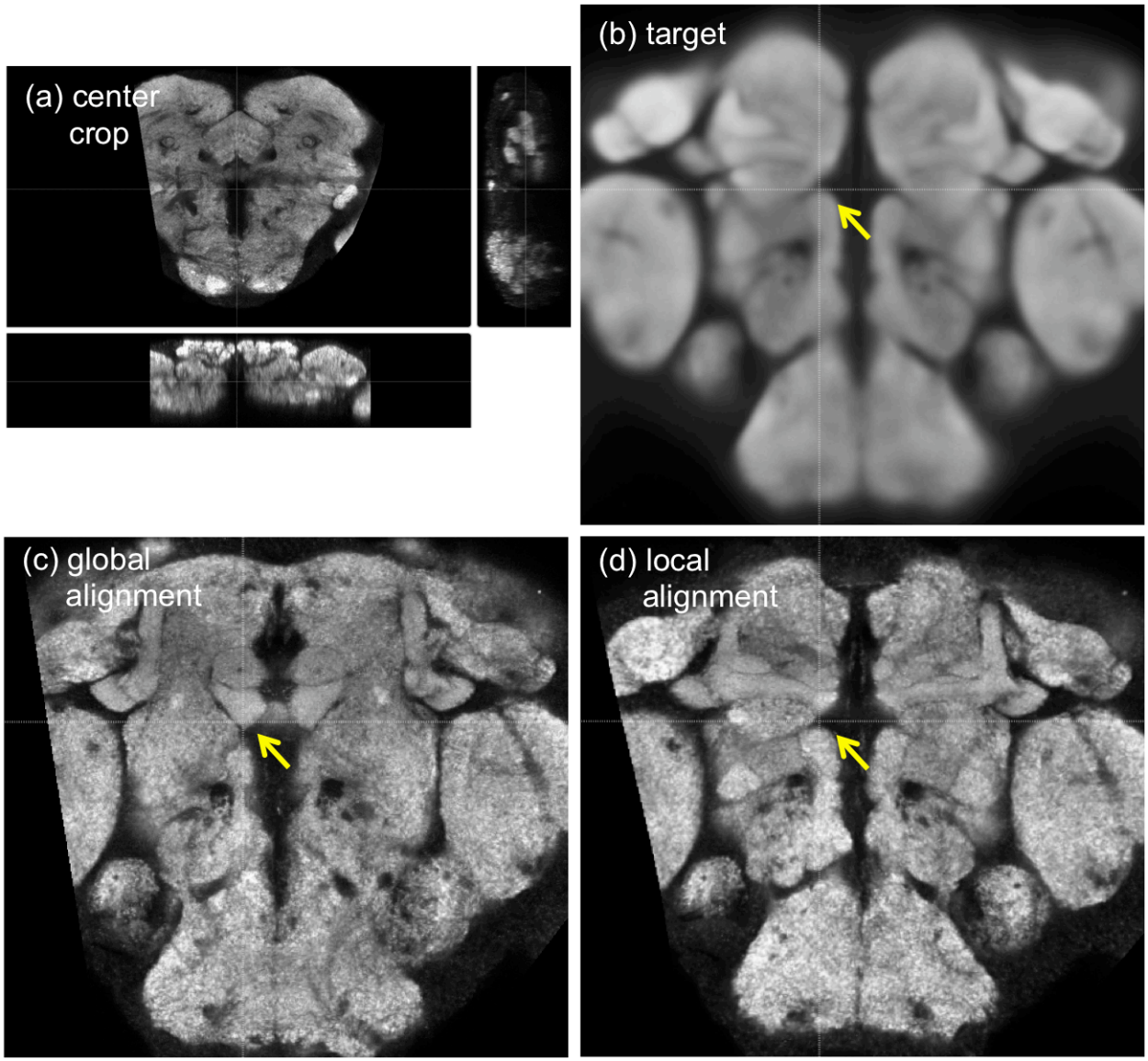


Figure SN8. Aligning the center portion of a brain to our standard target brain. (a) The initial tri-view of the center cropped brain image. (b) (c) (d) A check-point at the respective target, global alignment, local alignment brain regions. Yellow arrow: location where there is initial mismatch in the global alignment and correct match in the local alignment.

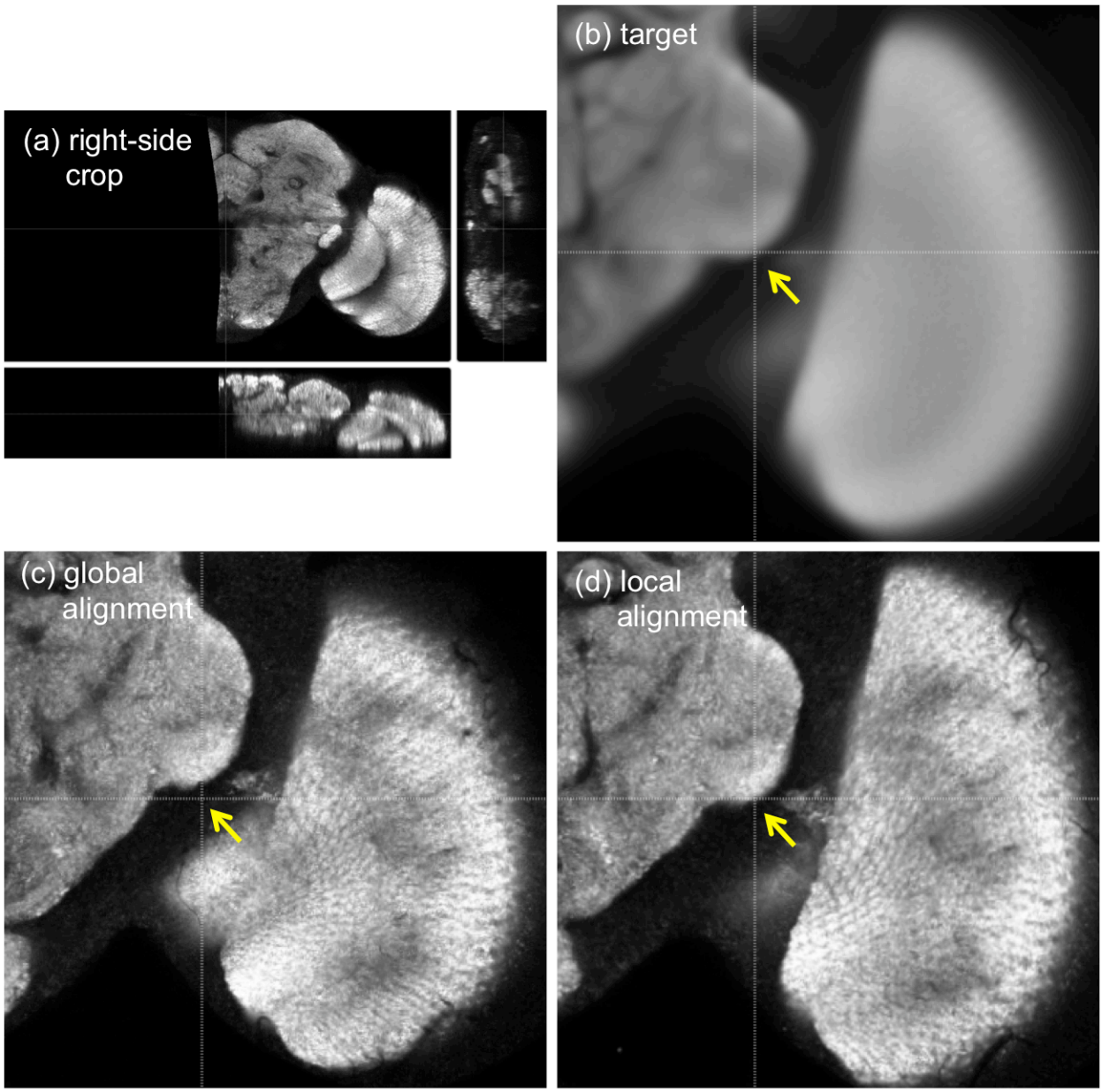


Figure SN9. Aligning the right portion of a brain to our standard target brain. (a) The initial tri-view of the right-side cropped brain image. (b) (c) (d) A check-point at the respective target, global alignment, local alignment brain regions. Yellow arrow: location where there is initial mismatch in the global alignment and correct match in the local alignment.

Discussion and conclusion

Compared to BrainAligner, all automatic and manual methods compared above either have not

produced satisfactory results for our LSM images quickly, and/or do not provide a way to automatically assess the quality of alignment (see the main text), which is critical for large or massive scale alignment applications.

However, it should be noted that many components of these methods, such as mutual information, invariant moments, etc., have been integrated in BrainAligner. We plan to enhance a future version of BrainAligner by using Demons and level set algorithms to refine the registration.

References

1. Jefferis, G.S., Potter, C.J., Chan, A.M., Marin, E.C., Rohlfing, T., Maurer, C.R., Jr., and Luo, L. Comprehensive maps of Drosophila higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell* **128**, 1187-1203 (2007).
2. Qu, L., & Peng, H. Q-FFD for aligning 5D time series of microscopic images. *Manuscript Under Preparation* (2010a).
3. Qu, L., & Peng, H. A principal skeleton algorithm for standardizing confocal images of fruit fly nervous systems. *Bioinformatics* **26**, 1091-1097 (2010b).
4. Kroon, D.J. K-FFD software (http://www.mathworks.com/matlabcentral/fileexchange/20057-b-spline-grid-image-and-point-based-registration)
5. Thirion, J.P. Image matching as a diffusion process: an analogy with Maxwell's demons, *Medical Image Analysis* **2**, 243-260 (1998).
6. Shen, D. & Davatzikos, C. HAMMER: hierarchical attribute matching mechanism for elastic

registration, *IEEE Transactions on Medical Imaging* **21**, 1421-1439 (2002).

7. Vercauteren, T., Pennec, X. Perchant, A. & Ayache, N. Symmetric log-domain diffeomorphic registration: a demons-based approach. *Lecture Notes in Computer Science* **5241**, 754–761 (2008).
8. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. & Suetens, P. "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, **16**: 187-198 (1997).
9. Rohlfing, T., Brandt, R., Maurer, C.R.J., & Menzel, R. Bee brains, B splines and computational democracy: generating an average shape atlas, *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, 187-194 (2001).
10. Cachero, S., Ostrovsky, A.D., Yu, J.Y., Dickson, B.J. & Jefferis, G.S. Sexual dimorphism in the fly brain, *Current Biology* **20**, 1589-1601 (2010).
11. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O. & D. J. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, *IEEE Transactions on Medical Imaging* **18**, 712-721 (1999).
12. Periaswamy, S. & Farid, H., Medical image registration with partial data, *Medical Image Analysis* **10**, 452-464 (2006).
13. Rohlfing, T. & Maurer, C.R., Nonrigid image registration in shared memory multiprocess environments with application to brains, breasts, and bees, *IEEE Transactions on Information Technology in Biomedicine* **7**, 16-25 (2003).
14. Cachero, S., Ostrovsky, A.D., Yu, J.Y., Dickson, B.J. & Jefferis, G.S.X.E., Sexual dimorphism in the fly brain, *Current Biology* **20**, 1589-1601 (2010).

15. Peng, H., Ruan, Z. Long, F., Simpson, J.H. & Myers E.W. V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nature Biotechnology* **28**, 348-353 (2010).

16. Lam, S.C., Ruan, Z., Zhao, T., Long, F., Jenett, A., Simpson, J., Myers, E.W. & Peng, H. (2010). Segmentation of center brains and optic lobes in 3D confocal images of adult fruit fly brains. *Methods* **50**, 63-69 (2010).