High precision registration between zebrafish brain atlases  
using symmetric diffeomorphic normalization

Gregory D. Marquart¹,², Kathryn M. Tabor¹, Eric J. Horstick¹, Mary Brown¹,  
Alexandra K. Geoca¹, Nicholas F. Polys³, Damian Dalle Nogare¹ and Harold A. Burgess¹*

¹ Division of Developmental Biology, Eunice Kennedy Shriver National Institute of Child Health  
and Human Development, Bethesda, MD 20892  
² Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD  
20742  
³ Advanced Research Computing, Department of Computer Science, Virginia Polytechnic  
Institute and State University, Blacksburg, VA 24061

* Contact: burgessha@mail.nih.gov ; 301-402-6018

Running title: High fidelity inter-atlas registration using ANTs SyN transform  
Keywords: zebrafish, atlas, ANTs, SyN, registration, diffeomorphism, normalization, brain  
imaging, transgenic, virtual reality, Unity, X3D
Abstract

Atlases provide a framework for spatially-mapping information from diverse sources into a common reference space. Specifically, brain atlases allow annotation of gene expression, cell morphology, connectivity, and activity. In larval zebrafish, advances in genetics, imaging, and computational methods now allow the collection of such information brain-wide. However, due to technical considerations, disparate datasets may use different references and may not be aligned to the same coordinate space. Two recent larval zebrafish atlases exemplify this problem: Z-Brain, containing gene expression, neural activity and neuroanatomical segmentations, was acquired using immunohistochemical stains, while the Zebrafish Brain Browser (ZBB) was constructed from live scans of fluorescent reporters in transgenic larvae. Although different references were used, the atlases included several common transgenic patterns that provide potential 'bridges' for transforming each into the other's coordinate space. We tested multiple bridging channels and registration algorithms and found that the symmetric diffeomorphic normalization (SyN) algorithm improved live brain registration precision while better preserving cell morphology than B-spline based registrations. SyN also corrected for tissue distortion introduced during fixation. Multi-reference channel optimization provided a transformation that enabled Z-Brain and ZBB to be co-aligned with precision of approximately a single cell diameter and minimal perturbation of cell and tissue morphology. Finally, we developed software to visualize brain regions in 3-dimensions, including a virtual reality neuroanatomy explorer. This study demonstrates the feasibility of integrating whole brain datasets, despite disparate reference templates and acquisition protocols, when sufficient information is present for bridging. Increased accuracy and interoperability of zebrafish digital brain atlases will facilitate neurobiological studies.

Background

Larval stage zebrafish are an increasingly popular model for neurobiological studies. With a brain that contains an estimated $10^5$ neurons, larvae are similar in complexity to adult *Drosophila*, another established neuroscience model. In both systems, researchers can deploy a wide range of genetic tools in efforts to decode patterns of neural structure and connectivity. In larval zebrafish, optical transparency and constrained physical dimensions (fitting within an imaging volume of 1000 $\times$ 600 $\times$ 350 $\mu$m) allow the entire brain to be rapidly scanned at cellular resolution using diffraction-limited microscopy. In principle, this enables researchers to systematically analyze effects of manipulations on a brain-wide level. However, such efforts have been hampered by the absence of a comprehensive digital atlas that would provide researchers with a unified framework in which to aggregate data from different experiments and gain deeper insights from correlations between neuronal cell identity, connectivity, gene
expression and function within the brain. Additionally, digital atlases may more clearly delineate structural boundaries that are difficult to accurately identify within individual brains, allowing for a more rigorous mapping of neuroanatomical regions onto experimental data.

These longstanding problems in zebrafish neuroscience have recently been addressed by the construction of digital atlases using 3-dimensional (3D) image registration techniques: the Virtual Brain Explorer for Zebrafish (ViBE-Z), Z-Brain and the Zebrafish Brain Browser (ZBB) [1–3]. In these atlases, information on gene expression, structure (neuronal cell bodies, glia, vasculature, ventricles, neuropil or axon tracts) and measures of activity (calcium or secondary messenger activity) are consolidated within a common spatial framework. By using widely-available transgenic lines or immunohistochemical stains as reference templates for brain alignment, each of these atlases provides other researchers the opportunity to register their own datasets into these digital spaces and take advantage of the information contained within.

ViBE-Z was the first comprehensive 3D digital brain atlas in zebrafish that used a nuclear stain for the alignment of 85 high resolution scans comprising 17 immunohistochemical patterns at 2-4 days post-fertilization (dpf) [3,4]. In ViBE-Z, custom algorithms were developed to correct for variations in fluorescent intensity with scan depth, and a landmark approach taken to perform accurate image registration and segmentation into 73 neuroanatomic regions.

In contrast, two more recent approaches (Z-Brain and ZBB) have generated brain atlases at 6 dpf through non-linear B-spline registration using the freely available Computational Morphometry Toolkit (CMTK) [5,6]. Z-Brain includes 29 immunohistochemical patterns from 899 scans which form the basis for expert manual segmentation of the brain into 294 neuroanatomic regions. These partitions facilitate the analysis of phospho-ERK expression for mapping neural activity [2]. In Z-Brain, each expression pattern was co-scanned with tERK immunoreactivity, and registered to a single tERK-stained reference brain. For ZBB, we live-imaged 354 brains from 109 transgenic lines and manually annotated the expression found in each [1]. In place of tERK, a single vglut2a:dsRed transgenic brain was used as the reference in ZBB with transgenic lines crossed and co-imaged with this channel for registration. Brain browser software enables researchers to select a transgenic line labeling a selected set of neurons for monitoring and manipulating circuit function.

While Z-Brain and ZBB are powerful datasets on their own, we saw an opportunity to merge the two atlases because they are both based on confocal scans of 6 dpf larvae. This would bring to Z-Brain a large number of additional transgenic lines and to ZBB, the expert manual segmentation of Z-Brain. Several
similarities between Z-Brain and ZBB suggested that bridging the atlases would be possible. First as zebrafish rearing conditions are standardized across laboratories and fish were imaged at the same time post-fertilization, Z-Brain and ZBB likely reflect the same developmental timepoint. Second, images in both atlases were acquired at similar resolution (0.8 x 0.8 x 2 μm for Z-Brain; 1 x 1 x 1 or 1 x 1 x 2 μm for ZBB) and orientation (dorsal to ventral horizontal scans). Third, despite using distinct templates (tERK for Z-Brain and vglut2a for ZBB), Z-Brain and ZBB have several transgenic markers in common, which provide the possibility of bridging the datasets by using these shared patterns as references for a secondary registration step.

One of the strengths of larval zebrafish is the ability to rapidly image at cellular resolution and visualize brain-wide neuronal morphology, providing valuable information on cell type and potential connectivity. Z-Brain and ZBB both illustrate the feasibility of performing whole-brain registration with precision sufficient to ensure that the ‘same’ neurons from different fish are aligned to within a cell diameter (∼8 μm). However, a challenge for brain registration in zebrafish is to minimize local distortions, so that cellular morphology is preserved while still allowing sufficient deformation to overcome biological variability between individual brains or malformations due to tissue processing.

Here we describe a method to co-register ZBB and Z-Brain, bridging the two existing 6 dpf larval zebrafish brain atlases. By using the diffeomorphic algorithm SyN in the Advanced Normalization Tools (ANTs) software package [7,8], we were able to overcome differences in tissue shape due to fixation, optimize the trade-off between preservation of cell morphology and global alignment, and provide precise registration in all tested brain regions. Additionally, ANTs provided superior image registration for live-scanned larvae, enabling us to improve the precision of registration and neuron morphology within our original ZBB atlas, allowing us to compile a new version with increased fidelity (ZBB0.2).

Methods

Zebrafish lines

In order to provide additional options for bridging ZBB and Z-Brain, we scanned two transgenic lines that were not in the original ZBB release: Et(gata2a:EGFP)zf81 (vmat2:GFP) and Tg(isl1:GFP)rw0 (isl1:GFP) [9,10]. Other lines referenced in this study are Tg(slc6a3:EGFP)ot80 (DAT:GFP) [11], Tg(-3.2fev:EGFP)ne0214 (pet1:GFP) [12], y264Et [13], s1181Et [14], Tg(gad1b:GFP)nns25 (gad1b:GFP) [15], Tg(slc6a5:GFP)cf3 (glyT2:GFP) [16], Tg(-17.6isl2b:GFP)zc7 (isl2b:GFP) [17], Tg(-3.4tph2:Gal4ff)y228 (tph2:Gal4) [18], TgBAC(slc17a6b:lox-DsRed-lox-GFP)nns14 (vglut2a:DsRed)
[19], Tg(slc17a6:EGFP)zf139 [20], Tg(elavl3:CaMPARI(W391F+V3987L))jf9 [21],
Tg(phox2b:GFP)w37 [22], JI229aGt [23] and several Gal4 enhancer traps from ZBB: y304Et, y332Et,
y341Et, y351Et and y393Et [1]. All in vivo experimental protocols were approved by the NICHD animal
care and use committee.

Immunohistochemistry

Immunolabeling was as described [2] with the following adaptations. Larvae were fixed overnight at 4°C
in PBS with 4% paraformaldehyde and 0.25% Triton X-100. Samples were then washed in PBS
containing 0.1% Triton X-100 (PBT) 3 times for 5 min. For antigen retrieval, samples were incubated in
150 mM Tris-HCl ph 9.0 for 5 min at room temperature (RT), followed by 15 min at 70°C and washed in
PBT 2 times for 5 min at RT [24]. Critically, samples were then permeabilized on ice in fresh 0.05%
trypsin-EDTA for no more than 5 minutes. If pigmented, samples were incubated in PBT with 1.5% H2O2
and 50 mM KOH for 15 min, rinsed 2 times in PBT and washed again for 10 min, all at RT. Samples
were then blocked in PBT containing 5% normal goat serum (NGS) and 0.2% bovine serum albumin
(BSA) for 1 hr at RT before incubation at 4°C with tERK antibodies (Cell Signaling, 4696) diluted 1:500
in PBT with 5% NGS and 0.2% BSA for a minimum of 6 hr. Samples were then washed with PBT 4
times for 30 min at RT before incubation at 4°C for a minimum of 2 hr with fluorescent secondary
antibodies (Alexa Fluor 488 or 548) diluted 1:1000 in PBT with 5% NGS and 0.2% BSA. Samples were
finally rinsed 4 times for 30 min at RT prior to imaging.

Registration

Registrations were performed using CMTK version 3.2.3 and ANTs version 2.1.0 running on the National
Institute of Health’s Biowulf Linux computing cluster. Registrations were parallelized using Slurm-based
bash scripts available upon request. For CMTK, previously optimized registration parameters that
minimize computation time while maximizing precision were used (affine parameters: registrationx --dofs
12 --min-stepsize 1 ; elastic parameters: warpx --fast --grid-spacing 100 --smoothness-constraint-weight
1e-1 --grid-refine 2 --min-stepsize 0.25 --adaptive-fix-thresh 0.25). For ANTs registrations, the
parameters used are cited in the relevant text and figures with optimized parameters listed in Table 1. All
deformable transformations are initiated with a rigid and affine step (parameters included in Table 1).
Aside from the use of ANTs, the basic imaging and registration workflow was performed as previously
described [1]. Image volumes were rendered within the Zebrafish Brain Browser (ZBB), ImageJ [25] or
code written in IDL (Harris Geospatial Solutions). For the conversion to/from NIfTi format required for
ANTS, we used the ImageJ plugin nifti_io.jar written by Guy Williams [26].
Volume rendering & 3D visualization

Binary masks corresponding to 25 anatomical regions from Z-Brain aligned to ZBB were converted into meshes using the Create Surfaces tool in the IntSeg_3D.jar plugin for ImageJ [27]. Edges for individual meshes were iteratively reduced below 5000 and vertices (single-precision floating-points of the triangular meshes) written as OBJ files. As there is no intrinsic color or color conventions as of yet for these brain structures, we used color hue as a nominal categorical coding for each region. To maximize accessibility, we rendered meshes in Extensible 3D (X3D) format, an ISO (International Organization for Standardization) standard developed by the not-for-profit Web3D Consortium [28]. This format allows portability between numerous tools and applications as well as deployment across a broad spectrum of platforms. For the rendering, OBJs were transcoded into ImageTextureAtlas PNGs using X3D’s standard IndexedFaceSet to represent mesh information and then tiled at different resolutions (4096 & 8192 pixels squared) using AtlasConversionScripts [29]. Additionally, dask and pyimg python libraries were used to generate volume norms (image and ImageTextureAtlas files) by gradient descent. All renderings were then merged into a single X3D XML scene which was losslessly compressed (in SRC/glTF) to a final size of 4.5 MB. This makes the scene compact enough to be visualized on a cell phone, while retaining details for visualization and editing in immersive virtual reality environments. Finally, X3D files were published to HTML5 via the X3DOM library and a simple user interface created that allows for the visibility of different structures to be toggled on and off. Brain meshes were converted to FBX files for import into Unity using Blender 2.78a (Blender foundation, Amsterdam, NL) and mobile app development for Google Cardboard VR headsets performed in Unity 5.4.2 (Unity Technologies SF, San Francisco CA) using the Google VR for Unity SDK (Google, Mountain View CA). Custom scripts controlling movement and mesh display were written for Unity in C#.

Measurements

Mean Landmark Distance (MLD). To assess registration precision using MLDs, corresponding landmarks were located and annotated on the reference brain, and on unregistered brains. In each case, landmarks were chosen to be widely distributed within the brain, and readily recognized in corresponding brain scans. In addition, to verify recognizability, the vglut2a landmarks in the reference brain were located by 3 blinded scorers; mean distance from each of the 10 reference points ranged from 1.7 to 11.8 μm (mean, 4.5 ± 0.9 μm). Using ImageJ, we positioned a 3 micron cube centered on each landmark in a second channel for each brain scan, then, after registering the brain scan using the first channel, applied the resulting transformation matrix to the second channel, using Nearest Neighbor interpolation for both
CMTK and ANTs. Landmark distance was taken as the distance between the geometrical center of the corresponding cubes in the reference image, and in the registered image.

**Hausdorff distance.** We manually segmented cells in a vglut:DsRed brain scan in a second channel and applied transformation matrices for registration to this second channel. Segmented cells were broadly drawn within the boundaries of the soma. We then compared the morphology of cells after registration (A), to their original shape (B) by calculating the partial Hausdorff distance [30]. Briefly, for every point in a segmented cell mask before registration, we found the minimum distance to a point in the same mask after registration. The Hausdorff distance is the maximum of all such distances, calculated for both A→B and B→A. Because the Hausdorff distance is highly sensitive to cell alignment, and registration displaces cells from their original location, we found the optimal alignment for comparisons using a two-step process. First we aligned the geometric center of each cell in the original and transformed images. Second we searched for the minimal Hausdorff distance across 4940 rigid transformations of the aligned cell, within a 3 x 3 x 3 micron cube, (0.25 micron steps in each dimension). Finally, as Hausdorff distances are sensitive to outliers, we used the 95th percentile distances instead of the maximum Hausdorff distance for all measures [30].

**Cell volume.** For each segmented cell, we calculated its change in volume as the absolute value of the fractional change in the number of pixels after application of a transformation matrix.

**Elongation index.** For each pixel in a segmented cell, we found the maximal distance (MD) to any other pixel in the mask. The elongation index for a given cell was the 95th percentile largest value of MD, which we take as an approximation of the diameter of the cell across its longest axis.

**Cross correlation.** Cross correlation between the tERK-stained reference brain, and registered tERK stains, was performed using the c_correlate function within IDL version 7.0. Correlations were run within eighteen 50 µm-side cube sub-regions of the image volumes that were manually selected to encompass high contrast boundaries and the mean of the 18 values taken as the mean cross correlation (MCC) for each brain in Fig. 3.

**Jaccard index.** Anti-tERK immunohistochemistry intensely stains tectal neuropil. Thus for measuring the accuracy of registration of the tectal neuropil, we manually segmented the left tectal neuropil area in 6 confocal scans of tERK stained larvae and our reference brain. We applied transformation matrices to these masks, then calculated the Jaccard index as the volume of the intersection between each registered mask (A) and the reference brain (B), divided by the total volume of the union of A and B.

**Results**
Optimization of ANTs based registration of live vglut2a:DsRed image scans

Brain registration in Z-Brain and ZBB used the B-spline elastic transformation in CMTK. Before attempting to co-align Z-Brain and ZBB, we tested an alternate algorithm for brain alignment, the diffeomorphic symmetric normalization (SyN) method in ANTs, because: (1) SyN has been shown to outperform B-spline transformations for deformable image registration in a variety of imaging modalities [31,32]. (2) ANTs permits registration using multiple reference channels, potentially allowing the use of multiple complementary expression patterns as references for improved registration fidelity. (3) By calculating forward and reverse transformations simultaneously, SyN transformation matrices are intrinsically symmetric, ensuring that bridging registrations would be unbiased and that we could easily perform reciprocal transformations to register each dataset into the other's coordinate system.

To calibrate registration parameters, we assessed the alignment precision and distortion of cell morphology after the registration of six representative vglut2a:DsRed scans to the original vglut2a:DsRed reference brain in ZBB (vglut2aZBB; file vglut-dsred-ref-01.nii.gz, available from [33], procedure summarized in Fig. 1a). Similar to CMTK we employed a three step registration within ANTs where rigid and affine steps were used to initialize a deformable registration using the SyN diffeomorphic transformation with cross correlation (CC) as the similarity metric. We tested a range of values for each of the SyN parameters as well as the radius of the region used for cross correlation.

To measure registration precision, we visually located 10 point-based landmarks in the vglut2aZBB pattern (Additional File 1a-c) and in each of the 6 vglut2a:DsRed confocal scans. We then used the vglut:DsRed channel for registration, and applied the resulting transformation matrix to the landmarks in each of the 6 brains. We measured the distance of each landmark from its location in the vglut2aZBB reference brain (Additional File 1a,d). We designated the average of the 10 distances the 'Mean Landmark Distance' (MLD). To assess the amount of distortion in cell shapes produced by the parameter sets, we segmented 107 cells in an unregistered vglut2a:DsRed confocal scan (Additional File 2), and applied each transformation matrix to this set of cell masks. Changes in cell shape were measured using the partial Hausdorff distance for each cell after registration compared to its original shape (see Methods).

Next we plotted the MLD against the Hausdorff distance and located points along the Pareto frontier (Fig. 1b) of these two measures. These points represent potentially optimal transformations, where registration accuracy can only be improved by increasing distortion, or vice versa. To distinguish between these points, we examined two additional measures of distortion: the change in cell volume (Fig. 1c) and
maximal elongation (Fig. 1d). Three transformations showed statistically significantly reduced distortion compared to CMTK for both measures, and we selected the one (Fig. 1b, point d) with the greatest precision for further testing. With this set of parameters (see Table 1, live registration), mean registration error was within the diameter of a single neuron for both ANTs and CMTK (MLD for ANTs 6.7 ± 0.3 μm, for CMTK 7.6 ± 0.4 μm; N = 6 brains, paired t-test p=0.056). However, cell morphology was better preserved using ANTs (Hausdorff Distance for ANTs 2.30 ± 0.14, CMTK 2.37 ± 0.14; N = 107 cells, paired t-test p=0.013), especially within ventral structures such as the hypothalamus and the caudal medulla oblongata (Fig. 1e).

We next examined whether these registration parameters also improved precision for the co-aligned transgenic lines. For ZBB, we co-scanned transgene and enhancer trap expression patterns with the vglut2a:dsRed transgene, allowing us to register each expression pattern to vglut2aZBB. We first compared the overlap and morphology of the Mauthner cells from brain scans of three different individuals of transgenic line J1229aGt [23]. Overlap of Mauthner cell bodies was similar for CMTK and ANTs (Fig. 1f,g). However, in CMTK registered images, the Mauthner axon was distorted in the caudal medulla, whereas axon morphology was preserved with ANTs. Second, in our previous work, we assessed the precision of CMTK registration using line y339Et by independently scanning two sets of three larvae, producing an average for each set, and visually comparing the result. With CMTK we had noted misalignment of approximately 1 cell diameter in the neuropil of the optic tectum (Fig. 1h). This was substantially improved with ANTs, where there was much closer alignment of the two averages (Fig. 1i).

**Improved precision of ZBB after registration using ANTs**

We next recompiled ZBB using ANTs to register the entire set of 354 brain scans from 109 different transgenic lines that were part of ZBB, then as before, averaged multiple larvae to create a representation of each transgenic line, masked the average stacks to remove expression outside the brain and re-imported the resulting images into our Brain Browser software. We refer to this new recompilation of the atlas as ZBB1.2. Unprocessed and registered brain images are available online [35].

To determine whether ZBB1.2 was a quantitative improvement over ZBB, we identified two conspicuously labeled cells or landmarks in each of 12 transgenic lines from the atlas (Additional File 3). We marked these positions in each of the three brain scans for each line, then, after registration, calculated the distance between corresponding points in each pair of brains. The mean of these distances measures how precisely landmarks are registered across the three brains. We performed this procedure...
first for brains registered using CMTK, then for the same set of brains registered using ANTs, allowing us to compare MLDs for the two methods (Fig. 2a-b). Overall, landmark distances decreased from ZBB to ZBB\(_2\) (10.8 ± 1.02 \(\mu\)m to 8.1 ± 0.83 \(\mu\)m; \(N = 24\) landmarks, paired t-test \(p=0.008\)), indicating that ZBB\(_2\) has significantly improved precision, and confirming that the new atlas is accurate to approximately the diameter of a single neuron. The improvement was greatest deeper in the brain (Fig. 2c; linear regression, \(N=24\), \(p=0.003\)) with the largest improvement for the caudal hypothalamus in line \(y341\), where increased alignment precision was associated with noticeably reduced distortion between the three brain scans (Fig. 2d).

Additionally, we inspected regions of ZBB\(_2\), where we had noticed poor registration precision or pronounced cell distortion in the original ZBB. One such area was the dorsal thalamus, where cell morphology was noticeably perturbed after elastic registration with CMTK, with cell somas stretching across the midline (Fig. 2e). In ZBB\(_2\) cells retained a rounded morphology with distinct cell clusters on the left and right sides of the brain (Fig. 2f). Similarly, distortions in cell shape that were apparent in the caudal hypothalamus in ZBB, were absent in ZBB\(_2\) (Fig. 2g,h). In the caudolateral medulla, we previously obtained poor registration, with expression extending to regions outside the neural tube (Fig. 2i). In ZBB\(_2\), patterns had improved bilateral symmetry and were correctly confined to the neural tube (Fig. 2j). Finally, we noticed that the posterior commissure was poorly aligned between larvae leading to a defasciculated appearance in ZBB (Fig. 2k), whereas this tract had the expected tightly bundled appearance in ZBB\(_2\) (Fig. 2l). Together, these observations confirm that ZBB\(_2\) is a more faithful representation of the transgenic lines. Not only is cell morphology better preserved, but global registration precision is improved compared to the original ZBB atlas.

**Optimization of ANTs registration parameters for fixed tissue**

The Z-Brain atlas was derived by registering brain scans to a single brain that was fixed, permeabilized and immunostained for tERK expression. We therefore anticipated that tERK would be a useful channel for bridging the two atlases, if we could first successfully register a tERK stained vglut2a:DsRed expressing brain to ZBB\(_2\). Therefore, we fixed and stained a transgenic vglut2a:DsRed larva for tERK, and registered the tERK pattern to ZBB\(_2\) using the vglut2a pattern. We used the resulting image as our ZBB tERK reference brain (tERK\(_{\text{ZBB}}\); file terk-ref-02.nii.gz available from [33]).

In addition to the tERK reference brain, Z-Brain contains an average tERK representation from 197 tERK stained larvae, which we thought might serve as a bridge between atlases. During studies on pERK-based
We therefore varied the registration parameters that were optimal for live vglut2a registration, to find settings that best rectified the variable tissue morphology following fixation and permeabilization (process summarized in Fig. 3a). For optimization of fixed tissue registration, we used a set of 6 tERK stained brains (including the Z-Brain tERK reference), iteratively varied parameters for registration to tERKZBB and assessed registration fidelity. For measuring precision, we were not able to identify unambiguous landmarks within the optic tectum, so we instead calculated the cross-correlation between each of the aligned tERK stains and tERKZBB within small volumes, including parts of the tectum (Fig. 3f,g). To verify that the 'fixed brain' parameters that yielded the greatest cross correlation did in fact improve registration within the tectum, we manually segmented the tectal neuropil in the same 6 brains, applied the transformation matrix to each mask, and calculated the Jaccard index for overlap with the segmented neuropil in tERKZBB. Parameters for fixed brain registration produced a significant increase in overlap, compared to the live brain parameters (Fig. 3h,i) and visual inspection confirmed that the morphology of the optic tectum neuropil after registration was greatly improved (Fig. 3j,k). We therefore used ANT's with the fixed brain parameters (Table 1, fixed registration) to register our 167 tERK stained brains to tERKZBB, and generated an average tERK representation comparable to the 197 tERK average in Z-Brain (Fig. 3l,m).

**Inter-atlas registration using multi-channel diffeomorphic transformation**

Z-Brain and ZBB incorporated eight expression patterns that we judged sufficiently similar to act either as templates for bridging the datasets and/or to provide metrics for assessing the precision of a bridging
registration (Table 2, Additional File 4). For example, vglut2aZBB is a confocal scan of DsRed in a single larva from transgenic line TgBAC(slc17a6b:loxP-DsRed-loxP-GFP)ms14, whereas Z-Brain includes Tg(slc17a6:EGFP)zf139. In both cases, reporter expression is regulated by the same bacterial artificial chromosome [15,20]. Crossing these two lines allowed us to scan DsRed and EGFP in the same larva and confirm that the patterns were largely congruous, potentially allowing us to use vglut2a expression to bridge the two atlases. Likewise, the expression patterns of tERK, elavl3, isl2b, vmat2 in Z-Brain and ZBB appeared sufficiently similar to provide templates for atlas co-registration.

Taking advantage of the ability of ANTs to use of multiple reference channels concurrently, we compared the effect of combinatorial use of complementary reference channels for inter-atlas registration (process summarized in Fig. 4a). We used seven expression patterns to evaluate registration precision: vglut2a, isl2b, vmat2, tERK, isl1, gad1b and glyT2. For each pattern we identified a set of 4-10 point-based landmarks that could be identified in corresponding ZBB and Z-Brain images and that were widely distributed to represent diverse brain regions (total of 41 landmarks; Additional File 5). We marked these points in each set of images, registered Z-Brain images to ZBB1.2 images, measured the distance between cognate landmarks and calculated the mean landmark distance for each of the seven expression patterns.

We used two summary measures of registration precision. The first metric (M1) was the mean of MLDs for the three patterns that were not used to drive registration (isl1, gad1b and glyT2). Although these channels measure precision independent of the patterns for atlas registration, they are relatively sparse and do not assess precision across the whole brain. Thus, to provide a global measure of precision, we also used a second metric (M2) that was the mean of all seven MLDs: those in M1 plus four of the patterns used as references for registration — vglut2a, tERK, isl2b and vmat2.

Using CMTK, minimal M1 and M2 scores were obtained using the average vmat2 pattern as the reference (Fig. 4b; mean MLD for 41 landmarks 14.9 ± 1.3 μm). We therefore registered all images in Z-Brain to ZBB using the vmat2 average in each dataset as the reference channel. We observed severe tissue distortions in several brain regions, with noticeable flattening of the torus longitudinalis as well as gross tissue distortions, particularly in ventral brain regions (Fig. 4c,d; ZBrain-CMTK). Next we used the ANTs SyN algorithm to register the atlases. Ideally, patterns for registration should include information throughout the brain. Because ANTs can use multiple concurrent reference channels to derive an optimal transformation matrix, we speculated that the best possible transformation would be achieved by a combination of channels with complementary information. We therefore produced an inter-atlas transformation matrix for every combination of the elavl3, isl2b, vglut2aAV (vglut2a average brain), vmat2, tERKZBB (tERK single brain) and tERKAV (tERK average brain) patterns as references. Because Z-
Brain used fixed samples, we used the registration parameters optimized for the greater variability present in fixed tissue. Multi-channel registration significantly reduced $M_1$ and $M_2$ values compared to any single channel alone and to transformations obtained using CMTK. The registration obtained with $vglut2a$, tERKZBB, $vmat2$ and $isl2b$ gave the lowest global metric ($M_2$) value and an $M_1$ score within 10% of the lowest scoring combination (Fig. 4b). With these parameters, the MLD was $9.1 \pm 0.8 \mu m$ (N=41)

landmarks) and the overt tissue distortions noted after elastic registration were far less salient (Fig. 4c,d; ZBrain-SyN). We therefore applied the transformation matrix obtained with this set of channels to the database of gene expression patterns in Z-Brain to align them to ZBB$_{1.2}$, and used the inverse of the transformation generated by SyN to register ZBB$_{1.2}$ to the Z-Brain coordinate system. We imported all Z-Brain expression patterns not previously represented in the database into ZBB$_{1.2}$, producing a total of 133 expression patterns.

The accuracy of the inter-atlas registration is evident when comparing the location of cells that are present in both datasets, such as those labeled by petl:GFP. The Z-Brain transformed pattern closely matches the transgene expression pattern in ZBB$_{1.2}$ within the superior raphe (Fig. 4e — note however that unexpectedly, the line in ZBB$_{1.2}$ also labels a set of more rostral cells not apparent in Z-Brain). Both atlases also include lines labeling the Mauthner cells. After registration, Mauthner cells in the atlases substantially overlapped, although they were several microns more medially positioned in ZBB$_{1.2}$ (Fig. 4f). Expression in the $DAT:GFP$ line in ZBB$_{1.2}$ overlapped well with the tyrosine hydroxylase stain from Z-Brain in the pretectum (Fig. 4g), although again, the ZBB$_{1.2}$ pattern was slightly more medial than in Z-Brain. Caudally, the $glyT2:GFP$ transgenic line labels glycinergic neurons in longitudinal columns in the medulla oblongata [36]. These columns were closely aligned after ZBB$_{1.2}$ was registered to Z-Brain (Fig. 4h).

Although best practice is to align directly to either ZBB or Z-Brain, because many researchers will have already registered data sets to either ZBB or Z-Brain, or for cases where it may not be possible to directly register a dataset, we have provided transformation matrixes and detailed instructions to quickly re-align datasets to either of the coordinate systems ([37] ; Additional File 6).

**Neuroanatomical visualization**

Z-Brain includes 294 masks that represent anatomically defined brain regions or discrete clusters of cells present in transgenic lines. We selected 113 of these masks that delineate neuroanatomical regions and transformed them into the ZBB$_{1.2}$ coordinate system. We had previously defined a small number of our
own anatomical masks by thresholding clusters of neuronal cell bodies located in well-defined brain regions. However the Z-Brain masks are more comprehensive, have smoother boundaries and include both the cell bodies and neuropil for a given region (Fig. 4i-l). We therefore imported the Z-Brain masks into ZBB1.2, replacing most of our existing masks. We also modified the Brain Browser software to automatically report the neuroanatomical identity of a selected pixel, or to display the boundaries of the region encompassing a selected point. The updated software and rebuilt database in ZBB1.2 can be downloaded from our website [38].

Finally, as the Zebrafish Brain Browser's strength is primarily in two dimensions (i.e., the visualization of horizontal, transverse, and sagittal slices through the brain), we decided to develop interactive tools to better facilitate 3D exploration. The use of 3D graphics to represent complex structure can also provide a more intuitive sensory experience that avoids cognitive bias or misinterpretation inadvertently introduced by two dimensional reductions [39,40]. By taking advantage of stereoscopy and vestibular-enhanced parallax (head tracking), the more immersive and holistic experience of Virtual Reality (VR) can also significantly improve performance of basic tasks like searching and making comparisons [41,42]. We therefore implemented our Zebrafish Brain Browser in both an open Web3D platform (X3D) and a custom game engine (Unity). First, we converted masks representing anatomical regions to meshes and built a Web3D interface using X3D to inspect the spatial relationship between different brain regions (Fig. 5a,b), available online [43]. Users can navigate within the brain using any web browser, rotating and zooming into brain regions to better interrogate larval neuroanatomy. Second, using the Unity platform we wrote a VR app to view the brain and neuroanatomical regions. By running the app on a cell phone, and inserting it into an inexpensive Google cardboard viewer, users can 'walk into' the brain, and see from the inside the inter-relationship between neuroanatomical domains (Fig. 5c,d), available for download [44].

Discussion

Digitized data-derived brain atlases provide an opportunity to continuously integrate new information and iteratively improve data accuracy within a common spatial framework. Thus, as methods evolve and technology improves, new insights can be easily added to existing data to provide an increasingly rich view of brain structure and function. Because the entire larval zebrafish brain can be rapidly imaged at cellular resolution, it is possible to envisage an atlas that combines detailed information on cell type (including gene expression and morphology), connectivity and activity under a variety of different physiological conditions. At present, biological variability presents an obstacle, as brain regions contain
multiple intermingled cell types that are not positioned in precisely the same manner between larvae. To compensate for this in the existing zebrafish brain atlases, multiple individuals of a given line are sampled and averaged to generate a representative expression pattern. Current atlases are thus essentially heat maps of gene expression or activity. Despite this spatial ambiguity, aggregating information from different sources into the same spatial framework still provides valuable indicators of cell type, gene co-expression, and neural activity under defined conditions.

Ideally different atlas projects might use the same reference brain, however in practice the choice of a reference is often dictated by study-specific experimental requirements. For example, despite the deformations introduced by fixation and permeabilization, a fixed brain is essential for activity mapping using pERK immunohistochemistry. In contrast, we were able to take advantage of the optical transparency of larvae to rapidly scan and register several hundred individuals representing more than 100 different transgenic lines. For our purposes, the TgBAC(slc17a6b:loxP-DsRed-loxP-GFP)tns14 line was ideal, because through Cre injection, we generated a vglut2a:GFP line with an almost identical pattern, allowing us to co-register lines with either GFP or RFP fluorescence. However, we have also used pan-neuronal Cerulean or mCardinal as a reference channel when green and red channels both contain useful information on transgene expression. Our work now demonstrates that it is feasible to contribute to community efforts at building an integrated map of brain structure, expression and activity, while allowing reference image selection to be guided by technical considerations.

One caveat to this conclusion is that deformable image registration can easily introduce artifacts into cell morphology if parameters are not carefully monitored and constrained. Indeed, a special challenge for brain registration in zebrafish is preserving the local morphology of neuronal cell bodies and axons, while permitting sufficient deformation to correct for biological differences and changes in brain structure arising from tissue fixation and permeabilization. Thus, while B-spline registration with CMTK produced acceptable inter-atlas alignment, it also introduced noticeable distortions into local brain structure that affected neuronal cell morphology. Such artifacts were particularly severe in ventral brain regions such as the caudal hypothalamus, and may therefore be due to differences in ventral signal intensity between the datasets. In ZBB, in order to compensate for the increase in light diffraction with tissue depth, we systematically increased laser intensity with confocal scan progression (z-compensation). As a result, the Z-Brain and ZBB datasets are comparable in dorsal brain regions, but there is a noticeable discrepancy ventrally which may account for the loss of registration fidelity. Alternatively, although z-compensation partially corrects for reduced fluorescent intensity, there is a noticeable drop-off in image resolution in ventral regions; the resulting loss of information may lead to lower quality registration. Registration
algorithms that allow parameters to vary by depth may ameliorate the effects of these physical imaging
constraints.

Nevertheless, the symmetrical diffeomorphic transformation in ANTs provides a solution to these
problems. For live tissue, we found parameters that allowed the ANTs SyN transform to achieve similar
or better registration precision than previously achieved using CMTK, while significantly reducing
distortions in tissue structure and neuronal cell morphology. In our hands, permeabilization of fixed tissue
tended to produce variable changes in neuropil structure which was most salient in the optic tectum.
Specifically, neuropil volume was diminished when fresh aliquots of trypsin were used for extended
durations. These artifacts can be minimized by stringent oversight of reagent viridity. However, by
calibrating SyN parameters to permit larger deformations, we were able to accommodate the variability
introduced in tissue processing.

Currently, limitations of the SyN registration algorithm in ANTs are the large memory demands (73 GB
for a single channel registration) and long computational times (3-5 hours for a single channel using 24
cores) required for registration of images with a resolution sufficient for the brain-wide visualization of
neuronal morphology (e.g., 1000 x 600 x 350 pixels). For multi-channel registrations, memory demands
and computation time were even greater: 106 GB for 6 channels taking over 16 hours on 24 cores.
However, our present ANTs SyN parameters likely can be further optimized to reduce these demands. For
instance, our parameters currently include 10 iterations of transformation matrix optimization at full
image resolution. From our experience, these full resolution registration cycles do not significantly
increase precision, but greatly increase computation time. Thus, computation time may be reduced by
adjusting registration resolution as well as other parameters without adversely affecting registration
quality. Although computational resources did not present a bottleneck for registering a small number of
samples, this increase in the demands of a single registration made it difficult to optimize registration
parameters as extensively as we had done previously with CMTK [1]. By reducing computation time, we
would be able to explore more comprehensively the parameter space available with SyN and evaluate
alternative diffeomorphic transforms available with ANTs that may provide still better registration
fidelity.

An obstacle to systematically calibrating registration parameters is finding a suitable metric to
quantitatively evaluate precision. This is a recognized problem, and it is not clear that a general solution
exists [34]. Here, we primarily assessed precision by measuring the distance between visually-located
landmarks in the reference brain, and registered images. However, this method has two drawbacks: (1) it
relies on the accuracy with which these landmarks are located, and (2) at least for our sample set, a relatively limited set of landmarks could reliably be identified. We obtained similar results when we assessed precision using cross-correlation within localized image neighborhoods that included high contrast internal image boundaries (data not shown). In registering live \textit{vglut2a:DsRed} image stacks, we noted the trade-off between accurate global brain alignment and biologically plausible cell morphology. Thus we also used a set of measures to assess changes in the morphology of manually segmented cells (Hausdorff distance, elongation index and cell volume). Finally, we also inspected the output of every transformation to subjectively judge registration quality.

This study demonstrates that the ANTs diffeomorphic symmetric normalization algorithm (SyN) advances upon elastic registration for precise registration of whole brain images in larval zebrafish and is markedly better at preserving neuronal cell morphology. By systematically testing SyN registration parameters for registering images acquired using live scans, we improved the ZBB atlas. Then, after calibrating registration parameters for fixed tissue and using multi-channel optimization, we were able to align the Z-Brain atlas into the ZBB coordinate space, and vice-versa, achieving co-registration accuracy to approximately the diameter of a single neuron. We believe that integrating the information present in each of these atlases produces a richer framework for future studies of structural and functional relationships within the nervous system. Large digital datasets such as those present in brain atlases can be used for many types of bioinformatic analysis. Z-Brain and ZBB already include software that can be used to explore the larval zebrafish brain, and we hope that integrating these datasets into a single coordinate system, will help to stimulate the development of additional computational tools and methods for querying this information.

**Availability of supporting data**

All individual brain scans, both before and after registration to a ZBB reference brain, are available in the \textit{GigaScience} repository, GigaDB \cite{35}. The GigaDB repository also includes the set of reference brains used for ZBB \cite{33} and the transformation matrices used to convert between ZBB and Z-Brain coordinate systems \cite{37}.

**Abbreviations**

- ac, anterior commissure
- Ce, cerebellum
- DT, Thalamus
- GT, Griseum tectale
Ha, Habenula  
Hc, Hypothalamus caudal zone  
Hi, Hypothalamus intermediate zone  
MO, Medulla oblongata  
NXm, Vagus motor neurons  
OB, Olfactory bulb  
OE, Olfactory epithelium  
ON, Olfactory nerve  
IO, Inferior olive  
LC, Locus coeruleus  
MN, Mauthner neuron  
MO, Medulla oblongata  
Pal, Pallium  
p, posterior commissure  
Po, preoptic region  
Pr, Pretectum  
SR, Superior raphe  
Teg, Tegmentum  
TeOn, Optic tectum neuropil  
TG, Trigeminal ganglion  
TL, Torus longitudinalis  
TS, Torus semicircularis

Competing Interests

The authors declare that they have no competing interests.

Authors' Contribution

GDM and HAB conceived the experiments. GDM, KMT, EJH and HAB optimized ANTS for zebrafish brain registration. GDM, MB and AKG contributed confocal brain scans and generated meshes. NFP developed the X3D/HTML5 based browser. DDN developed the Unity VR browser. GDM and HAB wrote the manuscript. All authors approved the final manuscript.

Acknowledgements

This work was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD) and utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD [45]. We thank Owen Randlett for valuable discussion and help checking the correspondence of vglut2a.
expression patterns. We are grateful to Sinisa Pajevic (NIH/CIT) for advice on computational procedures and to M. Okan Irfanglu and Neda Sadeghi (NICHD) for guidance in optimizing ANTs' parameters.

References

1. Marquart GD, Tabor KM, Brown M, et al. A 3D Searchable Database of Transgenic Zebrafish Gal4 and Cre Lines for Functional Neuroanatomy Studies. Front Neural Circuits 2015; 9:78.

2. Randlett O, Wee CL, Naumann EA, et al. Whole-brain activity mapping onto a zebrafish brain atlas. Nat Methods 2015; 12:1039–1046.

3. Ronneberger O, Liu K, Rath M, et al. ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains. Nat Methods 2012; 9:735–42.

4. Rath M, Nitschke R, Filippi A, Ronneberger O, Driever W. Generation of high quality multi-view confocal 3D datasets of zebrafish larval brains suitable for analysis using Virtual Brain Explorer (ViBE-Z) software. Protoc Exch 2012.

5. Portugues R, Feierstein CE, Engert F, Orger MB. Whole-Brain Activity Maps Reveal Stereotyped, Distributed Networks for Visuomotor Behavior. Neuron 2014; 81:1328–1343.

6. Rohlfing T, Maurer CR Jr. Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. IEEE Trans Inf Technol Biomed 2003; 7:16–25.

7. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 2008; 12:26–41.

8. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage 2011; 54:2033–2044.

9. Higashijima S, Hotta Y, Okamoto H. Visualization of cranial motor neurons in live transgenic zebrafish expressing green fluorescent protein under the control of the islet-1 promoter/enhancer. J Neurosci 2000; 20:18.

10. Wen L, Wei W, Gu W, et al. Visualization of monoaminergic neurons and neurotoxicity of MPTP in live transgenic zebrafish. Dev Biol 2008; 314:84–92.

11. Xi Y, Yu M, Godoy R, Hatch G, Poitras L, Ekker M. Transgenic zebrafish expressing green fluorescent protein in dopaminergic neurons of the ventral diencephalon. Dev Dyn 2011; 240:2539–47.

12. Lillesaar C, Stigloher C, Tannhauser B, Wullimann MF, Bally-Cuif L. Axonal projections originating from raphe serotonergic neurons in the developing and adult zebrafish, Danio rerio, using transgenics to visualize raphe-specific pet1 expression. J Comp Neurol 2009; 512:158–82.

13. Tabor KM, Bergeron SA, Horstick EJ, et al. Direct activation of the Mauthner cell by electric field pulses drives ultra-rapid escape responses. J Neurophysiol 2014; 112:834–844.
14. Scott EK, Baier H. The cellular architecture of the larval zebrafish tectum, as revealed by gal4 enhancer trap lines. *Front Neural Circuits* 2009; 3:13.

15. Satou C, Kimura Y, Hirata H, Suster ML, Kawakami K, Higashijima S. Transgenic tools to characterize neuronal properties of discrete populations of zebrafish neurons. *Development* 2013; 140:3927–31.

16. McLean DL, Fan J, Higashijima S, Hale ME, Fetcho JR. A topographic map of recruitment in spinal cord. *Nature* 2007; 446:71–5.

17. Pittman AJ, Law MY, Chien CB. Pathfinding in a large vertebrate axon tract: isotypic interactions guide retinotectal axons at multiple choice points. *Development* 2008; 135:2865–71.

18. Yokogawa T, Hannan MC, Burgess HA. The dorsal raphe modulates sensory responsiveness during arousal in zebrafish. *J Neurosci* 2012; 32:15205–15215.

19. Satou C, Kimura Y, Higashijima S. Generation of multiple classes of V0 neurons in zebrafish spinal cord: progenitor heterogeneity and temporal control of neuronal diversity. *J Neurosci* 2012; 32:1771–83.

20. Bae YK, Kani S, Shimizu T, *et al.* Anatomy of zebrafish cerebellum and screen for mutations affecting its development. *Dev Biol* 2009; 330:406–26.

21. Fosque BF, Sun Y, Dana H, *et al.* Labeling of active neural circuits in vivo with designed calcium integrators. *Science* 2015; 347:755–760.

22. Nechiporuk A, Linbo T, Poss KD, Raible DW. Specification of epibranchial placodes in zebrafish. *Development* 2007; 134:611–623.

23. Burgess HA, Johnson SL, Granato M. Unidirectional startle responses and disrupted left-right coordination of motor behaviors in robo3 mutant zebrafish. *Genes Brain Behav* 2009; 8:500–11.

24. Inoue D, Wittbrodt J. One for All—A Highly Efficient and Versatile Method for Fluorescent Immunostaining in Fish Embryos. *PLOS ONE* 2011; 6:e19713.

25. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012; 9:671–675.

26. Williams G. *NIfTi Input/Output*. https://imagej.nih.gov/ij/plugins/nifti.html (last accessed, 22 March 2017)

27. Heisenberg M, Schmid B. *IntSeg 3D*. http://3dviewer.neurofly.de/IntSeg_3D (last accessed, 22 March 2017)

28. Web3D Consortium. *Open Standards for Real-Time 3D Communication*. http://www.web3d.org/ (last accessed, 22 March 2017)

29. Vicomtech-IK4. *AtlasConversionScripts*. https://github.com/VolumeRC/AtlasConversionScripts (last accessed, 22 March 2017)

30. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Mach Intell* 1993; 15:850–863.
31. Klein A, Andersson J, Ardekani BA, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* 2009; **46**:786–802.

32. Murphy K, Ginneken B van, Reinhardt JM, et al. Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge. *IEEE Trans Med Imaging* 2011; **30**:1901–1920.

33. ZBB Reference Brains. http://helix.nih.gov/~BurgessLab/reference_brains.zip (last accessed, 22 March 2017) To be replaced with GigaDB DOI upon publication

34. Rohlfing T. Image Similarity and Tissue Overlaps as Surrogates for Image Registration Accuracy: Widely Used but Unreliable. *IEEE Trans Med Imaging* 2012; **31**:153–163.

35. ZBB Raw and Registered Brain Images GigaDB DOI to be provided upon publication

36. Kinkhabwala A, Riley M, Koyama M, et al. A structural and functional ground plan for neurons in the hindbrain of zebrafish. *Proc Natl Acad Sci U A* 2011; **108**:1164–9.

37. ZBB-ZBrain Transformation Matrices. https://helix.nih.gov/~BurgessLab/zbb-zbrain.zip (last accessed, 22 March 2017) To be replaced with GigaDB DOI upon publication

38. Zebrafish Brain Browser. https://science.nichd.nih.gov/confluence/display/burgess/Brain+Browser (last accessed, 22 March 2017)

39. Ota D, Loftin B, Saito T, Lea R, Keller J. Virtual reality in surgical education. *Comput Biol Med* 1995; **25**:127–137.

40. Ware C, Franck G. Viewing a graph in a virtual reality display is three times as good as a 2D diagram. In: *Proceedings of 1994 IEEE Symposium on Visual Languages*.; 1994:182–183.

41. Bowman DA, McMahan RP. Virtual Reality: How Much Immersion Is Enough? *Computer* 2007; **40**:36–43.

42. Henry JAG, Polys NF. The effects of immersion and navigation on the acquisition of spatial knowledge of abstract data networks. *Procedia Comput Sci* 2010; **1**:1737–1746.

43. Zebrafish Anatomy Explorer. http://metagrid2.sv.vt.edu/~npolys/NIH/Burgess/ZebraFishBrowser/ZebrafishBrowser.html (last accessed, 22 March 2017)

44. Virtual reality brain browser. https://science.nichd.nih.gov/confluence/display/burgess/Brain+Browser#BrainBrowser-vrfishbrain (last accessed, 22 March 2017)

45. *High Performance Computing at the NIH. 2016*. https://hpc.nih.gov (last accessed, 22 March 2017)

**Figure legends**

**Table 1. ANTs command parameters for image registration**
Table 2. Brain images in ZBB and Z-Brain that were used as templates for registration and/or for measurement of registration precision.

Figure 1. Optimization of parameters for registration of live brain scans using ANTs

(a) Overview of parameter optimization for live brain scans using ANTs. A calibration set of 6
$vglut2a:DsRed$ confocal stacks with 10 point-based landmarks and 107 cell masks were registered to the
$vglut2a_ZBB$ reference with the same 10 point-based landmarks defined (left). MLDs for landmarks and
Hausdorff distance for transformed cell masks compared to their originals were measured for each
parameter set (middle). Optimal parameters selected from these metrics (b-d) were used to re-register all
lines generating ZBB$_{1.2}$ where MLD was measured from 2 additional landmarks in each of 12 co-aligned
patterns (right).

(b) Hausdorff distance for cell shape comparison plotted against MLD for 68 sets of parameters tested
using ANTs (grey and blue circles) and after registration using CMTK (orange). Blue circles labeled a-f
indicate the Pareto frontier.

(c) Mean absolute change in cell volume (as a fraction of the original volume) produced by
transformations resulting from parameter sets a-f and CMTK in (b). * p < 0.05, compared to CMTK.

(d) Mean elongation index for cells after registration using parameter sets a-f and CMTK in (b). Dashed
line shows index for cells before registration — all transformations produced a significant increase in
compared to the untransformed cells. * p < 0.05, compared to CMTK.

(e) Horizontal section through the medulla oblongata in $vglut2a_ZBB$, and of a representative $vglut2a:DsRed$ brain after registration using
CMTK or ANTs. Distortion artifacts are indicated (arrow). Scale bar 50 μm

(f, g) Horizontal section in JI229aGt showing expression of GFP in the Mauthner cell and axon
(arrowheads) for three individual larvae (pseudo-colored red, green and blue). Registration was performed
with CMTK (f) or ANTs (g). Scale bar 100 μm.

(h, i) Transverse section through the optic tectum in two separate average brain images (colored green and
magenta) for y393Et. For each brain image, we independently scanned three individual brains and
registered them using CMTK (h) or ANTs (i). Scale bar 100 μm.

Figure 2. Improved precision of transgene representations in ZBB$_{1.2}$

(a) Mean landmark distances for 24 landmarks, after registration with CMTK and ANTs. Dotted line
indicates 1:1 ratio.

(b) Boxplot of data in (a). * paired t-test, N=12 lines, p = 0.019
(c) Difference in MLD between ANTs and CMTK plotted against distance from the dorsal-most point in the brain.

(d) Horizontal section through the caudal hypothalamus of three individual y341Et larvae as well as their pseudo-colored superimposition following registration with CMTK (top row) or ANTs (bottom row).

(e,f) Horizontal section through the thalamus showing the averaged representation of enhancer trap line y304Et, where individual brains were registered with CMTK for ZBB (e), or with ANTs for ZBB₁₂ (f). Arrow indicates neurons that are artificially elongated across the midline. Scale bar 100 µm.

(g,h) Transverse section through the caudal hypothalumus showing the average enhancer trap line y269Et brain registered with CMTK (g) or with ANTs (h). Arrow shows distortion of cells causing the caudal hypothalamus to appear dorsally elongated. Scale bar 50 µm.

(i,j) Transverse section through the medulla oblongata showing the average phox2b:GFP brain with CMTK (i) or ANTs (j). Scale bar 50 µm.

(k,l) Horizontal projection through the posterior commissure (arrow) for the average y351Et brain obtained with CMTK (k) or ANTs (l). Scale bar 100 µm.

Figure 3. Optimization of ANTs registration parameters for fixed tissue

(a) Overview of parameter optimization for fixed brain scans using ANTs. A calibration set of 6 tERK confocal stacks with segmentations of the tectal neuropil were registered to tERKZBB, a tERK and vglut2a:DsRed confocal scan previously aligned to the vglut2aZBB reference (left). MCCs were calculated between eighteen 50 µm-side cube high-contrast sub-regions in the calibration set and in the tERKZBB reference to identify parameters that maximized MCC (f,g) and improved the Jaccard index of tectal neuropil segmentation (h) while compensating for fixation artifacts (c,e,k) (middle). These optimized ANTs parameters allow for the accurate registration of fixed tissue and the generation of a tERK average reference (tERKᵥAᵥ) useful for bridging live and fixed tissue registrations (right).

(b,c) Horizontal section through the optic tectum of tERK immunostained(red) vglut2a:DsRed (green) larvae, using diluted (b, sample A) or fresh trypsin (c, sample B). Asterisk indicates missing area of tectal neuropil due to permeabilization artifact.

(d,e) Horizontal section through the same stacks as in (b,c) registered to tERKZBB using the parameters previously optimized for live registration. Gray shows the ZBB₁₂ vglut2a:DsRed pattern. Arrowheads highlight regions where tERK in the optic tectum neuropil fails to closely abut the adjacent glutamatergic cellular layer.

(f) MCC for tERK expression after registration of 6 brains to tERKZBB, varying each of the parameters for the ANTs SyN transform, starting with the parameters that gave the best registration for live
vglut2a:DsRed based registration (SyN[0.05,6,0.5]). Bottom right: MCCs after varying the radius of the cross-correlation metric used during registration.

(g) MCCs for tERK in the same brains as in (f), after combining the two best parameter sets from (f) (SyN[0.1,6,0.5] and SyN[0.05,6,0]) to assess further improvement in registration precision. Yellow box highlights the final optimal parameter set.

(h) Jaccard index for overlap of the manually segmented tectal neuropil of the reference brain, with each of the 6 brains in the calibration set. p < 0.01

(i) 3D view of overlap between segmented tectal neuropils from tERK-ZBB (red) and the Z-Brain tERK reference brain (green), after registration with ANTs using parameter optimal for live registration, fixed registration and CMTK.

(j,k) Same brains as in (d,e), but after registration to tERK-ZBB using the parameters optimized for fixed tissue registration.

(l,m) Horizontal section through the optic tectum showing tERK expression (red) and vglut2a:DsRed expression (green) in ZBB1,2 (l) and Z-Brain (m). Matching slices within the optic tectum were selected; because the rotation around the y-axis is slightly different, sections are different within the medulla.

Figure 4. Transformation between Z-Brain and ZBB coordinate systems using multi-channel registration

(a) Overview of bridging Z-Brain and ZBB using ANTs multi-channel registration. Combinations of 5 patterns common between Z-Brain and ZBB (vglut2aAV, tERKZBB, vmat2, isl2b and elavl3) were used to guide multi-channel bridging registrations (left). MLDs for 41 landmarks in gad1b, glyT2, isl1, isl2b, tERK, vglut2a, and vmat2 expression were measured for all reference channel combinations (middle). The combination of vglut2aAV, tERKZBB, vmat2, and isl2b enabled the most accurate bridging of ZBB and Z-Brain allowing the combination of the large collection of live transgenic lines of ZBB with the fixed tissue techniques and expert neuroanatomic segmentations of Z-Brain (right).

(b) MLDs for the expression patterns of gad1b, glyT2, isl1, isl2b, tERKZBB, vglut2a and vmat2 and M1 and M2 metrics after registration of Z-Brain to ZBB1,2 using either CMTK or ANTs SyN with fixed-tissue registration parameters and the indicated combination of reference channels (vglut2a, tERKZBB, vmat2, isl2b, and elavl3). Note, similar results were obtained using tERKAV instead of the tERKZBB, but are omitted for clarity. The combination of reference channels selected for co-registration of Z-Brain and ZBB is highlighted.
Transverse section through the caudal optic tectum showing the vglut2a pattern in ZBB1.2, Z-Brain, Z-Brain after registration to ZBB with CMTK (Z-Brain-CMTK), or with ANTs (Z-Brain-SyN). The torus longitudinalis (TL) is well separated from tectal neurons in live scans, but less so in fixed tissue (arrows). The TL appears flattened after CMTK registration, but retains normal morphology after registration with ANTs SyN.

Transverse sections as in (e), but slightly more caudal with contrast increased to highlight ventral distortion artifacts produced by registration (arrowheads).

Brain Browser views in the ZBB1.2 coordinate (e,f) or Z-Brain coordinate (g,h) space. Scale bars 25 μm except 50 μm in (e).

Horizontal (top) and sagittal (bottom) sections, comparing the pet1:GFP expression pattern in the superior raphe in ZBB1.2 (red) and Z-Brain after transformation to the ZBB coordinate system (green).

Horizontal (top) and transverse (bottom) sections through the medulla oblongata, showing the expression of y264Et from ZBB1.2 (red) and s1181Et from Z-Brain after transformation to ZBB1.2 (green), which both label the Mauthner cells (arrowhead).

Horizontal (top) and transverse (bottom) sections through the pretectum, comparing the expression of DAT:GFP from ZBB1.2 after transformation to Z-Brain (red) and anti-tyrosine hydroxylase staining in Z-Brain (green).

Horizontal (top) and transverse (bottom) sections through the medulla oblongata for glyT2:GFP from ZBB1.2 after transformation to Z-Brain (red) and the same transgenic line in Z-Brain (green).

Brain Browser horizontal sections showing manually segmented regions transformed from the Z-Brain coordinate system to ZBB1.2 (white outlines) compared to regions previously defined in ZBB obtained by thresholding expression patterns in transgenic lines (magenta). Regions are the torus longitudinalis (i), habenula (j), anterior commissure (k) and trigeminal ganglion (l).

Figure 5. 3D visualization of brain browser data
(a) X3D zebrafish brain shown in HTML5 Web browser and (b) Virginia Tech HyperCube (CAVE) (c) Virtual reality brain rendered using the Unity Game Engine for stereoscopic viewing using the Google Cardboard viewer. (d) In the VR browser, brain regions are selected using a menu on the floor of the virtual arena.

Additional Material

Additional File 1.pdf

Point-based landmarks for quantification of live-scan registration precision.
(a) Landmarks used for measuring registration precision. Position specifies the coordinates on vglut2a ZBB (transverse, sagittal, horizontal planes). View indicates whether the image plane shown in (b) is or transverse (T), horizontal (H), or sagittal (S). MLDs represent the average precision for each landmark for the set of 6 calibration brains, after registration with CMTK or ANTs.

(b) Images of the landmarks in vglut2a ZBB (red) used for measuring precision superimposed on elavl3 (gray).

(c) Position of the landmarks superimposed on horizontal (top) and sagittal (bottom) maximum projections of elavl3 through the brain.

(d) Horizontal maximum projections showing the landmark point (red dot), and the position of the corresponding landmarks in the six calibration brains after registration (green dots) superimposed on vglut2a ZBB. Scale bar 20 µm.

Additional File 2.pdf
Cells segmented for assessing distortion introduced by registration
(a) Position of manually segmented cells for measurement of distortion introduced by registration. Views show the same cells (individually color coded) superimposed on horizontal (top) and sagittal (bottom) maximum elavl3 brain projections.

(b) Two examples of cells showing (left to right): original confocal image, segmentation mask, mask after alignment with CMTK, and mask after alignment with ANTs.

Additional File 3.pdf
Point-based landmarks labeled by transgenic lines
(a) Transgenic line landmarks used for measuring registration precision of the zebrafish brain browser atlas. Coordinates give the transverse, horizontal, sagittal position. Letter in square brackets designates Right side [R], Left side [L], or Midline [M]. The mean and standard error of the landmark distances for the three brains per landmark are indicated for CMTK and ANTs.

(b) Position of the landmarks superimposed on horizontal (top) and sagittal (bottom) maximum elavl3 brain projections.

Additional File 4.pdf
ZBB and Z-Brain expression patterns used for atlas registration
Brain Browser 3D projections of corresponding expression patterns in Z-Brain (left) and ZBB (right) used for calibrating and verifying the precision of inter-atlas registration. The top 5 patterns were
combinatorially used to drive registration, while the bottom 3 were used for assessing precision. Middle images show Z-Brain patterns after registration to ZBB.

Additional File 5.pdf

**Point-based landmarks for measuring precision of Z-Brain/ZBB co-registration**

(a) Transgenic line and tERK-stain landmarks used for measuring registration precision of registration between Z-Brain and ZBB. Coordinates are in transverse, horizontal, sagittal sections. Letters in square brackets designate Right side [R], Left side [L], or Midline [M]. Color blocks correspond to points in (b).

(b) Position of the landmarks superimposed on horizontal (top) and sagittal (bottom) maximum brain projections.

Additional File 6.doc

**Instructions for using transformation matrices to convert between ZBB and Z-Brain coordinate systems.**
High precision registration between zebrafish brain atlases using symmetric diffeomorphic normalization

Gregory D. Marquart¹², Kathryn M. Tabor¹, Eric J. Horstick¹, Mary Brown¹, Alexandra K. Geoca¹, Nicholas F. Polys³, Damian Dalle Nogare¹ and Harold A. Burgess¹*

¹ Division of Developmental Biology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD 20892
² Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742
³ Advanced Research Computing, Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

* Contact: burgessha@mail.nih.gov ; 301-402-6018

Running title: High fidelity inter-atlas registration using ANTs SyN transform
Keywords: zebrafish, atlas, ANTs, SyN, registration, diffeomorphism, normalization, brain imaging, transgenic, virtual reality, Unity, X3D
Abstract

Atlases provide a framework for information from diverse sources to be spatially mapped and integrated into a common reference space. In particular, brain atlases allow regional annotation of gene expression, cell morphology, connectivity and activity. In larval zebrafish, advances in genetics, imaging and computational methods have enabled the collection of large datasets providing such information on a whole-brain scale. However, datasets from different sources may not be aligned to the same spatial coordinate system, because technical considerations may necessitate use of different reference templates. Two recent brain atlases for larval zebrafish exemplify this problem. The Z-Brain atlas contains information on gene expression, neural activity and neuroanatomical segmentations acquired using immunohistochemical staining of fixed tissue. In contrast, the Zebrafish Brain Browser (ZBB) atlas was constructed from live scans of fluorescent reporter genes in transgenic larvae. Although different reference brains were used, the two atlases included several transgenic patterns in common that provide potential ‘bridges’ for transforming each into the other’s coordinate space. We tested multiple bridging channels and registration algorithms. The symmetric diffeomorphic normalization (SyN) algorithm in ANTs improved the precision of live brain registration while better preserving cell morphology than the previously used B-spline elastic registration algorithm. SyN could also be calibrated to correct for tissue distortion introduced during fixation and permeabilization. Multi-reference channel optimization provided a transformation that enabled Z-Brain and ZBB to be co-aligned with high precision and minimal perturbation of cell and tissue morphology. Finally, we developed software to visualize brain regions in 3-dimensions, including a virtual reality neuroanatomy explorer. This study demonstrates the feasibility of integrating whole brain datasets, despite disparate reference templates and acquisition protocols, when sufficient information is present for bridging. This increased accuracy and interoperability of digital brain atlases in larval zebrafish will facilitate future neurobiological studies.
bridging channels and registration algorithms and found that the symmetric diffeomorphic normalization (SyN) algorithm improved live brain registration precision while better preserving cell morphology than B-spline based registrations. SyN also corrected for tissue distortion introduced during fixation. Multi-reference channel optimization provided a transformation that enabled Z-Brain and ZBB to be co-aligned with precision of approximately a single cell diameter and minimal perturbation of cell and tissue morphology. Finally, we developed software to visualize brain regions in 3-dimensions, including a virtual reality neuroanatomy explorer. This study demonstrates the feasibility of integrating whole brain datasets, despite disparate reference templates and acquisition protocols, when sufficient information is present for bridging. Increased accuracy and interoperability of zebrafish digital brain atlases will facilitate neurobiological studies.

Background

Larval stage zebrafish are an increasingly popular model for neurobiological studies. With a brain that contains an estimated $10^5$ neurons, larvae are similar in complexity to adult Drosophila, another established neuroscience model. In both systems, researchers can deploy a wide range of genetic tools in efforts to decode patterns of neural structure and connectivity. In larval zebrafish, optical transparency and constrained physical dimensions (fitting within an imaging volume of 1000 x 600 x 350 μm) allow the entire brain to be rapidly scanned at cellular resolution using diffraction-limited microscopy. In principle, this enables researchers to systematically analyze effects of manipulations on a brain-wide level. However, such efforts have been hampered by the absence of a comprehensive digital atlas that would provide researchers with a unified framework in which to aggregate data from different experiments and gain deeper insights from correlations between neuronal cell identity, connectivity, gene expression and function within the brain. Additionally, digital atlases may more clearly delineate structural boundaries that are difficult to accurately identify within individual brains, allowing for a more rigorous mapping of neuroanatomical regions onto experimental data.

These longstanding problems in zebrafish neuroscience have recently been addressed by the construction of digital atlases using three-dimensional (3D) image registration techniques: the Virtual Brain Explorer for Zebrafish (ViBE-Z), Z-Brain and the Zebrafish Brain Browser (ZBB) [1–3]. In these atlases, information on gene expression, structure (neuronal cell bodies, glia, vasculature, ventricles, neuropil or axon tracts) and measures of activity (calcium or secondary messenger activity) are consolidated within a common spatial framework. By using widely-available transgenic lines or immunohistochemical stains as reference templates for brain alignment, each of these atlases provides other researchers the opportunity
to register their own datasets into these digital spaces and take advantage of the information contained within.

ViBE-Z was the first comprehensive three-dimensional digital brain atlas in zebrafish that used a nuclear stain for the alignment of 85 high resolution scans comprising 17 immunohistochemical patterns at 2-4 days post-fertilization (dpf) [3,4]. In ViBE-Z, custom algorithms were developed to correct for variations in fluorescent intensity with scan depth, and a landmark approach taken to perform accurate image registration and segmentation into 73 neuroanatomic regions.

In contrast, two more recent approaches (Z-Brain and ZBB) have generated brain atlases at 6 dpf through non-linear B-spline registration using the freely available Computational Morphometry Toolkit (CMTK) [5,6]. Z-Brain includes 29 immunohistochemical patterns from 899 scans which form the basis for expert manual segmentation of the brain into 294 neuroanatomic regions. These partitions facilitate the analysis of phospho-ERK expression for mapping neural activity [2]. In Z-Brain, each expression pattern was co-scanned with tERK immunoreactivity, and registered to a single tERK-stained reference brain. For ZBB, we live-imaged 354 brains from 109 transgenic lines and manually annotated the expression found in each [1]. In place of tERK, a single vglut2a:dsRed transgenic brain was used as the reference in ZBB with transgenic lines crossed and co-imaged with this channel for registration. Brain browser software enables researchers to select a transgenic line labeling a selected set of neurons for monitoring and manipulating circuit function.

While Z-Brain and ZBB are powerful datasets on their own, we saw an opportunity to merge the two atlases because they are both based on confocal scans of 6 dpf larvae. This would bring to Z-Brain a large number of additional transgenic lines and to ZBB, the expert manual segmentation of Z-Brain. Several similarities between Z-Brain and ZBB suggested that bridging the atlases would be possible. First as zebrafish rearing conditions are standardized across laboratories and fish were imaged at the same time post-fertilization, Z-Brain and ZBB likely reflect the same developmental timepoint. Second, images in both atlases were acquired at similar resolution (0.8 x 0.8 x 2 μm for Z-Brain; 1 x 1 x 1 or 1 x 1 x 2 μm for ZBB) and orientation (dorsal to ventral horizontal scans). Third, despite using distinct templates (tERK for Z-Brain and vglut2a for ZBB), Z-Brain and ZBB have several transgenic markers in common, which provided the possibility of bridging the datasets by using these shared patterns as references for a secondary registration step.
One of the strengths of larval zebrafish is the ability to rapidly image at cellular resolution and visualize brain-wide neuronal morphology, providing valuable information on cell type and potential connectivity. Z-Brain and ZBB both illustrate the feasibility of performing whole-brain registration with precision sufficient to ensure that the 'same' neurons from different fish are aligned to within a cell diameter (~10.8 μm). However, a challenge for brain registration in zebrafish is to minimize local distortions, so that cellular morphology is preserved while still allowing sufficient deformation to overcome biological variability between individual brains or malformations due to tissue processing.

Here we describe a method to co-register ZBB and Z-Brain, bridging the two existing 6 dpf larval zebrafish brain atlases. By using the diffeomorphic algorithm SyN in the Advanced Normalization Tools (ANTS) software package [7,8], we were able to overcome differences in tissue shape due to fixation, optimize the trade-off between preservation of cell morphology and global alignment, and provide precise registration in all tested brain regions. Additionally, ANTs provided superior image registration for live-scanned larvae, enabling us to improve the precision of registration and neuron morphology within our original ZBB atlas, allowing us to compile a new version with increased fidelity (ZBB1.2).

Methods

Zebrafish lines.

In order to provide additional options for bridging ZBB and Z-Brain, we scanned two transgenic lines that were not in the original ZBB release: Et(gata2a:EGFP)zf81 (vmat2:GFP) and Tg(isl1:GFP)rw0 (isl1:GFP) [9,10]. Aside from the use of ANTs, the basic imaging and registration workflow was performed as previously described [1]. Other lines referenced in this study are Tg(slc6a3:EGFP)ot80 (DAT:GFP) [11], Tg(-3.2fev:EGFP)ne0214 (pet1:GFP) [12], y264Et [13], s118Et [14], Tg(gad1b:GFP)mn25 (gad1b:GFP) [15], Tg(slc6a5:GFP)cf3 (glyT2:GFP) [16], Tg(-17.6isl2b:GFP)z7 (isl2b:GFP) [17], Tg(-3.4ph2:Gal4ff)y228 (tph2:Gal4) [18], TgBAC(slc17a6b:lox-DsRed-lox-GFP)mn14 (vglut2a:dsRedDsRed) [19], Tg(slc17a6:EGFP)zf139 [20], Tg(elavl3:CaMPARI(W391F+V3987L))yf9 [21], Tg(phox2b:GFP)w37 [22], J1229aGt [23] and several Gal4 enhancer traps from ZBB: y304Et, y332Et, y341Et, y351Et and y393Et [1]. All in vivo experimental protocols were approved by the NICHD animal care and use committee.

Immunohistochemistry.

Immunolabeling was as described [2] with the following adaptations. Larvae were fixed overnight at 4°C in PBS with 4% paraformaldehyde and 0.25% Triton X-100. Samples were then washed in PBS
containing 0.1% Triton X-100 (PBT) 3 times for 5 min. For antigen retrieval, samples were incubated in
150 mM Tris-HCl pH 9.0 for 5 min **at room temperature (RT)**, followed by 15 min at 70°C and washed in
PBT 2 times for 5 min **at RT** [24]. Critically, samples were then permeabilized on ice in fresh 0.05%
trypsin-EDTA for no more than 5 minutes. If pigmented, samples were incubated in PBT with 1.5% H₂O₂
and 50 mM KOH for 15 min, rinsed 2 times in PBT and washed again for 10 min, **all at RT**. Samples
were then blocked in PBT containing 5% normal goat serum (NGS) and 0.2% bovine serum albumin
(BSA) for 1 hr **at RT** before incubation at 4°C with tERK antibodies (Cell Signaling, 4696) diluted 1:500
in PBT with 5% NGS and 0.2% BSA for a minimum of 6 hr. Samples were then washed with PBT 4
times for 30 min **at RT** before incubation at 4°C for a minimum of 2 hr with fluorescent secondary
antibodies (Alexa Fluor 488 or 548) diluted 1:1000 in PBT with 5% NGS and 0.2% BSA. Samples were
finally rinsed 4 times for 30 min **at RT** prior to imaging.

**Registration.**

Registrations were performed using CMTK version 3.2.3 and ANTs version 2.1.0 running on the National
Institute of Health’s Biowulf Linux computing cluster. Registrations were parallelized using Slurm-based
bash scripts available upon request. For CMTK, previously optimized registration parameters that
minimize computation time while maximizing precision were used (**affine parameters**: registrationx --dofs
12 --min-stepsize 1: **elastic parameters**: warp x --fast --grid-spacing 100 --smoothness-constraint-
weight 1e-1 --grid-refine 2 --min-stepsize 0.25 --adaptive-fix-thresh 0.25). For ANTs registrations, the
parameters used are cited in the relevant text and figures with optimized parameters listed in Table 1. **All
deformable transformations are initiated with a rigid and affine step** (parameters included in Table 1).

Aside from the use of ANTs, the basic imaging and registration workflow was performed as previously
described [11]. Image volumes were rendered within the Zebrafish Brain Browser (ZBB), ImageJ [25] or
code written in IDL (Harris Geospatial Solutions). For the conversion to/from NIfTi format required for
ANTS, we used the ImageJ plugin nifti_io.jar written by Guy Williams [26].

**Volume rendering & 3D visualization.**

Binary masks corresponding to 25 anatomical regions from Z-Brain aligned to ZBB were converted into
meshes using the Create Surfaces tool in the IntSeg_3D.jar plugin for ImageJ [27]. Edges for individual
meshes were iteratively reduced below 5000 and vertices (single-precision floating-points of the
triangular meshes) written as OBJ files. As there is no intrinsic color or color conventions as of yet for
these brain structures, we used color hue as a nominal categorical coding for each region. To maximize
accessibility, we rendered meshes in Extensible 3D (X3D) format, an ISO (International Organization for
Standardization) standard developed by the not-for-profit Web3D Consortium [28]. This format allows
portability between numerous tools and applications as well as deployment across a broad spectrum of
platforms. For the rendering, previously generated OBJs were transcoded into ImageTextureAtlas PNGs
using X3D's standard IndexedFaceSet to represent mesh information and then tiled at different
resolutions (4096 & 8192 pixels squared) using AtlasConversionScripts [29]. Additionally, dask and
pyimg python libraries were used to generate volume norms (image and ImageTextureAtlas files) by
gradient descent. All renderings were then merged into a single X3D XML scene which was losslessly
compressed (in SRC/glTF) to a final size of 4.5 MB. This makes the scene compact enough to be
visualized on a cell phone, yet still retaining details for visualization and editing in more immersive
virtual reality environments. Finally, X3D files were published to HTML5 via the X3DOM library and a
simple user interface created that allows for the visibility of different structures to be toggled on and off.

Brain meshes were converted to FBX files for import into Unity using Blender 2.78a (Blender foundation,
Amsterdam, NL) and mobile app development for Google Cardboard VR headsets performed in Unity
5.4.2 (Unity Technologies SF, San Francisco CA) using the Google VR for Unity SDK (Google,
Mountain View CA). Custom scripts controlling movement and mesh display were written for Unity in C#.

Measurements

Cross correlation between registered image sets was performed using the c_correlate function within IDL
version 7.0. Correlations were run within small sub-regions of the registered image volumes. In Fig. 1,3 & 4, 50 μm side cube sub-regions were manually defined by selecting volumes containing high contrast
boundaries. For cross correlations between individual brains scanned for each transgenic line in ZBB (Fig
2a,b), 40 μm side cubes were drawn around the three computationally identified brightest sub-regions
within the expression pattern, with cross-correlation then calculated between all pairs of brains. The mean
of all cross-correlations was used to estimate registration precision.

Mean Landmark Distance (MLD). To assess registration precision using MLDs, corresponding landmarks
were located and annotated on the reference brain, and on unregistered brains. In each case, landmarks
were chosen to be widely distributed within the brain, and readily recognized in corresponding brain
scans. In addition, to verify recognizability, the vglut2a landmarks in the reference brain were located by
3 blinded scorers; mean distance from each of the 10 reference points ranged from 1.7 to 11.8 μm (mean,
4.5 ± 0.9 μm). Using ImageJ, we positioned a 3 micron cube centered on each landmark in a second
channel for each brain scan, then, after registering the brain scan using the first channel, applied the
resulting transformation matrix to the second channel, using Nearest Neighbor interpolation for both
CMTK and ANTs. Landmark distance was taken as the distance between the geometrical center of the
corresponding cubes in the reference image, and in the registered image.
Hausdorff distance. We manually segmented cells in a vglut:DsRed brain scan in a second channel and applied transformation matrices for registration to this second channel. Segmented cells were broadly distributed to ensure that distortion measures sampled the entire brain, and cell masks conservatively drawn within the boundaries of the soma. We then compared the morphology of cells after registration (A), to their original shape (B) by calculating the partial Hausdorff distance [30]. Briefly, for every point in a segmented cell mask before registration, we found the minimum distance to a point in the same mask after registration. The Hausdorff distance is the maximum of all such distances, calculated for both A→B and B→A. Because the Hausdorff distance is highly sensitive to cell alignment, and registration displaces cells from their original location, we found the optimal alignment for comparisons using a two-step process. First we aligned the geometric center of each cell in the original and transformed images. Second we searched for the minimal Hausdorff distance across 4940 rigid transformations of the aligned cell, within a 3 x 3 x 3 micron cube, (0.25 micron steps in each dimension). Finally, as Hausdorff distances are sensitive to outliers, we used the 95th percentile distances instead of the maximum Hausdorff distance for all measures [30].

Cell volume. For each segmented cell, we calculated its change in volume as the absolute value of the fractional change in the number of pixels after application of a transformation matrix.

Elongation index. For each pixel in a segmented cell, we found the maximal distance (MD) to any other pixel in the mask. The elongation index for a given cell was the 95th percentile largest value of MD, which we take as an approximation of the diameter of the cell across its longest axis.

Cross correlation. Cross correlation between the tERK-stained reference brain, and registered tERK stains, was performed using the c_correlate function within IDL version 7.0. Correlations were run within eighteen 50 µm-side cube sub-regions of the image volumes that were manually selected to encompass high contrast boundaries and the mean of the 18 values taken as the mean cross correlation (MCC) for each brain in Fig. 3.

Jaccard index. Anti-tERK immunohistochemistry intensely stains tectal neuropil. Thus for measuring the accuracy of registration of the tectal neuropil, we manually segmented the left tectal neuropil area in 6 confocal scans of tERK stained larvae and our reference brain. We applied transformation matrices to these masks, then calculated the Jaccard index as the volume of the intersection between each registered mask (A) and the reference brain (B), divided by the total volume of the union of A and B.

Results

Optimization of ANTs based registration of live vglut2a:dsRed-DsRed image scans
Brain registration in Z-Brain and ZBB used the B-spline elastic transformation in CMTK. Before attempting to co-align Z-Brain and ZBB, we tested an alternate algorithm for brain alignment, the diffeomorphic symmetric normalization (SyN) method in ANTs, because: (1) SyN has been shown to outperform B-spline transformations for deformable image registration in a variety of imaging modalities [31,32], (2) ANTs permits registration using multiple reference channels, potentially allowing the use of multiple complementary expression patterns as references for improved registration fidelity. (3) By calculating forward and reverse transformations simultaneously, SyN transformation matrices are intrinsically symmetric, ensuring that bridging registrations would be unbiased and that we could easily perform reciprocal transformations to register each dataset into the other’s coordinate system.

To calibrate registration parameters, we assessed the alignment precision and distortion of cell morphology after the registration of six representative vglut2a:DsRed scans to the original vglut2a:DsRed reference brain in ZBB (vglut2azbb; file vglut-dsred-ref-01.nii.gz, available from [33], procedure summarized in Fig. 1a). Similar to CMTK we employed a three step registration within ANTs where rigid and affine steps were used to initialize a deformable registration using the SyN diffeomorphic transformation with cross correlation (CC) as the similarity metric. We tested a range of values for each of the SyN parameters as well as the radius of the region used for cross correlation.

To measure registration precision, we visually located 10 point-based landmarks in the vglut2azbb pattern (Additional File 1a-c) and in each of the 6 vglut2a:DsRed confocal scans. We then used the vglut:DsRed channel for registration, and applied the resulting transformation matrix to the landmarks in each of the 6 brains. We measured the distance of each landmark from its location in the vglut2azbb reference brain (Additional File 1a,d). We designated the average of the 10 distances the 'Mean Landmark Distance' (MLD). To assess the amount of distortion in cell shapes produced by the parameter sets, we segmented 107 cells in an unregistered vglut2a:DsRed confocal scan (Additional File 2), and applied each transformation matrix to this set of cell masks. Changes in cell shape were measured using the partial Hausdorff distance for each cell after registration compared to its original shape (see Methods).

Next we plotted the MLD against the Hausdorff distance and located points along the Pareto frontier (Fig. 1b) of these two measures. These points represent potentially optimal transformations, where registration accuracy can only be improved by increasing distortion, or vice versa. To distinguish between these points, we examined two additional measures of distortion: the change in cell volume (Fig. 1c) and maximal elongation (Fig. 1d). Three transformations showed statistically significantly reduced distortion compared to CMTK for both measures, and we selected the one (Fig. 1b, point d) with the greatest
precision for further testing. With this set of parameters (see Table 1, live registration), mean registration error was within the diameter of a single neuron for both ANTs and CMTK (MLD for ANTs 6.7 ± 0.3 μm, for CMTK 7.6 ± 0.4 μm; N = 6 brains, paired t-test p=0.056). However, cell morphology was better preserved using ANTs (Hausdorff Distance for ANTs 2.30 ± 0.14, CMTK 2.37 ± 0.14; N = 107 cells, paired t-test p=0.013), especially within ventral structures such as the hypothalamus and the caudal medulla oblongata (Fig. 1e).

We first calibrated registration conditions by assessing alignment precision for a representative vglut2a:DsRed scan registered to the original vglut2a:DsRed reference brain in ZBB (vglut2a_ZBB-file vglut dsred ref 01.nii.gz, available from [33]). Similar to CMTK we employed a three step registration within ANTs where rigid and affine steps were used to initialize a deformable registration using the SyN diffeomorphic transformation with cross correlation (CC) as the similarity metric. We tested a range of values for each of the SyN parameters as well as the radius of the region used for cross correlation.

While we previously used brain-wide normalized cross correlation (NCC) to evaluate registration [1], correlation within local anatomical regions that contain discrete landmarks has been shown to be a more reliable criterion for quantitatively assessing registration precision [34]. Accordingly, in this work we quantified precision in two ways: by measuring local registration errors both computationally as well as manually. For the computational measure, we identified a set of 10 landmarks within the vglut2a pattern, each within a 50 μm side cube. Landmarks were broadly distributed in the hope of representing diverse brain regions and minimizing the bias for any single structure. We measured the cross-correlation between the corresponding regions in vglut2a_zbb and the registered image, then calculated the mean of the cross-correlation between all regions (MCC; Fig. 1a). For the manual measure of precision, we identified 10 landmarks in the vglut2a_zbb pattern that could be visually recognized (landmark positions are described in Additional File 1), and calculated the mean landmark distance (MLD) from the corresponding points in the registered image as assessed by three blinded experts. We also assessed the results visually to subjectively assess the severity of tissue distortion. Unsurprisingly, similar to our previous work with brain-wide NCC, images with the highest MCCs generally showed more conspicuous tissue distortion — thus although greater precision was achieved with increased deformation, we preferred results where cell shape and axon tract morphology were preserved (Fig. 1b,c). Disregarding parameter combinations that resulted in overt distortion, we identified a set of values (Table 1, live registration) where cell morphology remained intact, but registration precision (MCC) was maximized. With these parameters, although the MCC for vglut2a improved only slightly from 0.79 using CMTK to 0.81 using SyN, cell morphology was noticeably better preserved, especially within ventral structures.
such as the hypothalamus (Fig. 1d). Manual measurement of precision registration confirmed these findings: ANTs registration using values that avoided noticeable morphological distortion showed similar MLDs to images registered using CMTK (Table 2).

We next tested whether these registration parameters also improved precision for the co-aligned transgenic lines. For ZBB, we co-scanned transgene and enhancer trap expression patterns with the v glut 2a: dsRed transgene, allowing us to register each expression pattern to v glut 2a ZBB. We first compared the overlap and morphology of the Mauthner cells from brain scans of three different individuals of transgenic line J1229aGt [23]. Overlap of Mauthner cell bodies was similar for CMTK and ANTs (Fig. 1e, f, g). However, in CMTK registered images, the Mauthner axon was distorted in the caudal medulla, whereas axon morphology was preserved with ANTs. Second, in our previous work, we assessed the precision of CMTK registration using line y339Et by independently scanning two sets of three larvae, producing an average for each set, and visually comparing the result. With CMTK we had noted misalignment of approximately 1 cell diameter in the neuropil of the optic tectum (Fig. 1h, g). This was substantially improved with ANTs, where there was much closer alignment of the two averages (Fig. 1i, h).

For quantification we calculated the cross correlation for 8 landmarks within the y339Et pattern, and found that the mean increased from 0.52 with CMTK to 0.63 with ANTs.

Improved precision of ZBB after registration using ANTs

Based on the improved registration precision and reduced distortion of cell morphology achieved using SyN, we next recompiled ZBB using ANTs for registration to create a more accurate atlas (unprocessed and registered brain images are available from [35]). We used ANTs to register the entire set of 354 brain scans from 109 different transgenic lines that were part of ZBB, then as before, averaged multiple larvae to create a representation of each transgenic line, masked the average stacks to remove expression outside the brain and re-imported the resulting images into our Brain Browser software. We refer to this new recompilation of our atlas as ZBB1.2. Unprocessed and registered brain images are available online [35].

To determine whether ZBB1.2 was a quantitative improvement over ZBB, we calculated a cross-correlation score for each pattern in the browser. To avoid manually defining landmarks for each line, we instead computationally identified three regions inside each pattern with strong expression to serve as landmarks. For each of these regions, we iteratively performed pair-wise cross correlations between all individual brains from the same transgenic line, allowing us to calculate a mean cross-correlation (MCC).
value for each line. We performed this procedure first for brains registered using CMTK, then for the
same set of brains registered using ANTs, allowing us to compare MCCs for the two methods (Fig. 2a).
Overall, the correlations increased slightly from ZBB to ZBB$_{1.2}$ (0.32±0.02 to 0.34±0.02; paired t-test
p=0.15). Although this was not statistically significant, it was instructive to examine instances with large
echanges in mean cross correlation. Line y332Et labels a small set of cells with a salt and pepper pattern in
the right habenula. Here, cross correlation was greater after registration with CMTK (CMTK, 0.50;
ANTS 0.39), due at least in part to greater distortion of cells resulting in increased overlap between
individual fish despite the biological variability (Fig. 2b). In y341Eh, distortion artifacts also appeared to
account for the large increase in MCC obtained with ANTs (CMTK, 0.19; ANTs 0.58). Here, cells in the
caudal hypothalamus had an elongated morphology after registration with CMTK, often stretching
outside the boundaries of the nucleus. Consequently, in this case distortion reduced rather than increased
the cross correlation score (Fig. 2c).

To determine whether ZBB$_{1.2}$ was a quantitative improvement over ZBB, we identified two
conspicuously labeled cells or landmarks in each of 12 transgenic lines from the atlas (Additional File 3).
We marked these positions in each of the three brain scans for each line, then, after registration,
calculated the distance between corresponding points in each pair of brains. The mean of these distances
measures how precisely landmarks are registered across the three brains. We performed this procedure
first for brains registered using CMTK, then for the same set of brains registered using ANTs, allowing us
to compare MLDs for the two methods (Fig. 2a-b). Overall, landmark distances decreased from ZBB to
ZBB$_{1.2}$ (10.8 ± 1.02 μm to 8.1 ± 0.83 μm; N = 24 landmarks, paired t-test p=0.008), indicating that
ZBB$_{1.2}$ has significantly improved precision, and confirming that the new atlas is accurate to
approximately the diameter of a single neuron. The improvement was greatest deeper in the brain (Fig. 2c
; linear regression, N=24, p=0.003) with the largest improvement for the caudal hypothalamus in line
y341, where increased alignment precision was associated with noticeably reduced distortion between the
three brain scans (Fig. 2d).

Additionally, we inspected regions of ZBB$_{1.2}$ where we had noticed poor registration precision or
pronounced cell distortion in the original ZBB. One such area was the dorsal thalamus, where cell
morphology was noticeably perturbed after elastic registration with CMTK, with cell somas stretching
across the midline (Fig. 2d). In ZBB$_{1.2}$ cells retained a rounded morphology with distinct cell clusters
on the left and right sides of the brain (Fig. 2e). Similarly, distortions in cell shape that were apparent in
the caudal hypothalamus in ZBB, were absent in ZBB$_{1.2}$ (Fig. 2f). In the caudalateral medulla, we
previously obtained poor registration, with expression extending to regions outside the neural tube (Fig.
h2i). In ZBB$_{1.2}$, patterns had improved bilateral symmetry and were correctly confined to the neural tube
Finally, we noticed that the posterior commissure was poorly aligned between larvae leading to a defasciculated appearance in ZBB (Fig. 2j2k), whereas this tract had the correct expected tightly bundled appearance in ZBB\textsubscript{1.2} (Fig. 2k2l).

Together, these observations confirm that ZBB\textsubscript{1.2} is a more faithful representation of the transgenic lines. Not only is cell morphology better preserved, but metrics of global registration precision as measured by mean cross-correlation are nevertheless improved from those of compared to the original ZBB atlas.

**Optimization of ANTs registration parameters for fixed tissue**

The Z-Brain atlas was derived by registering brain scans to a single brain that was fixed, permeabilized and immunostained for tERK expression. We therefore presumed-anticipated that tERK would be a useful channel for bridging the two atlases, if we could first successfully register a tERK stained vglut2a:DsRed expressing brain to ZBB\textsubscript{1.2}. Therefore, we fixed and co-stained a transgenic vglut2a:DsRed larva for DsRed-and-tERK, and registered the tERK pattern to ZBB\textsubscript{1.2} using the vglut2a pattern. We used the resulting image as our ZBB tERK reference brain (tERK\textsubscript{ZBB}; file terk-ref-02.nii.gz available from [33]).

In addition to the tERK reference brain, Z-Brain contains an average tERK representation from\textsubscript{1.2} 197 tERK stained larvae, which we thought might serve as a bridge between atlases. During studies on pERK-based activity mapping, we had previously generated a dataset of 167 tERK stained brains and sought to use therefore used these to create an-our own average tERK representation by registering them to tERK\textsubscript{ZBB}. However, during this process, we noticed a high degree of variability between tERK stained brains, most salient notably in either poor labeling of ventral brain structures and or in deformation of the optic tectum neuropil. Immunohistochemistry for tERK proved highly sensitive to staining parameters with the trypsin activity, permeabilization duration, and antigen retrieval having the strongest effects. This variability in fixed tissue was most apparent in the optic tectum, where high trypsin activity tended to disrupt morphology and reduce the volume of the tectal neuropil (Fig. 3a3b,bg). These local distortions were not corrected-resolved by deformable image registration: alignment to tERK\textsubscript{ZBB} with the same parameters optimized for live vglut2a based registration, failed to correct the reduced tectal neuropil volume (Fig. 3e3d,d–e; asterisk) and often created an artifact where the neuropil zone failed to abut the underlying cellular layer labeled by vglut2a expression (Fig. 3e3d,d–e; arrowheads).
We therefore varied the registration parameters that were optimal for live vglut2a registration, to find settings that best rectified the variable tissue morphology following fixation and permeabilization (process summarized in Fig. 3a).

For tERK-optimization of fixed tissue registration-optimization, we used a set of 6 tERK stained brains (including the Z-Brain tERK reference). We iteratively varied parameters for registration to tERK\textsubscript{ZBB} and calculated the mean cross-correlation between each of the aligned tERK stains and tERK\textsubscript{ZBB} (e.g., Fig. 3e,f). Again when visually inspected, we noted a trade-off between the quality of global alignment and local distortion artifacts, with the parameters which yielded the greatest increase in MCC often producing abnormally elongated cell profiles throughout the brain (Fig. 3g). However, visual inspection confirmed that parameters which increased MCC for fixed tissue greatly improved the morphology of the optic tectum neuropil (Fig. 3h,i). We therefore used ANTs with the fixed brain parameters (Table 1, fixed registration) to register 167 tERK stained brains to tERK\textsubscript{ZBB} and generated an average tERK representation comparable to the Z-Brain tERK average (Fig. 3j,k). Iteratively varied parameters for registration to tERK\textsubscript{ZBB} and assessed registration fidelity. For measuring precision, we were not able to identify unambiguous landmarks within the optic tectum, so we instead calculated the cross-correlation between each of the aligned tERK stains and tERK\textsubscript{ZBB} within small volumes, including parts of the tectum (Fig. 3f,g). To verify that the 'fixed brain' parameters that yielded the greatest cross correlation did in fact improve registration within the tectum, we manually segmented the tectal neuropil in the same 6 brains, applied the transformation matrix to each mask, and calculated the Jaccard index for overlap with the segmented neuropil in tERK\textsubscript{ZBB}. Parameters for fixed brain registration produced a significant increase in overlap, compared to the live brain parameters (Fig. 3h,i) and visual inspection confirmed that the morphology of the optic tectum neuropil after registration was greatly improved (Fig. 3j,k). We therefore used ANTs with the fixed brain parameters (Table 1, fixed registration) to register our 167 tERK stained brains to tERK\textsubscript{ZBB}, and generated an average tERK representation comparable to the 197 tERK average in Z-Brain (Fig. 3l,m).

Inter-atlas registration using multi-channel diffeomorphic transformation

By chance, both Z-Brain and ZBB incorporated seven additional eight gene or transgene-expression patterns that we judged were sufficiently similar to act either as templates for bridging the datasets and/or to provide metrics for assessing the precision of a bridging registration (Table 32, Additional File 4). For example, vglut2a\textsubscript{ZBB} is a confocal scan of DsRed in a single larva from transgenic line TgBAC(slcl7a6b:loxP-DsRed-loxP-GFP)mns14, whereas Z-Brain includes Tg(slcl7a6:EGFP)zf139. In both cases, reporter expression is regulated by the same bacterial artificial chromosome [15,20]. Crossing...
these two lines allowed us to scan DsRed and EGFP in the same larva and confirm that the patterns were largely congruous, potentially allowing us to use vglut2a expression to bridge the two atlases. Likewise, the expression patterns of tERK, elavl3, isl2b, vmat2 in Z-Brain and ZBB appeared sufficiently similar to provide templates for atlas co-registration.

We used seven expression patterns to evaluate registration precision using cross correlation: vglut2a, isl2b, vmat2, elavl3, isl1, gad1b and glyT2. For each pattern we identified a set of 5-18 landmarks that were widely-distributed to represent diverse brain regions. For each landmark, we measured the cross-correlation between the corresponding volumes in ZBB and Z-Brain. We then calculated the mean of all cross-correlation (MCC) values for landmarks associated with a given expression pattern. We used two measures of registration precision. The first metric ($M_1$) was the mean of the MCCs for isl1, gad1b and glyT2 expression patterns in ZBB and in Z-Brain after registration to ZBB. These three expression patterns do not provide sufficient coverage across all brain regions to use for registration, but served as independent channels to estimate registration precision. However, as these patterns are relatively sparse, they do not comprehensively assess precision across all brain regions. To provide a global measure of precision, we therefore also used a second metric ($M_2$) that was the mean of all seven MCCs: those in $M_1$ plus four of the patterns used as references for registration—vglut2a, tERK, isl2b and vmat2. Although $M_2$ uses expression patterns that together provide good coverage for the entire brain, we expected that the four patterns that were also used to guide the deformable registration, would artificially inflate the MCC.

We first used CMTK to register Z-Brain to ZBB. Maximal $M_1$- and $M_2$-scores were obtained using the average vglut2a pattern as the reference (Fig. 4a). We therefore registered all images in Z-Brain to ZBB using the vglut2a average in each dataset as the reference channel. We observed severe tissue distortions in several brain regions, with noticeable flattening of the torus longitudinalis and gross tissue distortions, particularly in ventral brain regions (Fig. 4b,c; ZBrain-CMTK).

Next, for comparison, we used the ANTs SyN algorithm to register the atlases. Ideally, patterns for registration should include information throughout the brain. Because ANTs can use multiple concurrent reference channels to derive an optimal transformation matrix, we speculated that the best possible transformation would be achieved by a combination of channels with complementary information. We therefore produced an inter-atlas transformation matrix using every combination of the elavl3, isl2b, vglut2a, vmat2, tERK single (tERK single brain) and tERK average (tERK average brain) patterns as references. As Z-Brain used fixed samples, we used the registration parameters previously optimized for the greater variability present in fixed tissue. Multi-channel registration significantly improved $M_1$- and $M_2$-values...
compared to any single-channel alone and to transformations obtained using CMTK. The registration obtained with *vglut2a*, *tERK*<sub>SER</sub>-*vmat2* and *isl2b* gave the highest M<sub>2</sub> value and an M<sub>1</sub> score within 1% of the highest scoring combination (Fig. 4a). Moreover, the overt tissue distortions noted after elastic registration with CMTK were far less salient using these parameters (Fig. 4b,c; ZBrain-SyN). This conclusion was supported when we assessed registration precision by visually locating landmarks in the *vglut2a* pattern after registration with CMTK, or multi-channel ANTs registration. After calculating the distance from the same points in the *vglut2a* pattern we found the multi-channel ANTs registered points were on average 9.9 μm away from the reference points, compared to 17.9 μm for CMTK (Table 4). We therefore applied the transformation matrix obtained with this set of channels to the database of gene expression patterns in Z-Brain to align them to ZBB<sub>1-2</sub>

The precision of the inter-atlas registration is apparent when comparing the location of cells that are present in both datasets, such as those labeled by *Pet1:GFP*. The Z-Brain transformed pattern closely matches the transgene expression pattern in ZBB<sub>1-2</sub> within the superior raphe (Fig. 4d — note however that unexpectedly, the line in ZBB<sub>1-2</sub> also labels a set of more rostral cells not apparent in Z-Brain). Both atlases also include lines labeling the Mauthner cells. After registration, Mauthner cells in the atlases substantially overlapped, although they were several microns more medially positioned in ZBB<sub>1-2</sub> (Fig. 4e). Similarly, we used the inverse of the transformation generated by SyN to register ZBB<sub>1-2</sub> to the Z-Brain coordinate system. As expected, expression in the DAT:*GFP* line in ZBB<sub>1-2</sub> overlapped well with the tyrosine hydroxylase stain from Z-Brain in the pretectum (Fig. 4f), although again, the ZBB<sub>1-2</sub> pattern was slightly more medial than in Z-Brain. More caudally, the *glyT2:*GFP transgenic line labels glycinergic neurons in longitudinal columns in the medulla oblongata [36]. These columns were closely aligned after ZBB<sub>1-2</sub> was registered to Z-Brain (Fig. 4g). Although best practice is to align directly to either ZBB or Z-Brain, because many researchers will have already registered data sets to either ZBB or Z-Brain, or for cases where it may not be possible to directly register a dataset, we have provided transformation matrices and detailed instructions to quickly re-align datasets to either of the coordinate systems ([37]; instructions are provided in Additional File 2).

Taking advantage of the ability of ANTs to use of multiple reference channels concurrently, we compared the effect of combinatorial use of complementary reference channels for inter-atlas registration (process summarized in Fig. 4a). We used seven expression patterns to evaluate registration precision: *vglut2a*, *isl2b*, *vmat2*, *tERK*, *isl1*, *gad1b* and *glyT2*. For each pattern we identified a set of 4-10 point-based landmarks that could be identified in corresponding ZBB and Z-Brain images and that were widely distributed to represent diverse brain regions (total of 41 landmarks ; Additional File 5). We marked these points in each set of images, registered Z-Brain images to ZBB<sub>1-2</sub> images, measured the distance between
cognate landmarks and calculated the mean landmark distance for each of the seven expression patterns.

We used two summary measures of registration precision. The first metric ($M_1$) was the mean of MLDs for the three patterns that were not used to drive registration ($isl1$, $gad1b$ and $glyT2$). Although these channels measure precision independent of the patterns for atlas registration, they are relatively sparse and do not assess precision across the whole brain. Thus, to provide a global measure of precision, we also used a second metric ($M_2$) that was the mean of all seven MLDs; those in $M_1$ plus four of the patterns used as references for registration — $vglut2a$, tERK, $isl2b$ and $vmat2$.

Using CMTK, minimal $M_1$ and $M_2$ scores were obtained using the average $vmat2$ pattern as the reference (Fig. 4b; mean MLD for 41 landmarks $14.9 \pm 1.3$ μm). We therefore registered all images in Z-Brain to ZBB using the $vmat2$ average in each dataset as the reference channel. We observed severe tissue distortions in several brain regions, with noticeable flattening of the torus longitudinalis as well as gross tissue distortions, particularly in ventral brain regions (Fig. 4c,d; ZBrain-CMTK). Next we used the ANT’s SyN algorithm to register the atlases. Ideally, patterns for registration should include information throughout the brain. Because ANT’s can use multiple concurrent reference channels to derive an optimal transformation matrix, we speculated that the best possible transformation would be achieved by a combination of channels with complementary information. We therefore produced an inter-atlas transformation matrix for every combination of the $elavl3$, $isl2b$, $vglut2a\_AV$ ($vglut2a$ average brain), $vmat2$, tERK\_ZBB (tERK single brain) and tERK\_AV (tERK average brain) patterns as references. Because Z-Brain used fixed samples, we used the registration parameters optimized for the greater variability present in fixed tissue. Multi-channel registration significantly reduced $M_1$ and $M_2$ values compared to any single channel alone and to transformations obtained using CMTK. The registration obtained with $vglut2a$, tERK\_ZBB, $vmat2$ and $isl2b$ gave the lowest global metric ($M_2$) value and an $M_1$ score within 10% of the lowest scoring combination (Fig. 4b). With these parameters, the MLD was $9.1 \pm 0.8$ μm ($N=41$ landmarks) and the overt tissue distortions noted after elastic registration were far less salient (Fig. 4c,d; ZBrain-SyN). We therefore applied the transformation matrix obtained with this set of channels to the database of gene expression patterns in Z-Brain to align them to ZBB\_1.2, and used the inverse of the transformation generated by SyN to register ZBB\_1.2 to the Z-Brain coordinate system. We imported all Z-Brain expression patterns not previously represented in the database into ZBB\_1.2, producing a total of 133 expression patterns.

The accuracy of the inter-atlas registration is evident when comparing the location of cells that are present in both datasets, such as those labeled by $pet1\_GFP$. The Z-Brain transformed pattern closely matches the transgene expression pattern in ZBB\_1.2 within the superior raphe (Fig. 4e — note however that
unexpectedly, the line in \( \text{ZBB}_1.2 \) also labels a set of more rostral cells not apparent in \( \text{Z-Brain} \). Both atlases also include lines labeling the Mauthner cells. After registration, Mauthner cells in the atlases substantially overlapped, although they were several microns more medially positioned in \( \text{ZBB}_1.2 \) (Fig. 4f). Expression in the \( \text{DAT}:GFP \) line in \( \text{ZBB}_1.2 \) overlapped well with the tyrosine hydroxylase stain from \( \text{Z-Brain} \) in the pretectum (Fig. 4g), although again, the \( \text{ZBB}_1.2 \) pattern was slightly more medial than in \( \text{Z-Brain} \). Caudally, the \( \text{glyT2}:GFP \) transgenic line labels glycinergic neurons in longitudinal columns in the medulla oblongata [36]. These columns were closely aligned after \( \text{ZBB}_1.2 \) was registered to \( \text{Z-Brain} \) (Fig. 4h).

Although best practice is to align directly to either \( \text{ZBB} \) or \( \text{Z-Brain} \), because many researchers will have already registered data sets to either \( \text{ZBB} \) or \( \text{Z-Brain} \), or for cases where it may not be possible to directly register a dataset, we have provided transformation matrixes and detailed instructions to quickly re-align datasets to either of the coordinate systems ([37]; Additional File 6).

**Neuroanatomical visualization**

\( \text{Z-Brain} \) includes 294 masks that represent anatomically defined brain regions or discrete clusters of cells present in transgenic lines. We selected 113 of these masks that delineate neuroanatomical regions and transformed them into the \( \text{ZBB}_1.2 \) coordinate system. We had previously defined a small number of our own anatomical masks by thresholding clusters of neuronal cell bodies located in well-defined brain regions. In contrast, however, the \( \text{Z-Brain} \) masks are more comprehensive, have smoother boundaries and include both the cell bodies and neuropil for a given region (Fig. 4h–k). We therefore imported the \( \text{Z-Brain} \) masks into \( \text{ZBB}_1.2 \), replacing most of our existing masks. We also modified the Brain Browser software to automatically report the neuroanatomical identity of a selected pixel, or to display the boundaries of the region encompassing a selected point. The updated software and rebuilt database in \( \text{ZBB}_1.2 \) can be downloaded from our website [38].

Finally, as the Zebrafish Brain Browser's strength is primarily in two dimensions (i.e., the visualization of horizontal, transverse, and sagittal slices through the brain), we decided to develop interactive tools to better facilitate three dimensional 3D exploration. The use of 3D graphics to represent complex structure can also provide a more intuitive sensory experience that avoids cognitive bias or misinterpretation inadvertently introduced by sometimes largely arbitrary two dimensional reductions [39,40]. By taking advantage of stereoscopy and vestibular-enhanced parallax (head tracking), the more immersive and holistic experience of Virtual Reality (VR) can also significantly improve performance of basic tasks like
searching and making comparisons [41,42]. We therefore implemented our Zebrafish Brain Browser in both an open Web3D platform (X3D) and a custom game engine (Unity). First, we converted masks representing anatomical regions to meshes and built a Web3D interface using X3D to inspect the spatial relationship between different brain regions (Fig. 5a,b), available online [43]. Users can navigate within the brain using any web browser, rotating and zooming into brain regions to better interrogate larval neuroanatomy. Second, using the Unity platform we wrote a VR app to view the brain and neuroanatomical regions. By running the app on a cell phone, and inserting it into an inexpensive Google cardboard viewer, users can 'walk into' the brain, and see from the inside the inter-relationship between neuroanatomical domains (Fig. 5c,d), available for download [44].

Discussion

Digitized data-derived brain atlases provide an opportunity to continuously integrate new information and iteratively improve data accuracy within a common spatial framework. Thus, as methods evolve and technology improves, new insights can be easily added to existing data to provide an increasingly rich view of brain structure and function. Because the entire larval zebrafish brain can be rapidly imaged at cellular resolution, it is possible to envisage an atlas that combines detailed information on cell type (including gene expression and morphology), connectivity and activity under a variety of different physiological conditions. At present, biological variability presents an obstacle, as brain regions contain multiple intermingled cell types that are not positioned in precisely the same manner between larvae. To circumvent-compensate for this in the existing zebrafish brain atlases, multiple individuals of a given line are sampled and averaged to generate a representative expression pattern. Current atlases are thus essentially heat maps of gene expression or activity. Despite this spatial ambiguity, aggregating information from different sources into the same spatial framework still provides valuable indicators of cell type, gene co-expression, and neural activity under defined conditions.

Ideally different atlas projects might use the same reference brain, however in practice the choice of a reference is often dictated by study-specific experimental requirements. For example, despite the deformations introduced by fixation and permeabilization, a fixed brain is essential for activity mapping using pERK immunohistochemistry. In contrast, we were able to take advantage of the optical transparency of larvae to rapidly scan and register several hundred individuals representing more than 100 different transgenic lines. For our purposes, the TgBAC(slc17a6b:loxP-DsRed-loxP-GFP)nns14 line was ideal, because through Cre injection, we generated a vglut2a:GFP line with an almost identical pattern, allowing us to co-register lines with either GFP or RFP fluorescence. However, we have also used pan-
neuronal Cerulean or mCardinal as a reference channel when useful information on transgene expression. Our work now demonstrates that it is feasible to contribute to community efforts at building an integrated map of brain structure, expression and activity, while allowing reference image selection to be guided by technical considerations.

One caveat to this conclusion is that deformable image registration can easily introduce artifacts into cell morphology if parameters are not carefully monitored and constrained. Indeed, a special challenge for brain registration in zebrafish is preserving the local morphology of neuronal cell bodies and axons, while permitting sufficient deformation to correct for biological differences and changes in brain structure arising from tissue fixation and permeabilization. Thus, while B-spline registration with CMTK produced acceptable inter-atlas alignment, it also introduced noticeable distortions into local brain structure that affected neuronal cell morphology. Such artifacts were particularly severe in ventral brain regions such as the caudal hypothalamus, and may therefore be due to differences in ventral signal intensity between the datasets. In ZBB, in order to compensate for the increase in light diffraction with tissue depth, we systematically increased laser intensity with confocal scan progression (z-compensation). As a result, the Z-Brain and ZBB datasets are comparable in dorsal brain regions, but there is a noticeable discrepancy ventrally which may account for the loss of registration fidelity. Alternatively, although z-compensation partially corrects for reduced fluorescent intensity, there is a noticeable drop-off in image resolution in ventral regions; the resulting loss of information may lead to lower quality registration. Registration algorithms that allow parameters to vary by depth may ameliorate the effects of these physical imaging constraints.

Nevertheless, the symmetrical diffeomorphic transformation in ANTs provides a satisfactory solution to these problems. For live tissue, we found parameters that allowed the ANTs SyN transform to achieve similar or better registration precision than previously achieved using CMTK, while significantly reducing overt distortions in tissue structure and neuronal cell morphology. In our hands, permeabilization of fixed tissue tended to produce variable changes in neuropil structure which was most salient in the optic tectum. Specifically, neuropil volume was diminished when fresh aliquots of trypsin were used for extended durations. These artifacts can be minimized by stringent oversight of reagent viridity. However, by calibrating SyN parameters to permit larger deformations, we were able to accommodate the variability introduced in tissue processing.

Currently, the main limitations for use of the SyN registration algorithm in ANTs are the large memory demands (73 GB for a single channel registration) and long computational times (3-5 hours for a single
channel using 24 cores) required for registration of images with a resolution sufficient for the brain-wide visualization of neuronal morphology (e.g., 1000 x 600 x 350 pixels). For multi-channel registrations, memory demands and computation time were even greater: 106 GB for 6 channels taking over 16 hours on 24 cores. However, our present ANTs SyN parameters likely can be further optimized to reduce these demands. For instance, our parameters currently include 10 iterations of transformation matrix optimization at full image resolution. From our experience, these full resolution registration cycles do not significantly improve cross correlation scores but greatly increase computation time.

Thus, computation time may be reduced by adjusting registration resolution as well as other parameters without adversely affecting registration quality. Although computational resources did not present a bottleneck for registering a small number of samples, this increase in the demands of a single registration made it difficult to optimize registration parameters as extensively as we had done previously with CMTK [1]. For example, during our initial effort to optimize registration parameters for live vglut2a expression, we used a single representative example rather than assessing parameters for a set of several independent scans. By reducing computation time, we would be able to explore more comprehensively the parameter space available with SyN and evaluate alternative diffeomorphic transforms available with ANTs that may provide still better registration fidelity.

An obstacle to systematically calibrating registration parameters was finding a suitable metric to quantitatively evaluate registration precision. This is a recognized problem, and it is not clear that a general solution exists [34]. Here, we primarily assessed precision by measuring the distance between visually-located landmarks in the reference brain, and registered images. However, this method has two drawbacks: (1) it relies on the accuracy with which these landmarks are located, and (2) at least for our sample set, a relatively limited set of landmarks could reliably be identified. We obtained similar results when we assessed precision using cross-correlation within localized image neighborhoods that included relatively-high contrast internal image boundaries (data not shown). However, in registering live vglut2a:DsRed image stacks, we found that the highest scoring transformations achieved the trade-off between accurate global brain alignment and at the expense of biologically plausible cell morphology. Thus we also used a set of measures to assess changes in the morphology of manually segmented cells (Hausdorff distance, elongation index and cell volume). Finally, we also Therefore, it was essential to visually compare the output of every transformation and make subjective judgments about registration quality. This was difficult, because distortions, when present, tended to be variable in different parts of the image, thus requiring the entire image stack produced by each transformation to be scrutinized to select optimal parameter settings.
Nevertheless, this study demonstrates that the ANTs diffeomorphic symmetric normalization algorithm (SyN) improves upon elastic registration for precise registration of whole brain images in larval zebrafish and is markedly better at preserving neuronal cell morphology. By systematically testing SyN registration parameters for registering images acquired using live scans, we improved the ZBB atlas. Then, after calibrating registration parameters for fixed tissue and using multi-channel optimization, we were able to align the Z-Brain atlas into the ZBB coordinate space, and vice-versa, achieving co-registration accuracy to approximately the diameter of a single neuron. We believe that integrating the information present in each of these atlases produces a richer framework for future studies of structural and functional relationships within the nervous system. Large digital datasets such as those present in brain atlases can be used for many types of bioinformatic analysis. Z-Brain and ZBB already include software that can be used to explore the larval zebrafish brain, and we hope that by integrating these datasets into a single coordinate system, we will help to stimulate the development of additional computational tools and methods for querying this information.

### Availability of supporting data

All individual brain scans, both before and after registration to a ZBB reference brain, are available in the *GigaScience* repository, GigaDB [35]. The GigaDB repository also includes the set of reference brains used for ZBB [33] and the transformation matrices used to convert between ZBB and Z-Brain coordinate systems [37].

### Abbreviations

- ac, anterior commissure
- Ce, cerebellum
- DT, Thalamus
- GT, Griseum tectale
- Ha, Habenula
- Hc, Hypothalamus caudal zone
- Hi, Hypothalamus intermediate zone
- MO, Medulla oblongata
- NXm, Vagus motor neurons
- OB, Olfactory bulb
- OE, Olfactory epithelium
- ON, Olfactory nerve
- IO, Inferior olive
- LC, Locus coeruleus
- MN, Mauthner neuron
- MO, Medulla oblongata
- Pal, Pallium
Competing Interests

The authors declare that they have no competing interests.

Authors' Contribution

GDM and HAB conceived the experiments. GDM, KMT, EJH and HAB optimized ANTS for zebrafish brain registration. GDM, MB and AKG contributed confocal brain scans and generated meshes. NFP developed the X3D/HTML5 based browser. DDN developed the Unity VR browser. GDM and HAB wrote the manuscript. All authors approved the final manuscript.

Acknowledgements

This work was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD) and utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD [45]. We thank Owen Randlett for valuable discussion and help checking the correspondence of vglut2a expression patterns. We are grateful to Sinisa Pajevic (NIH/CIT) for advice on computational procedures and to M. Okan Irfanglu and Neda Sadeghi (NICHD) for guidance in optimizing ANTs parameters.

References

1. Marquart GD, Tabor KM, Brown M, et al. A 3D Searchable Database of Transgenic Zebrafish Gal4 and Cre Lines for Functional Neuroanatomy Studies. Front Neural Circuits 2015; 9:78.

2. Randlett O, Wee CL, Naumann EA, et al. Whole-brain activity mapping onto a zebrafish brain atlas. Nat Methods 2015; 12:1039–1046.
3. Ronneberger O, Liu K, Rath M, et al. ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains. *Nat Methods* 2012; 9:735–42.

4. Rath M, Nitschke R, Filippi A, Ronneberger O, Driever W. Generation of high quality multi-view confocal 3D datasets of zebrafish larval brains suitable for analysis using Virtual Brain Explorer (ViBE-Z) software. *Protoc Exch* 2012.

5. Portugues R, Feierstein CE, Engert F, Orger MB. Whole-Brain Activity Maps Reveal Stereotyped, Distributed Networks for Visuomotor Behavior. *Neuron* 2014; 81:1328–1343.

6. Rohlfing T, Maurer CR Jr. Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. *IEEE Trans Inf Technol Biomed* 2003; 7:16–25.

7. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* 2008; 12:26–41.

8. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* 2011; 54:2033–2044.

9. Higashijima S, Hotta Y, Okamoto H. Visualization of cranial motor neurons in live transgenic zebrafish expressing green fluorescent protein under the control of the islet-1 promoter/enhancer. *J Neurosci* 2000; 20:206–18.

10. Wen L, Wei W, Gu W, et al. Visualization of monoaminergic neurons and neurotoxicity of MPTP in live transgenic zebrafish. *Dev Biol* 2008; 314:84–92.

11. Xi Y, Yu M, Godoy R, Hatch G, Poitras L, Ekker M. Transgenic zebrafish expressing green fluorescent protein in dopaminergic neurons of the ventral diencephalon. *Dev Dyn* 2011; 240:2539–47.

12. Lillesaar C, Stigloher C, Tannhauser B, Wullimann MF, Bally-Cuif L. Axonal projections originating from raphe serotonergic neurons in the developing and adult zebrafish, Danio rerio, using transgenics to visualize raphe-specific pet1 expression. *J Comp Neurol* 2009; 512:158–82.

13. Tabor KM, Bergeron SA, Horstick EJ, et al. Direct activation of the Mauthner cell by electric field pulses drives ultra-rapid escape responses. *J Neurophysiol* 2014; 112:834–844.

14. Scott EK, Baier H. The cellular architecture of the larval zebrafish tectum, as revealed by gal4 enhancer trap lines. *Front Neural Circuits* 2009; 3:13.

15. Satou C, Kimura Y, Hirata H, Suster ML, Kawakami K, Higashijima S. Transgenic tools to characterize neuronal properties of discrete populations of zebrafish neurons. *Development* 2013; 140:3927–31.

16. McLean DL, Fan J, Higashijima S, Hale ME, Fetcho JR. A topographic map of recruitment in spinal cord. *Nature* 2007; 446:71–5.

17. Pittman AJ, Law MY, Chien CB. Pathfinding in a large vertebrate axon tract: isotypic interactions guide retinotectal axons at multiple choice points. *Development* 2008; 135:2865–71.
18. Yokogawa T, Hannan MC, Burgess HA. The dorsal raphe modulates sensory responsiveness during arousal in zebrafish. *J Neurosci* 2012; **32**:15205–15215.

19. Satou C, Kimura Y, Higashijima S. Generation of multiple classes of V0 neurons in zebrafish spinal cord: progenitor heterogeneity and temporal control of neuronal diversity. *J Neurosci* 2012; **32**:1771–83.

20. Bae YK, Kani S, Shimizu T, *et al.* Anatomy of zebrafish cerebellum and screen for mutations affecting its development. *Dev Biol* 2009; **330**:406–26.

21. Fosque BF, Sun Y, Dana H, *et al.* Labeling of active neural circuits in vivo with designed calcium integrators. *Science* 2015; **347**:755–760.

22. Nechiporuk A, Linbo T, Poss KD, Raible DW. Specification of epibranchial placodes in zebrafish. *Development* 2007; **134**:611–623.

23. Burgess HA, Johnson SL, Granato M. Unidirectional startle responses and disrupted left-right co-ordination of motor behaviors in robo3 mutant zebrafish. *Genes Brain Behav* 2009; **8**:500–11.

24. Inoue D, Wittbrodt J. One for All—A Highly Efficient and Versatile Method for Fluorescent Immunostaining in Fish Embryos. *PLOS ONE* 2011; **6**:e19713.

25. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012; **9**:671–675.

26. Williams G. *NIfti Input/Output*. https://imagej.nih.gov/ij/plugins/nifti.html (last accessed, 22 March 2017)

27. Heisenberg M, Schmid B. *IntSeg 3D*. http://3dviewer.neurofly.de/IntSeg_3D (last accessed, 22 March 2017)

28. Web3D Consortium. *Open Standards for Real-Time 3D Communication*. http://www.web3d.org/ (last accessed, 22 March 2017)

29. Vicomtech-IK4. *AtlasConversionScripts*. https://github.com/VolumeRC/AtlasConversionScripts (last accessed, 22 March 2017)

30. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Mach Intell* 1993; **15**:850–863.

31. Klein A, Andersson J, Ardekani BA, *et al.* Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* 2009; **46**:786–802.

32. Murphy K, Ginneken B van, Reinhardt JM, *et al.* Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge. *IEEE Trans Med Imaging* 2011; **30**:1901–1920.

33. *ZBB Reference Brains*. http://helix.nih.gov/~BurgessLab/reference_brains.zip (last accessed, 22 March 2017) To be replaced with GigaDB DOI upon publication

34. Rohlfing T. Image Similarity and Tissue Overlaps as Surrogates for Image Registration Accuracy: Widely Used but Unreliable. *IEEE Trans Med Imaging* 2012; **31**:153–163.

35. *ZBB Raw and Registered Brain Images* GigaDB DOI to be provided upon publication
36. Kinkhabwala A, Riley M, Koyama M, et al. A structural and functional ground plan for neurons in the hindbrain of zebrafish. *Proc Natl Acad Sci U A* 2011; **108**:1164–9.

37. ZBB-ZBrain Transformation Matrices. https://helix.nih.gov/~BurgessLab/zbb-zbrain.zip (last accessed, 22 March 2017) To be replaced with GigaDB DOI upon publication

38. Zebrfish Brain Browser. https://science.nichd.nih.gov/confluence/display/burgess/Brain+Browser (last accessed, 22 March 2017)

39. Ota D, Loftin B, Saito T, Lea R, Keller J. Virtual reality in surgical education. *Comput Biol Med* 1995; **25**:127–137.

40. Ware C, Franck G. Viewing a graph in a virtual reality display is three times as good as a 2D diagram. In: *Proceedings of 1994 IEEE Symposium on Visual Languages*.; 1994:182–183.

41. Bowman DA, McMahan RP. Virtual Reality: How Much Immersion Is Enough? *Computer* 2007; **40**:36–43.

42. Henry JAG, Polys NF. The effects of immersion and navigation on the acquisition of spatial knowledge of abstract data networks. *Procedia Comput Sci* 2010; **1**:1737–1746.

43. Zebrfish Anatomy Explorer. http://metagrid2.sv.vt.edu/~npolys/NIH/Burgess/ZebraFishBrowser/ZebrafishBrowser.html (last accessed, 22 March 2017)

44. Virtual reality brain browser. https://science.nichd.nih.gov/confluence/display/burgess/Brain+Browser#BrainBrowser-vrfishbrain (last accessed, 22 March 2017)

45. High Performance Computing at the NIH. 2016. https://hpc.nih.gov (last accessed, 22 March 2017)

**Figure legends**

**Table 1. ANTs command parameters for image registration**

| Command Parameters |
|--------------------|
| ImageRegistration -Raff.JSON -aff -n 2 -s 200 -o temp.nii.gz | **Table 2. Registration precision for CMTK, ANTs (using parameters for optimal mean cross correlation), and ANTs (using parameters that minimize distortion), as measured by manually locating landmarks within the image.**

| Registration Method | Mean Distance (microns) |
|---------------------|-------------------------|
| CMTK                | 45.2 ± 3.2              |
| ANTs (optimal mean correlation) | 38.5 ± 2.7             |
| ANTs (distortion minimization) | 41.3 ± 2.8             |

Values are distances in microns from the corresponding landmarks in the reference brain. Three experts located, blind to the identity of the samples, located the landmarks in each registered image. The distance shown is the mean of the three distances from the same landmarks in the reference brain. To assess reproducibility of locating landmarks in the reference brain, the same three people also located the landmarks in the reference brain (fourth column) demonstrating that these landmarks can be located by experts to within 5 microns.
Table 3. Brain images in ZBB and Z-Brain that were used as templates for registration and/or for measurement of registration precision.

Table 4. Precision of ZBB and Z-brain co-alignment after using CMTK with tERK as the registration channel, CMTK with vglut:dsRed as the registration channel or multi-channel registration using ANTs.

Distances (in microns) were measured as in Table 2, by three experts who were blind to the identity of the samples.

Figure 1. Optimization of parameters for registration of live brain scans using ANTs

(a) Overview of parameter optimization for live brain scans using ANTs. A calibration set of 6 vglut2a:DsRed confocal stacks with 10 point-based landmarks and 107 cell masks were registered to the vglut2a/ZBB reference with the same 10 point-based landmarks defined (left). MLDs for landmarks and Hausdorff distance for transformed cell masks compared to their originals were measured for each parameter set (middle). Optimal parameters selected from these metrics (b-d) were used to re-register all lines generating ZBB12 where MLD was measured from 2 additional landmarks in each of 12 co-aligned patterns (right). (a) Dorsal maximum projections through the twelve 50 x 50 x 50 μm cubes used to calculate the mean cross correlation (MCC) for vglut2a expression patterns. Top row shows projections for the reference image, vglut2a/ZBB, and bottom row shows projections for a representative vglut2a:dsRed brain that was registered to the reference brain using CMTK. Correlation coefficients are indicated in the bottom row. For this example, the MCC is the mean of the indicated values, 0.73.

(b) Hausdorff distance for cell shape comparison plotted against MLD for 68 sets of parameters tested using ANTs (grey and blue circles) and after registration using CMTK (orange). Blue circles labeled a-f indicate the Pareto frontier.

(c) Mean absolute change in cell volume (as a fraction of the original volume) produced by transformations resulting from parameter sets a-f and CMTK in (b). * p < 0.05, compared to CMTK.

(d) Mean elongation index for cells after registration using parameter sets a-f and CMTK in (b). Dashed line shows index for cells before registration — all transformations produced a significant increase in compared to the untransformed cells. * p < 0.05, compared to CMTK.

(b-de) Comparison of a single plane Horizontal section through the medulla oblongata in vglut2a/ZBB, and of the a representative vglut2a:dsRed brain after registration using CMTK or ANTs, with parameters that produced the largest mean cross correlation score (0.85; MCC optimal), and ANTs with parameters where visual inspection showed cell morphology was best preserved (Visual optimal. MCC was 0.81).
Slices are through the optic tectum (b), medulla oblongata (c) and hypothalamus (d). Distortion artifacts introduced by CMTK in the hypothalamus (arrowhead) as well as poor cell morphology with CMTK and ANTs MCC-optimal (arrow) are indicated (arrow). Scale bar 50 µm.

Comparison of a single horizontal section plane in J1229aGt showing expression of GFP in the Mauthner cell and axon (arrowheads) for three individual larvae (pseudo-colored red, green and blue).

Registration was performed with CMTK (ef) or ANTs (fg). Scale bar 100 µm.

Single coronal-plane Transverse section through the optic tectum in two separate average brain images (colored green and magenta) for y393Et. For each brain image, we independently scanned three individual brains and registered them using CMTK (gh) or ANTs (hi). Scale bar 100 µm.

Figure 2. Improved precision of transgene representations in ZBB1,2

(a) Mean landmark distances for 24 landmarks. Mean of cross-correlation values derived from all pairwise comparisons of individual brains for each transgenic line in ZBB, after registration with CMTK and ANTs. Dotted line indicates 1:1 ratio.

(b) Horizontal slice through the right habenula in y332Et, showing three individual brain scans after registration with CMTK (top row), and the same slices pseudo-colored (red, green, blue) and superimposed. Bottom row shows the equivalent after registration using ANTs.

(b) Boxplot of data in (a). * paired t-test, N=12 lines, p = 0.019

(c) Difference in MLD between ANTs and CMTK plotted against distance from the dorsal-most point in the brain.

(ed) Horizontal slice-section through the caudal hypothalamus of three individual y341Et larvae as well as their pseudo-colored superimposition following registration with CMTK (top row) or ANTs (bottom row).

(de,ef) Horizontal slice-section through the thalamus showing the averaged representation of enhancer trap line y304Et, where individual brains were registered with CMTK for ZBB (de), or with ANTs for ZBB1,2 (ef). Arrow indicates neurons that are artificially elongated across the midline. Scale bar 100 µm.

(fg,gh) Coronal slice-Transverse section through the caudal hypothalamus showing the average enhancer trap line y269Et brain registered with CMTK (dg) and-or with ANTs (eh). Arrow shows distortion of cells causing the caudal hypothalamus to appear dorsally elongated. Scale bar 50 µm.

(hi,ij) Coronal slice-Transverse section through the medulla oblongata showing the average phox2b:GFP brain with CMTK (fi) and-or ANTs (gi). Scale bar 50 µm.

(jk,kl) Horizontal projection through the posterior commissure (arrow) for the average y351Et brain obtained with CMTK (jk) or ANTs (kl). Scale bar 100 µm.
Figure 3. Optimization of ANTs registration parameters for fixed tissue

(a) Overview of parameter optimization for fixed brain scans using ANTs. A calibration set of 6 tERK confocal stacks with segmentations of the tectal neuropil were registered to \( tERK_{ZBB} \), a tERK and \( vglut2a:DsRed \) confocal scan previously aligned to the \( vglut2a_{ZBB} \) reference (left). MCCs were calculated between eighteen 50 \( \mu \)m-side cube high-contrast sub-regions in the calibration set and in the \( tERK_{ZBB} \) reference to identify parameters that maximized MCC (f,g) and improved the Jaccard index of tectal neuropil segmentation (h) while compensating for fixation artifacts (c,e,k) (middle). These optimized ANTs parameters allow for the accurate registration of fixed tissue and the generation of a tERK average reference (\( tERK_{AV} \)) useful for bridging live and fixed tissue registrations (right).

(ab,be) Horizontal section through the optic tectum after \( tERK \) immunostaining for tERK (red) and DsRed in \( vglut2a:DsRed \) (green) larvae, using diluted (ab, sample A) or fresh trypsin (be, sample B). Asterisk indicates missing area of tectal neuropil due to permeabilization artifact.

(ef) Mean cross-correlation values MCC for the tERK expression pattern after registration of 6 brains to \( tERK_{ZBB} \), varying each of the parameters for the ANTs SyN transform, starting with the parameters that gave the best registration for live \( vglut2a:DsRed \) based registration (SyNa[0.05,6,0.5]). Bottom right: MCCs after varying the radius of the cross-correlation metric used during registration.

(fg) MCCs for tERK in the same brains as in (ef), after combining the two best parameter sets from (ef) (SyNa[0.1,6,0.5] and SyNa[0.05,6,0.5]) to assess further improvement in registration precision. Yellow box highlights the final optimal parameter set.

(g) Horizontal section for comparison of tERK stain revealing cell morphology in the pallium after registration with optimal parameters for live \( vglut2a \) registration (left), and optimal parameters for registering fixed and stained tissue (right).

(h) Jaccard index for overlap of the manually segmented tectal neuropil of the reference brain, with each of the 6 brains in the calibration set, \( p < 0.01 \)

(i) 3D view of overlap between segmented tectal neuropils from \( tERK_{ZBB} \) (red) and the Z-Brain tERK reference brain (green), after registration with ANTs using parameter optimal for live registration, fixed registration and CMTK.
Same brains as in (e),(g) (h,l), but after registration to tERK\textsubscript{ZBB} using the parameters optimized for fixed tissue registration.

Horizontal section through the optic tectum showing tERK expression (red) and \textit{vglut2a}:\textit{DsRed} expression (green) in ZBB\textsubscript{1,2} (j) and Z-Brain (km). Matching slices within the optic tectum were selected; because the rotation around the y-axis is slightly different, sections are different within the medulla.

**Figure 4. Transformation between Z-Brain and ZBB coordinate systems using multi-channel registration**

(a) Overview of bridging Z-Brain and ZBB using ANTs\textsuperscript{\textregistered}\textsubscript{SyN} multi-channel registration. Combinations of 5 patterns common between Z-Brain and ZBB (\textit{vglut2a}\textsubscript{AV}, tERK\textsubscript{ZBB}, \textit{vmat2}, \textit{isl2b} and \textit{elavl3}) were used to guide multi-channel bridging registrations (left). MLDs for 41 landmarks in \textit{gad1b}, \textit{glyT2}, \textit{isl1}, \textit{isl2b}, tERK, \textit{vglut2a}, and \textit{vmat2} expression were measured for all reference channel combinations (middle). The combination of \textit{vglut2a}\textsubscript{AV}, tERK\textsubscript{ZBB}, \textit{vmat2}, and \textit{isl2b} enabled the most accurate bridging of ZBB and Z-Brain allowing the combination of the large collection of live transgenic lines of ZBB with the fixed tissue techniques and expert neuroanatomic segmentations of Z-Brain (right).

(ab) MCC-MLDs for the expression patterns of \textit{gad1b}, \textit{glyT2}, \textit{isl1}, \textit{isl2b}, tERK\textsubscript{ZBB}, \textit{vmat2}, \textit{isl2b} and \textit{elavl3} after registration of Z-Brain to ZBB\textsubscript{1,2} using either CMTK or ANTs\textsuperscript{\textregistered}\textsubscript{SyN} with fixed-tissue registration parameters and the indicated combination of reference channels (vglut2a, tERK\textsubscript{ZBB}, \textit{vmat2}, \textit{isl2b}, and \textit{elavl3}). Note, similar results were obtained using the \textit{tERK}\textsubscript{AV} channel, but are omitted for clarity. The combination of reference channels selected for co-registration of Z-Brain and ZBB is highlighted.

(bg) Transverse view-section through the caudal optic tectum showing the \textit{vglut2a} pattern in ZBB\textsubscript{1,2}, Z-Brain, Z-Brain after registration to ZBB with CMTK (Z-Brain-CMTK), or with ANTs (Z-Brain-SyN). The torus longitudinalis (TL) is well separated from tectal neurons in live scans, but less so in fixed tissue (arrows). The TL appears flattened after CMTK registration, but retains normal morphology after registration with ANTs\textsuperscript{\textregistered} SyN.

(ed) A comparison of transverse view-sections as in (bc), but slightly more caudal with contrast increased to highlight ventral distortion artifacts produced by registration (arrowheads).

(de-gh) Brain Browser views in the ZBB\textsubscript{1,2} coordinate (d,e,f) or Z-Brain coordinate (g,h) space. Scale bars 25 \textmu m except 50 \textmu m in (e)
Horizontal (top) and sagittal (bottom) slices, comparing the Pet1:GFP expression pattern in the superior raphe in ZBB1.2 (red) and Z-Brain after transformation to the ZBB coordinate system (green).

Horizontal (top) and transverse coronal (bottom) slices through the medulla oblongata, showing the expression of y264Et from ZBB1.2 (red) and s1181Et from Z-Brain after transformation to ZBB1.2-transformed Z-Brain (green), which both label the Mauthner cells (arrowhead).

Horizontal (top) and transverse coronal (bottom) slice sections through the medulla oblongata for glyT2:GFP from ZBB1.2 after transformation to Z-Brain (red) and the same transgenic line in Z-Brain (red,green).

Brain Browser horizontal slice-sections showing manually segmented regions transformed from the Z-Brain coordinate system to ZBB1.2 (white outlines) compared to regions previously defined in ZBB obtained by thresholding expression patterns in transgenic lines (magenta). Regions are the torus longitudinalis (hi), habenula (ij), anterior commissure (jk) and trigeminal ganglion (kl).

Figure 5. 3D visualization of brain browser data
(a) X3D zebrafish brain shown in HTML5 Web browser and (b) Virginia Tech HyperCube (CAVE)
(c) Virtual reality brain rendered using the Unity Game Engine for stereoscopic viewing using the Google eCardboard viewer. (d) In the VR browser, brain regions are selected using a menu on the floor of the virtual arena.

Additional Material

Additional File 1.pdf

Point-based landmarks for manual quantification of live-scan registration precision.

(a) Landmarks used for manually measuring registration precision. Position specifies the coordinates on vglut2aZBB (transverse, sagittal, horizontal, sagittal planes). View indicates whether the image plane shown in (b) is or transverse (T), horizontal (H), or sagittal (S). MLDs represent the average precision for each landmark for the set of 6 calibration brains, after registration with CMTK or ANTs.

(b) Images of the landmarks in vglut2aZBB (red) used for measuring precision superimposed on elavl3 (gray).

(c) Position of the landmarks superimposed on horizontal dorsal (top) and sagittal (bottom) maximum projections of elavl3 through the larval-brain.
(d) Horizontal maximum projections showing the landmark point (red dot), and the position of the corresponding landmarks in the six calibration brains after registration (green dots) superimposed on vglut2a/ZBB. Scale bar 20 μm.

Additional File 2.pdf

Cells segmented for assessing distortion introduced by registration

(a) Position of manually segmented cells for measurement of distortion introduced by registration. Views show the same cells (individually color coded) superimposed on horizontal (top) and sagittal (bottom) maximum elavl3 brain projections.

(b) Two examples of cells showing (left to right): original confocal image, segmentation mask, mask after alignment with CMTK, and mask after alignment with ANTs.

Additional File 3.pdf

Point-based landmarks labeled by transgenic lines

(a) Transgenic line landmarks used for measuring registration precision of the zebrafish brain browser atlas. Coordinates give the transverse, horizontal, sagittal position. Letter in square brackets designates Right side [R], Left side [L], or Midline [M]. The mean and standard error of the landmark distances for the three brains per landmark are indicated for CMTK and ANTs.

(b) Position of the landmarks superimposed on horizontal (top) and sagittal (bottom) maximum elavl3 brain projections.

Additional File 4.pdf

ZBB and Z-Brain expression patterns used for atlas registration

Brain Browser 3D projections of corresponding expression patterns in Z-Brain (left) and ZBB (right) used for calibrating and verifying the precision of inter-atlas registration. The top 5 patterns were combinatorially used to drive registration, while the bottom 3 were used for assessing precision. Middle images show Z-Brain patterns after registration to ZBB.

Additional File 5.pdf

Point-based landmarks for measuring precision of Z-Brain/ZBB co-registration

(a) Transgenic line and tERK-stain landmarks used for measuring registration precision of registration between Z-Brain and ZBB. Coordinates are in transverse, horizontal, sagittal sections. Letters in square brackets designates Right side [R], Left side [L], or Midline [M]. Color blocks correspond to points in (b).
(b) Position of the landmarks superimposed on horizontal (top) and sagittal (bottom) maximum brain projections.

Additional File 26.doc
Instructions for using transformation matrices to convert between ZBB and Z-Brain coordinate systems.
Table 1

| Method          | Step | Function                  | Command                                                                 |
|-----------------|------|---------------------------|-------------------------------------------------------------------------|
| Live Registration | 1    | Register vglut2a pattern in fish1-01.nii.gz, to the reference brain ref/vglut-ref.nii | antsRegistration -d 3 --float 1 -o [fish1_1_Warped.nii.gz] --interpolation WelchWindowedSinc --use-histogram-matching 0 -r [ref/vglut-ref.nii,fish1-01.nii.gz,1] -t rigid[0.1] -m MI[ref/vglut-ref.nii,fish1-01.nii.gz,1,32,Regular,0.25] -c [200x200x200x10,1e-8,10] --shrink-factors 12x8x4x2 --smoothing-sigmas 4x3x2x1vox -o Affine[0.1] -m MI[ref/vglut-ref.nii,fish1-01.nii.gz,1,32,Regular,0.25] -c [200x200x200x10,1e-8,10] --shrink-factors 12x8x4x2 --smoothing-sigmas 4x3x2x1vox -o SyN[0.05,6,0.5] -m CC[ref/vglut-ref.nii,fish1-01.nii.gz,1,2] -c [200x200x200x10,1e-7,10] --shrink-factors 12x8x4x2x1 --smoothing-sigmas 4x3x2x1vox |
|                 | 2    | Apply transformation matrix from (1) to a second channel for fish 1, in file fish1-02.nii.gz | antsApplyTransforms -d 3 -v 0 --float -n WelchWindowedSinc -i fish1-02.nii.gz -r ref/vglut-ref.nii -o fish1-02_Warped.nii -t fish1_1Warp.nii -t fish1_0GenericAffine.mat |
| Fixed registration | 1    | Register erk pattern in fish1-01.nii.gz, to the reference brain ref/terk-ref.nii | antsRegistration -d 3 --float 1 -o [fish1_1_Warped.nii.gz] --interpolation WelchWindowedSinc --use-histogram-matching 0 -r [ref/terk-ref.nii,fish1-01.nii.gz,1] -t rigid[0.1] -m MI[ref/terk-ref.nii,fish1-01.nii.gz,1,32,Regular,0.25] -c [200x200x200x10,1e-8,10] --shrink-factors 12x8x4x2 --smoothing-sigmas 4x3x2x1vox -o Affine[0.1] -m MI[ref/terk-ref.nii,fish1-01.nii.gz,1,32,Regular,0.25] -c [200x200x200x10,1e-8,10] --shrink-factors 12x8x4x2 --smoothing-sigmas 4x3x2x1vox -o SyN[0.1,6,0.5] -m CC[ref/terk-ref.nii,fish1-01.nii.gz,1,2] -c [200x200x200x10,1e-7,10] --shrink-factors 12x8x4x2x1 --smoothing-sigmas 4x3x2x1vox |
|                 | 2    | Apply transformation matrix from (1) to a second channel for fish 1, in file fish1-02.nii.gz | antsApplyTransforms -d 3 -v 0 --float -n WelchWindowedSinc -i fish1-02.nii.gz -r ref/terk-ref.nii -o fish1-02_Warped.nii -t fish1_1Warp.nii -t fish1_0GenericAffine.mat |
| ZBB                          | Z-Brain                          | Registration channel? | Quantification metric? |
|-----------------------------|----------------------------------|-----------------------|------------------------|
| Tg(vGlut2a:DsRed)nns14      | Tg(VGlut2a:EGFP)zf139            | y                     | y                      |
| single reference brain      | Tg(vGlut2a:EGFP)zf139            | y                     | n                      |
| Tg(Isl1:Gal4)zc60           | Tg(IERK immunostain)             | y                     | y                      |
| mean of 3 brains            | Tg(IERK immunostain)             | y                     | n                      |
| Tg(isl1:GFP)rw0             | Tg(isl1:GFP)rw0                  | n                     | y                      |
| mean of 6 brains            | Tg(isl1:GFP)rw0                  | n                     | y                      |
| Tg(vGlut2a:DsRed)nns14      | Tg(vGlut2a:EGFP)zf139            | y                     | y                      |
| mean of 346 brains          | mean of 15 brains                | y                     | n                      |
| Tg(vGlut2a:EGFP)zf139       | Tg(vGlut2a:EGFP)zf139            | y                     | n                      |
| single reference brain      | Tg(vGlut2a:EGFP)zf139            | y                     | n                      |
| Tg(Isl1:Gal4)zc60           | Tg(IERK immunostain)             | y                     | y                      |
| mean of 167 brains          | Tg(IERK immunostain)             | y                     | y                      |
| Tg(Isl1:GFP)rw0             | Tg(isl1:GFP)rw0                  | n                     | y                      |
| mean of 17 brains           | Tg(isl1:GFP)rw0                  | n                     | y                      |
| TgBAC(gad1b:GFP)nns25       | TgBAC(gad1b:GFP)nns25            | n                     | y                      |
| mean of 4 brains            | mean of 10 brains                | n                     | y                      |
| Tg(glyT2:GFP)cf3            | Tg(glyT2:GFP)cf3                 | n                     | y                      |
| mean of 6 brains            | mean of 13 brains                | n                     | y                      |

Table 2
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Click here to access/download

Supplementary Material
Additional File 1.png
Click here to access/download

**Supplementary Material**
Additional File 2.png
Click here to access/download
Supplementary Material
Additional File 3.png
Click here to access/download
**Supplementary Material**
Additional File 4.png
Click here to access/download
Supplementary Material
Additional File 5.png
Click here to access/download
Supplementary Material
Additional File 6.docx
June 7, 2017

Nicole Nogoy, PhD
Editor, GigaScience

Dear Dr Nogoy,

We are resubmitting for publication in *GigaScience* manuscript GIGA-D-17-00066, that provides a large dataset of zebrafish brain expression patterns, based on a procedure for allowing brain atlases to be co-registered. The original submission was favorably reviewed, but the second referee asked that we directly measure cell deformation, and drop cross-correlation as a measure of precision. To address this, and other criticisms, we have substantially revised the manuscript, which now includes four new additional files, several new figure panels, and a complete replacement of panels using our original measure of precision with the method recommended by the reviewer. Specifically:

- We manually segmented more than 100 cells, then implemented three new measures of cell shape (Hausdorff distance, elongation index and soma volume), that together, quantitatively demonstrate that our new registration procedure avoids local distortion of brain tissue. These results are described in three new panels in Fig. 1, and a new additional File 2
- Eliminated cross-correlation in Figs 1, 3 and 4, and instead defined a large number of landmarks within the larval brain, enabling us to assess precision in terms of micron distance from reference points.
- Retained cross-correlation in Fig. 2, but added a new supporting measure, Jaccard Index, to assess precision within the optic tectum
- Added flowcharts to Figs 1, 3 and 4 to clarify what images were used for each experiment, and how the resulting transformation matrices were applied.

Finally, we have comprehensively addressed each other point raised by the reviewers. We believe that the manuscript is much stronger as a result.

Thank you for your consideration.

Sincerely,

Harold A. Burgess, Ph.D.
Division of Developmental Biology
*Eunice Kennedy Shriver* National Institute of Child Health and Human Development