ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains

Olaf Ronneberger[1,2], Kun Liu[1,2], Meta Rath[3], Dominik Rueβ[1], Thomas Mueller[3,4], Henrik Skibbe[1], Benjamin Drayer[1,2], Thorsten Schmidt[1,2], Alida Filippi[3], Roland Nitschke[2,5], Thomas Brox[1,2], Hans Burkhardt[1,2] & Wolfgang Driever[2–5]

Precise three-dimensional (3D) mapping of a large number of gene expression patterns, neuronal types and connections to an anatomical reference helps us to understand the vertebrate brain and its development. We developed the Virtual Brain Explorer (ViBE-Z), a software that automatically maps gene expression data with cellular resolution to a 3D standard larval zebrafish (*Danio rerio*) brain. ViBE-Z enhances the data quality through fusion and attenuation correction of multiple confocal microscope stacks per specimen and uses a fluorescent stain of cell nuclei for image registration. It automatically detects 14 predefined anatomical landmarks for aligning new data with the reference brain. ViBE-Z performs colocalization analysis in expression databases for anatomical domains or subdomains defined by any specific pattern; here we demonstrate its utility for mapping neurons of the dopaminergic system. The ViBE-Z database, atlas and software are provided via a web interface.

Understanding brain development at the molecular and cellular levels requires integration of gene expression information for thousands of genes at a resolution precise enough to define small groups of cells. Until now, automated volumetric analysis of cellular-resolution expression data has been possible only in the brains of small invertebrates including *Caenorhabditis elegans*[1,2] and *Drosophila melanogaster*[3–5]. For the larger vertebrate brains, automated analysis of expression patterns has been limited to serial-section data[6] or global anatomical annotations[7,8], mainly for two reasons: first, standard preparation techniques and confocal microscopy do not resolve fine structures deeper than approximately 250 μm in optically dense tissue. Thus even the small zebrafish larval brain, which is thicker than 400 μm in every dimension, is excluded. Second, existing image analysis methods cannot reliably cope with confocal data sets of complex brain structures and their natural or fixation- or staining-related variations.

Restricted optical depth during vertebrate brain imaging is primarily due to absorption and light scattering. Attempts have been made to overcome this problem, with limited success. Two-photon imaging can extend the depth to 500–800 μm (ref. 9), but it cannot resolve fine details in deeper layers. Clearing solutions have been optimized to render fixed samples optically transparent[10], but they differentially deform tissues, depending on the nature of the extracellular matrix. Optical sectioning deep inside zebrafish embryos has been achieved using selective-plane illumination microscopes[11], but this requires specialized equipment. Sophisticated volumetric image-registration techniques have been developed in the medical field. However, they are optimized for magnetic resonance imaging and computer tomography data and cannot readily deal with the particular challenges in confocal microscopy[4]. Alignment by an affine transformation, such as is used for *Drosophila*[3], is not flexible enough to compensate fixation-induced deformations. A registration approach based on landmarks and thin-plate splines has also been applied to analysis of the fly brain[4] but has not been tested yet on zebrafish data. Previous work has already established algorithms for global registration of zebrafish larval images[12], but the data did not approach single-cell resolution.

ViBE-Z overcomes these restrictions through two major technical contributions. By taking images from the dorsal and ventral views of the brain and fusing these recordings with attenuation correction, ViBE-Z resolves detailed structures deep inside brain tissue using standard confocal microscope data. Moreover, ViBE-Z can align complex volumetric data sets to an anatomical model by employing a trainable, automated landmark detector. We applied ViBE-Z to generate a high-resolution gene-expression atlas of 2-, 3- and 4-d-old zebrafish larval brains, to automatically annotate anatomical domains of gene expression and to identify colocalization events in selected anatomically defined gene-expression subdomains.

RESULTS

The ViBE-Z software, database and atlas

The ViBE-Z software creates high-quality data sets through fusion of multiple confocal stacks, it aligns 3D gene or protein expression

[1]Computer Science Department, Albert-Ludwigs-University Freiburg, Freiburg, Germany. [2]BIOSS Centre for Biological Signalling Studies, Freiburg, Germany. [3]Department of Developmental Biology, Albert-Ludwigs-University Freiburg, Freiburg, Germany. [4]Freiburg Institute for Advanced Studies, Albert-Ludwigs-University Freiburg, Freiburg, Germany. [5]Zentrum für Biosystemanalyse, Albert-Ludwigs-University Freiburg, Freiburg, Germany. Correspondence should be addressed to O.R. (ronneber@informatik.uni-freiburg.de) or W.D. (driever@biologie.uni-freiburg.de).

RECEIVED 16 DECEMBER 2011; ACCEPTED 25 MAY 2012; PUBLISHED ONLINE 17 JUNE 2012; CORRECTED AFTER PRINT 9 JULY 2012; DOI:10.1038/NMETH.2076

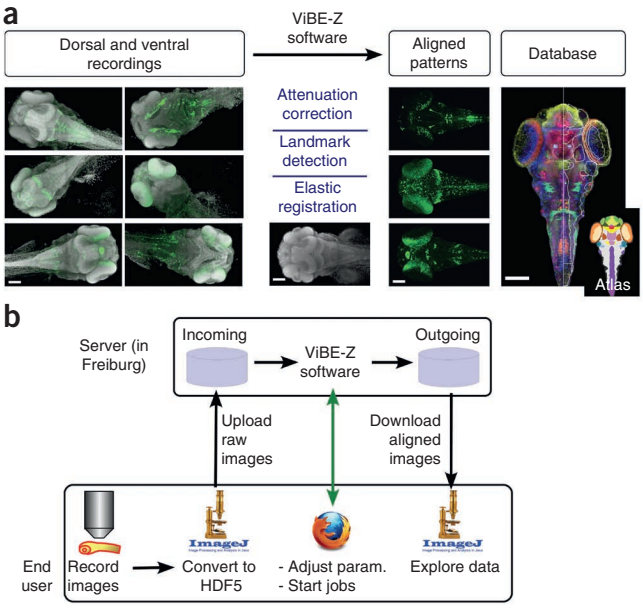

Figure 1 | The Virtual Brain Explorer for Zebrafish (ViBE-Z). (**a**) ViBE-Z overview. Input: different individuals recorded from two viewing directions (dorsal and ventral). Each recording consists of two channels: gray, nuclear stain; green, gene expression pattern. Output: aligned gene expression patterns for the ViBE-Z database and atlas. All scale bars are 100 µm. (**b**) ViBE-Z usage: the software will be provided to other scientists through a web-based interface.

patterns of different zebrafish larvae of a given developmental stage to a common reference, and it stores the aligned patterns in the ViBE-Z database. Together with the anatomical segmentation (ViBE-Z atlas), this database allows users to explore the colocalization of all contained patterns in their anatomical context. (**Fig. 1a**). The high-resolution gene-expression database (ViBE-Z database) created with the ViBE-Z software currently contains spatially aligned expression patterns of 85 individual larvae and represents the expression of 17 different genes, transgenes or antigens. The larvae were 72 h post fertilization (h.p.f.). At this stage, development is already advanced, anatomical structures are distinct, many neuronal systems have formed and larvae are still largely transparent. The ViBE-Z database is available as **Supplementary Data 1–3** (see also Online Methods). The ViBE-Z software is provided via a web interface (http://vibez.informatik.uni-freiburg.de/) (**Fig. 1b**) such that only standard open-source tools (ImageJ with the HDF5 plug-in (**Supplementary Software 1–3**) and a web browser) are required. An end user can record new gene expression patterns using our provided staining and recording protocol with any standard confocal microscope, convert the raw data to HDF5 format using ImageJ and upload the resulting file to the server in Freiburg, Germany. Processing parameters can be adjusted via a web interface. After approximately 0.5 h of processing time per larva, the user can download the aligned image, add it to a personal ViBE-Z database and explore it with ImageJ.

Attenuation-corrected high-quality data sets

Fluorescence confocal imaging of thick samples suffers from the attenuation of the excitation light and emitted light. Sources of attenuation include absorption and light-scattering effects. The common model for attenuation approximates the sample as a solid block with a constant attenuation coefficient, which results in an exponential intensity decay in the z direction[13,14]. This approximation is too coarse for objects such as the zebrafish larva head that are not of cuboidal shape and that consist of distinct tissues and spaces filled with liquid, which yield different local attenuation properties (see **Fig. 2** and **Supplementary Fig. 1**). For this reason, we propose a more realistic, physically motivated attenuation model with spatially varying attenuation coefficients.

After the preprocessing of the acquired confocal images (**Fig. 2a**) by high-dynamic-range (HDR) fusion, stitching and dorsal-ventral alignment (**Fig. 2b,c**) (see Online Methods and **Supplementary Note 1**), two intensity measurements are available at each voxel: one from the dorsal view and one from the ventral view. Combining these measurements, we have enough constraints to jointly estimate the local attenuation coefficient and the combined attenuation-free intensity at each position in the data set (**Fig. 2e**). The fused result is a single high-quality data set with homogeneous intensities and high resolution in all regions (**Fig. 2f**). The image formation is modeled using a ray-tracing approach (**Fig. 2g**). For each voxel, the excitation light and the emitted light are approximated with a cone-shaped bundle of rays and by applying the Beer-Lambert law with local attenuation coefficients. Starting with initial estimates of the attenuation-free

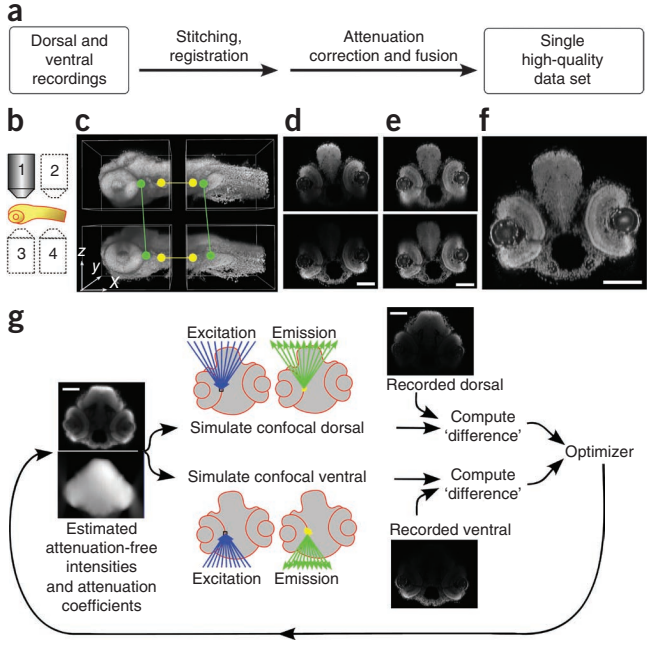

Figure 2 | High-quality data sets acquired through fusion of multiple recordings and attenuation correction. (**a**) Flow chart of the image preprocessing. (**b**) Four multichannel high–dynamic range data sets are recorded from each larva with a standard confocal microscope. (**c**) The nucleus channels of the four data sets are aligned to each other using stitching (yellow connectors) and rigid body registration (green connectors). (**d**) *yz* slices of the eyes. Top: dorsal view; bottom: ventral view. (**e**) *yz* slices of the eyes after variational attenuation correction. (**f**) *yz* slice of the eyes of the attenuation corrected fused data set. (**g**) Flow chart of the variational attenuation correction algorithm. The estimates for the attenuation-free intensities and the local attenuation coefficients are iteratively refined by minimizing the residuals between the simulated images and the recordings. All scale bars are 100 µm.

intensities and attenuation coefficients, both estimates are iteratively refined by variational optimization, which minimizes the residual between the simulated images and the recordings. This procedure takes about 4 min on a standard workstation with six cores; see **Supplementary Note 2** for more details on the variational model and its implementation.

Trainable automated landmark detection

Landmarks are required to align any new image to the reference database and thus correct for embedding-related differences in orientation across samples. We defined 14 anatomical landmarks within the larva's head (**Supplementary Fig. 2**) and manually marked them in seven samples. Using these examples, we trained a detector for each of these landmarks. After training, the detector locates the landmarks in any new sample regardless of its orientation (**Fig. 3a**–**e**).

To ensure high reliability and precise localization at the same time, the landmark detection consists of three steps. In the first step, the trained detectors are applied on a down-sampled version of the image to efficiently compute the probability of each landmark at each position. Because of the pairwise symmetric appearance of five landmarks, such as the lenses of the left and right eyes, this results in nine probability maps (**Fig. 3b**). Because the probability maps might contain false-positive detections, the second step selects the most plausible global constellation of potential landmarks through combinatorial optimization (**Fig. 3c**,**d**). The third step runs locally trained detectors on the high-resolution image to refine the position of the detected landmarks (**Fig. 3e**).

The performance of the landmark detector relies on a kernelized equivariant filter that is an extension of the harmonic filter[15–17] (**Fig. 3f**). Each local structure in the surroundings of a landmark 'votes' for the landmark position by adding a voting pattern to the landmark probability map. The mapping from a local structure to the appropriate voting pattern is learned from the seven training examples. An important property of the harmonic filter is its equivariance: that is, the voting pattern is rotated according to the orientation of the input structure. Consequently, the absolute orientation of the input (the new image) does not affect the detector. Equivariance is achieved

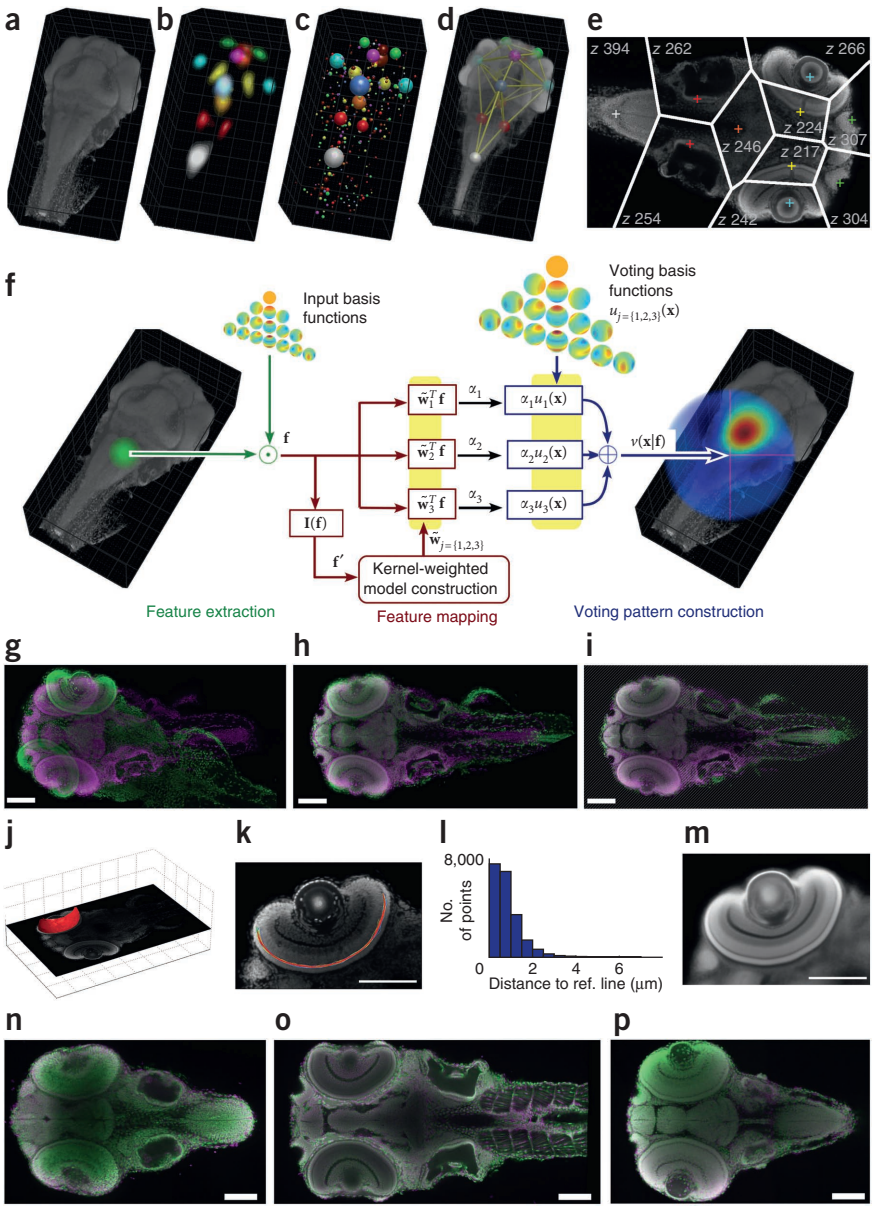

Figure 3 | Automated landmark detection and elastic registration. (**a**) Volume rendering of the nuclear-stain channel of the input image stack. (**b**) Coarse landmark probability map, computed by trained kernelized equivariant filters. (**c**) Local maxima of the probability map. The sphere size denotes the probability. (**d**) Most plausible global constellation of the landmarks based on the individual probabilities and their pairwise distances. (**e**) Refined landmark positions, computed by a second set of trained kernelized equivariant filters on the high-resolution data set. Each piece in the figure shows the *z* slice of a particular landmark. The colored cross marks the *xy* position of the landmark. (**f**) Flow chart of the kernelized equivariant filter for a certain voxel. **f**, feature vector; **f′**, rotation invariant descriptor; **I**(**f**), operator that maps **f** to **f′**; $\tilde{\mathbf{w}}_j$, filter weights; $\tilde{\mathbf{w}}^T\mathbf{f}$, scalar product of $\tilde{\mathbf{w}}$ and **f**; α_j, voting pattern coefficients; **x**, relative position in the 3D space; $u_j(\mathbf{x})$, voting basis functions; $v(\mathbf{x}|\mathbf{f})$, constructed voting pattern. (**g**) Overlay of subject (green) and reference (magenta) before registration. (**h**) Overlay after landmark-based elastic registration. (**i**) Overlay after subsequent intensity-based elastic registration. Parts outside the brain (shaded) are not considered. (**j**) Central surface of the retinal outer plexiform layer. (**k**) Outer plexiform layer of all larvae (colored lines) as overlay on the reference larva. (**l**) Registration error: distances of all plexiform layer points of all larvae to the reference line. (**m**) Average nucleus channel computed from all 72-h.p.f. larvae in the database. (**n**) Registration result of a 48-h.p.f. larva. (**o**) Registration result of a 96-h.p.f. larva. (**p**) Registration result of a 72-h.p.f. larva following WISH staining procedure. Scale bars, 100 µm.

by expressing both the input structure and the voting pattern not as a set of voxels but as a weighted sum of certain spatial-intensity distributions, so-called 'basis functions' (**Fig. 3f**). A small set of coefficients (weights) then describes the input structure and the voting pattern. The selected basis functions contain the spherical harmonics (wave functions on a spherical surface, **Supplementary Fig. 3**) in their angular part, which allows the mathematical decoupling of rotation-induced variations from the intrinsic structure. This decoupling is the basis for efficient algorithms with equivariance properties. The kernelized equivariant filter replaces the global feature mapping of the conventional harmonic filter by a locally adaptive scheme, which improves performance. A detailed description of the kernelized equivariant filter and a comparison to other detection schemes is provided in **Supplementary Note 3**.

A coarse elastic registration is performed using these detected landmarks (**Fig. 3g**,**h**). It is followed by a fine elastic registration using the image intensities, which aligns all structures precisely to the reference brain (**Fig. 3i**; Online Methods and **Supplementary Note 4**). The elastic registration approach is a state-of-the-art algorithm from the medical image–registration field[18] that was extended to work reliably on volumetric data from confocal microscopy. The total run time of the software is ~30 min for each larva, a suitable rate for application to large databases.

Validation of overall registration quality and general applicability

The proposed registration pipeline should be sufficiently precise for detailed anatomical annotation of gene expression patterns. We therefore evaluated how accurately surfaces of anatomical structures are aligned to the reference larval brain. For the retinal outer plexiform layer, we obtained a mean distance (registration error) of 0.82 µm (**Fig. 3j–l**). Details on the measurement protocol and results for other representative structures are provided in the Online Methods, **Supplementary Figures 4** and **5** and **Supplementary Table 1**. For qualitative validation of the registration accuracy, we provide the average intensity values of the nuclear-stain channel computed across all 72-h.p.f. larvae in the database (**Fig. 3m**, **Supplementary Video 1** and **Supplementary Data 1**). Because of the high registration accuracy, even the 3-µm-wide outer plexiform layer is visible.

Because the landmark detectors are trainable, only few labeled samples are required to adapt the ViBE-Z software for different experimental or reference staining procedures (**Supplementary Note 3**). We generated data sets from larvae stained for mRNA expression by the harsher whole-mount *in situ* hybridization (WISH) procedure and retrained the detectors on this type of data (**Supplementary Figs. 6–8** and Online Methods). We also demonstrate that the axonal scaffold labeled by anti-acetylated tubulin antibodies can be used as an anatomical reference instead of the nuclear stain (**Supplementary Fig. 9**). ViBE-Z also works with data from other developmental stages, such as 48-h.p.f. and 96-h.p.f. larvae (**Supplementary Figs. 10–13**). It is not necessary to change any of the registration parameters for this, and the final registration accuracy is approximately the same as with the 72-h.p.f. immunofluorescence larvae (**Fig. 3n–p**, **Supplementary Fig. 5** and **Supplementary Table 1**). This demonstrates that ViBE-Z can be easily adapted to work with different types of data as long as the recording quality of the confocal data sets is comparable to those used here.

Anatomical reference atlas and nomenclature

Working from published anatomy and nomenclature of the zebrafish larval brain[19], we segmented the 72-h.p.f. brain into 73 anatomical domains (**Fig. 4a**, **Supplementary Data 1**, **Supplementary Table 2** and **Supplementary Videos 2–4**). We used the nuclear stain to predefine these anatomical regions because it allows the identification of cell- and nucleus-enriched gray matter regions, white matter regions that have very few nuclei and ventricular cavities. The forebrain and midbrain distribution of nuclei allowed for a detailed anatomical segmentation, but the hindbrain could be annotated only globally based on the nuclear stain. We also provide an initial anatomical segmentation of major brain regions for the 48-h.p.f. and 96-h.p.f. brain (**Supplementary Data 2** and **3**).

Qualitative colocalization analysis

We generated a small database of confocal data stacks of 85 larvae representing 14 immunofluorescence and 3 WISH staining patterns: stains for individual neurons or neuronal groups (for example: dopaminergic neurons labeled using anti–tyrosine hydroxylase (anti-TH) immunofluorescence[20] and hypocretinergic neurons labeled using a transgenic line that expresses GFP under the *hypocretin* promoter[21]), axonal tracts (for example: tracts labeled using anti–acetylated tubulin immunofluorescence), or expression patterns extending over one or more anatomical domains (for example: telencephalic and hypothalamic neurons expressing YFP from the *emx3* transcription factor promoter[22]). The aligned expression pattern image stacks can be overlaid using standard channel merging tools (such as ImageJ).

Variations of expression patterns of one gene in different individuals may affect reliability of colocalization analysis. For staining patterns in distributed cell populations such as neurons of a given neurotransmitter phenotype, superimposition of stain patterns from more than one larva can help to define an anatomical volume that is representative of the pattern. This approach compensates for stochastic cell positioning within a volume. We determined the information gain by adding successive data stacks of a given pattern (Online Methods and **Supplementary Fig. 14**). For many staining patterns, the largest improvement was obtained by adding two or three data sets, whereas adding more data sets had successively smaller effects. Therefore, we used combined data volumes of three individual larvae to obtain a representative definition of the anatomical volume for each specific expression pattern.

Figure 4b shows a combined data volume obtained by adding image stacks for three anti-TH–stained larvae. An overlay with the anatomical atlas can be used to map any staining pattern to anatomical domains (**Fig. 4c** and **Supplementary Videos 5** and **6**). Three-dimensional volumes with white lines indicating borders of anatomical domains, as well as color-coded domains, are provided as **Supplementary Data 1–3** for the three stages. A combination of two or three different patterns can be used to define subpopulations of neurons. The colocalization of the expression of TH and vesicular monoamine carrier (Vmat2, also known as Slc18a2) identifies catecholaminergic neurons among aminergic neurons (**Fig. 4d** and **Supplementary Video 7**). Visualization of TH colocalizing with the expression domain of the transcription factor Otp in the posterior tuberculum and hypothalamus reveals the A11 subtype of dopaminergic neurons[23] (**Fig. 4e**

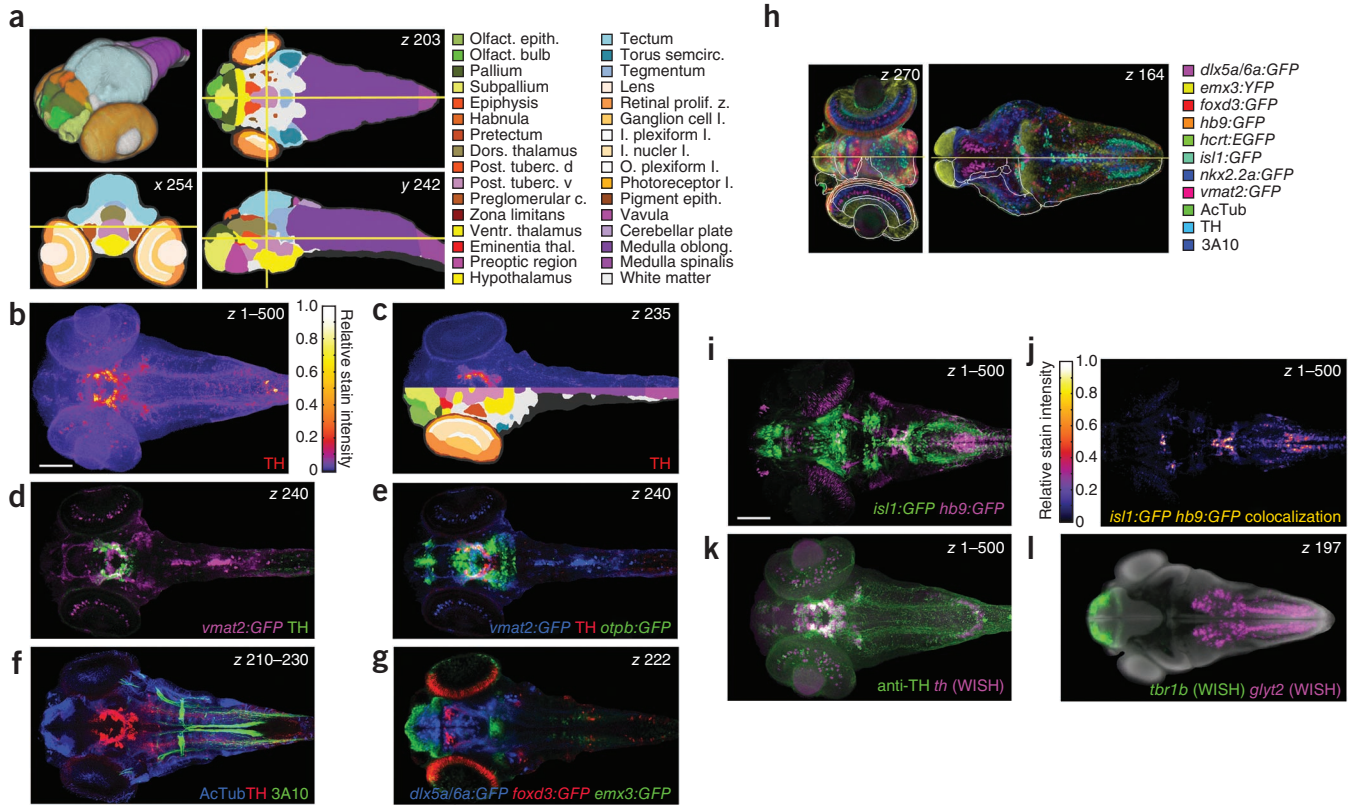

Figure 4 | Digital 3D anatomical atlas and qualitative colocalization analysis. (**a**) Anatomical segmentation of the reference zebrafish larva. Clockwise from top left: surface rendering, transversal section, lateral section, frontal section. Planes indicated by numbers, yellow lines. (**b**) Anti-TH immunofluorescence of three larvae registered into the anatomical model (maximum-intensity projection). (**c**) Combination with an anatomical model identifies anatomical domains of gene expression. (**d**–**l**) Colocalization analysis. For each antigen, GFP or mRNA pattern, registered data from three 72-h.p.f. larvae are shown; color representation is indicated in each image. (**d**) Anti-TH (wild type) and anti-GFP (*vmat2:GFP*) immunofluorescence. (**e**) Anti-TH (wild type) and anti-GFP (*vmat2:GFP* and *otpb:GFP*) immunofluorescence. (**f**) Analysis of axon bundles containing acetylated tubulin (anti-AcTub), projections of TH-expressing neurons, or axons expressing a neurofilament-associated antigen recognized by the 3A10 monoclonal antibody (anti-3A10). (**g**) Expression domains of *dlx5a/6a:GFP*, *foxd3:GFP* (reporting Foxd3 transcription factor expression in epiphysis, retina and other regions) and *emx3:GFP* transgenic markers. (**h**) Virtual qualitative colocalization analysis of 11 different expression patterns; left section of the eyes and forebrain. Right dorsal brain extending from the telencephalon through di- and mesencephalon into the hindbrain. White lines: borders of anatomical domains. *nkx2.2a:GFP* marks the ventral diencephalic expression domains of the Nkx2.2 transcription factor. (**i**,**j**) *isl1:GFP* and *hb9:GFP* expression in the zebrafish larval brain (**i**) and colocalization pattern (**j**). (**k**) Tyrosine hydroxylase mRNA (WISH) and protein (immunofluorescence) registration. (**l**) Registration of *tbr1b* transcription factor mRNA (WISH) and *glyt2* mRNA marker for glycinergic neurons (WISH), shown with average nuclear stain. (**b**–**l**) Dorsal views, anterior at left. Scale bars (**b**,**i**): 100 µm (for **b**–**h**, **i**–**l**); image planes indicated at top right. Abbreviations: c, complex; d, dorsal; i, inner; l, layer; o, outer; post, posterior; v, ventral; z, zone.

and **Supplementary Video 8**). Global qualitative colocalization analysis may also be used to analyze relative positioning of axonal tracts (**Fig. 4f** and **Supplementary Video 9**) as well as borders of expression domains of transcription factors (**Fig. 4g** and **Supplementary Video 10**). Finally, multiple simultaneous colocalization analysis can provide a tool to mine larger data sets and build hypotheses based on spatial relations of expression domains (**Fig. 4h** and **Supplementary Video 11**). We used colocalization analysis of the transgenic lines *isl1:GFP* and *hb9:GFP*, carrying transgenes expressing GFP under the control of *isl1* and *hb9* (also known as *mnx1*) regulatory elements, to determine where the Islet1 and Hb9 transcription factors may be coexpressed (**Fig. 4i**,**j**). In the spinal cord, both transcription factors are characteristic of new spinal motor neurons emerging from the ventricular zone and ventral motorneurons[24]. Colocalization analysis indicates that Islet1 and Hb9 are probably coexpressed in distinct populations of neurons adjacent to the midline in the zebrafish midbrain and hindbrain. We also determined the precision at which immunohistochemistry and WISH data can be aligned by registering anti-TH immunostain and *th* WISH data (**Fig. 4k** and **Supplementary Video 12**). The signal of the neuronal somata aligns well within the catecholaminergic anatomical groups, whereas, as expected, projections are detected only by the immunostain. We further demonstrate that WISH data for expression of mRNAs encoding the telencephalic transcription factor Tbr1b and glycinergic neuron marker Glyt2 (also known as Slc6a5) can be analyzed in the context of the anatomical brain model (**Fig. 4l** and **Supplementary Video 13**).

Automatic gene expression colocalization analysis

By mapping experimental data sets to an anatomically segmented reference larva, it becomes feasible to search for expression in defined anatomical parts. **Figure 5a** shows 14 patterns, each consisting of three overlaid individual larvae. **Figure 5b** shows

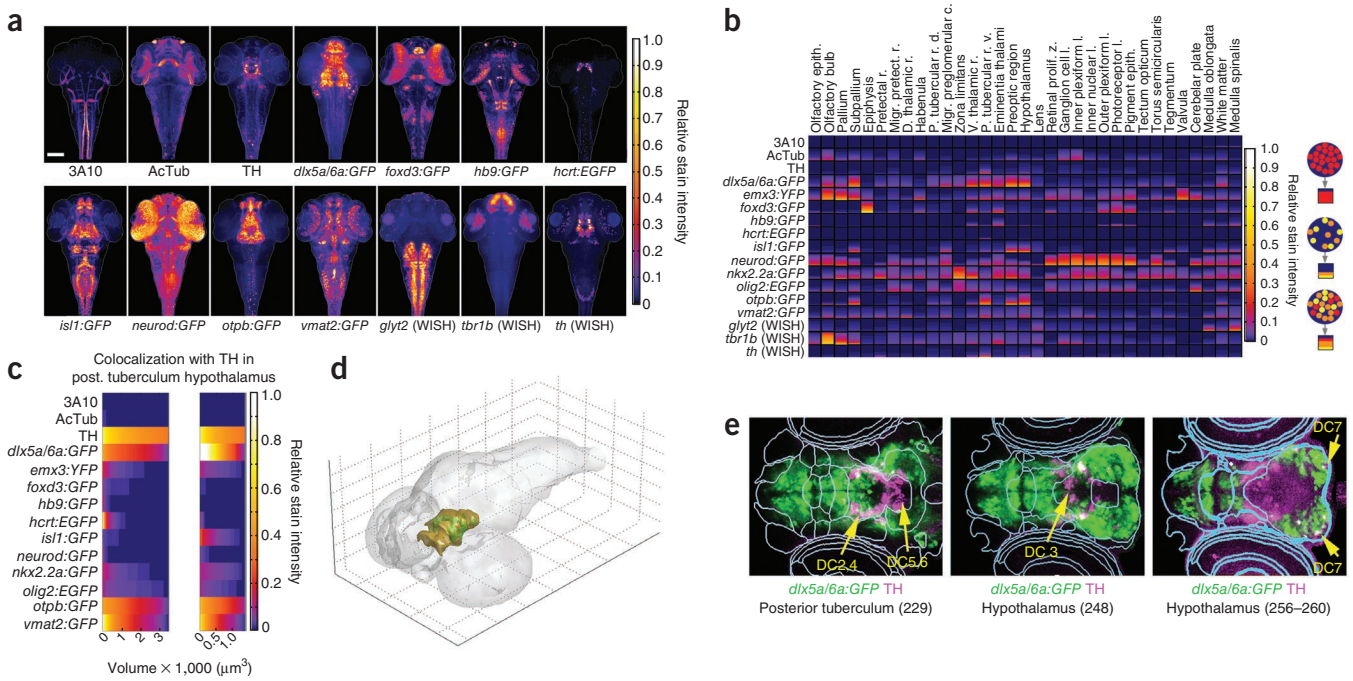

Figure 5 | Automatic anatomical mapping and colocalization analysis of expression domains. (**a**) Normalized expression data for 11 antigen and 3 mRNA (WISH) expression patterns. Each pattern represents the overlay (voxelwise maximum) of three experimental larvae. The *neurod:GFP* and *olig2:EGFP* transgenes mark the expression of the NeuroD and Olig2 transcription factors in differentiating neurons. (**b**) Expression data mapped to anatomical segments representing volumetric data of the signal as well as signal intensities. Each cell in the table represents the volume of an anatomical segment. Color represents relative stain intensities, whereas the fraction of area covered in each square represents the fraction of volume having the respective stain intensities in each anatomical domain volume (color code at right). This mapping paradigm is graphically depicted by the circles (right). Abbreviations: epith, epithelium; d, dorsal; r, region; v, ventral; l, layer; p, posterior; z, zone. (**c**) Visualization of expression volume and relative expression signal intensities in an anatomical subvolume defined by TH expression within the posterior tuberculum (left) or hypothalamus (right). The posterior tuberculum subvolume corresponds to group DC2, DC4 and some DC5 and DC6 dopaminergic neurons; the hypothalamic subvolume corresponds to DC3, some DC5 and DC6, and caudal DC7 dopaminergic neurons. The volume of expression is specified on the *x* axis and the relative signal intensities (expression levels) by a color code from high intensities (white) to background levels (dark blue; bar at right). (**d**) Visualization of the TH-expressing subvolume (green) within the posterior tuberculum (brown) and brain (gray). (**e**) Visualization of TH and *dlx5a/6a:GFP* colocalization in selected planes of the posterior tuberculum and hypothalamus. The anatomical segmentation is superimposed as blue lines. DC2–DC7 represent diencephalic dopaminergic clusters DC2 through DC7.

the result of automatic mapping of the normalized relative signal intensities to the anatomical domains of the reference larva.

Anatomical volumes used for coexpression searches may also be defined by expression patterns or by intersections of anatomical segments and expression patterns. We searched our database for colocalization in anatomical volumes defined by TH expression in the ventral diencephalic posterior tuberculum and hypothalamus (**Fig. 5c**,**d**). The expression volume diagram revealed colocalization of *vmat2:GFP* and *otpb:GFP* in dopaminergic neurons, consistent with previous published data[23,25]. In addition, the *dlx5a/6a:GFP* transgene reporting expression of the transcription factors Dlx5a and Dlx6a was coexpressed in the dopaminergic anatomical volume at low levels in the posterior tuberculum and at high levels in the hypothalamus. A closer inspection of the expression data (**Fig. 5e**) revealed that dopaminergic neurons of the posterior tubercular groups (diencephalic clusters (DC) 2 and 4) and the hypothalamic groups DC3 and DC7 appear mixed with *dlx5a/6a:GFP*, whereas group DC5 and DC6 dopaminergic neurons are located in a region devoid of high levels of *dlx5a/6a:GFP* expression. Using combined *dlx5a* WISH and anti-TH immunohistochemistry on wild-type larvae, and anti-TH and anti-GFP double immunohistochemistry on *dlx5a/6a:GFP* transgenic larvae, we experimentally confirmed that the findings predicted from our database colocalization analysis did indeed reflect coexpression (**Supplementary Fig. 15**). Rostral DC2 and DC4 neurons were intermingled with *dlx5a*, whereas caudal DC5 and DC6 are in a region devoid of *dlx5* expression. Although all four A11-type dopaminergic groups are specified by Otp, rostral and caudal groups differ in their axonal projection patterns[20]. In addition, the analysis confirmed significant coexpression of *dlx5a* in DC3 and DC7 hypothalamic dopaminergic neurons. Thus, zebrafish caudal hypothalamic groups share *dlx5a* as a potential transcriptional determinant with mammalian *Dlx5* expressing A12 arcuate nucleus dopaminergic neurons[26].

The analysis (**Fig. 5c**) further suggested colocalization of TH and *hcrt:EGFP* in the posterior tuberculum. Detailed analysis revealed that TH- and *hcrt:EGFP*-expressing cells were intermingled in this anatomical element and could colocalize at random based on the overlay of six different larval data sets in the analysis (**Supplementary Fig. 16**). However, TH and Hcrt are not coexpressed in the same cells in individual larvae (data not shown). These data emphasize that the 'virtual' colocalization analysis can serve as an indicator for potential colocalization of expression, but that it requires verification by pairwise high-resolution coexpression analysis in individual experimental larvae to demonstrate coexpression at the single-cell level.

DISCUSSION

We present a framework for building a high-resolution zebrafish brain atlas comprising anatomical information and gene expression data. The image acquisition and processing extends the applicability of standard confocal microscopy to high-quality recording of objects up to an imaging depth of at least 500 µm. The multiview recording and image-fusion algorithms are not limited to zebrafish but can also be applied to other anatomically or histologically defined biological objects. The trainable landmark detection works reliably in arbitrarily oriented larvae, which will allow application to fully automated embryo or larva handling and imaging environments. The detectors can be adapted to other developmental zebrafish stages, as well as many other structures, just by retraining with a few labeled samples.

ViBE-Z extends the set of potential applications further than related software packages, such as the BrainAligner[4]. Because our landmark detector initially learns different appearances of the same landmark and can detect this landmark in arbitrarily oriented larvae, it does not require manual prealignment of the data set (**Supplementary Fig. 17**). In contrast to the BrainAligner, which focuses on point-like structures, ViBE-Z can also align line- and plane-like structures (which dominate the zebrafish brain). This advantage arises from the dense registration that jointly optimizes with regard to point-wise similarity and global smoothness of the alignment field.

We provide a web-based interface for the scientific community to use ViBE-Z without the need to set up the computational infrastructure. The resulting standardized anatomical representation of data, together with the broadly used HDF5 file format, facilitate global sharing of data and offer new ways of combinatorial analysis of gene expression data. ViBE-Z will therefore provide a useful environment for mapping and analyzing the complexity of gene expression and anatomical brain structures in the zebrafish vertebrate model. Comprehensive gene expression analysis[7] and large-scale functional studies involving gene- or protein-trapping approaches[27] are currently in progress for zebrafish. ViBE-Z can integrate information from protein as well as mRNA expression detection. High-resolution integration of different data will help to establish neuronal identities and subtypes, to identify targets of projections and to postulate potential regulatory mechanisms based on genetic intersectional strategies. Therefore, it will complement efforts to curate neuroanatomical, expression and, eventually, functional data of the zebrafish brain (such as http://ZFIN.org/, http://www.zebrafishbrain.org/, http://zfatlas.psu.edu/). We have already demonstrated ViBE-Z's usefulness and its natural limitations in the context of dopaminergic neuron development. With this approach, we expect to integrate live fluorescence recordings and ultimately neuronal activity data into an anatomical framework at high resolution.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

The authors would like to thank S. Lin (University of California, Los Angeles), Z. Varga (University of Oregon), D. Gilmour (EMBL), M. Manoli (University of Freiburg), H. Okamoto (RIKEN), U. Strähle (Karlsruhe Institute of Technology), B. Appel (University of Colorado, Denver), D. Meyer (University of Innsbruck), J. Schweitzer (University of Freiburg), E. Zecchin and F. Argenton (University of Padova) and M. Ekker (University of Ottawa) for sharing transgenic zebrafish lines, and S. Götter for excellent fish care. Special thanks to the staff of the Center for Systems Biology Life Imaging Center for excellent confocal microscopy resources. We are grateful to A. Arrenberg, D. Onichtchouck and J. Schweitzer for critical discussion. This work was funded by the Excellence Initiative of the German Federal and State Governments (Centre for Biological Signalling Studies EXC 294; Freiburg Institute for Advanced Studies) as well as by the European Commission projects 223744 (ZF-HEALTH), 222999 (mesDANEURODEV) and 223744 (DOPAMINET).

AUTHOR CONTRIBUTIONS

O.R. designed and implemented the ViBE-Z platform and performed analyses. W.D. designed the biological experiments and performed analyses. O.R. developed the variational absorption correction. K.L., D.R., H.S. and O.R. developed the landmark detection; B.D. and O.R. developed the elastic registration. T.S. contributed software and libraries. M.R. and A.F. prepared the samples and acquired confocal images. W.D. and H.B. initiated the project. T.M. generated the brain segmentation. R.N. contributed to confocal microscopy analysis. O.R., K.L., T.B. and W.D. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nmeth.2076. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Liu, X. *et al.* Analysis of cell fate from single-cell gene expression profiles in *C. elegans*. *Cell* **139**, 623–633 (2009).
2. Long, F., Peng, H., Liu, X., Kim, S.K. & Myers, E. A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat. Methods* **6**, 667–672 (2009).
3. Chiang, A.S. *et al.* Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr. Biol.* **21**, 1–11 (2011).
4. Peng, H. *et al.* BrainAligner: 3D registration atlases of *Drosophila* brains. *Nat. Methods* **8**, 493–500 (2011).
5. Walter, T. *et al.* Visualization of image data from cells to organisms. *Nat. Methods* **7**, S26–S41 (2010).
6. Lein, E.S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
7. Thisse, B. *et al.* Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. *Methods Cell Biol.* **77**, 505–519 (2004).
8. Thisse, B. *et al.* Expression of the zebrafish genome during embryogenesis. *ZFIN Zebrafish Model Organism Database* <http://zfin.org/cgi-bin/webdriver?MIval=aa-pubview2.apg&OID=ZDB-PUB-010810-1> (2001).
9. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat. Methods* **2**, 932–940 (2005).
10. Hama, H. *et al.* Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nat. Neurosci.* **14**, 1481–1488 (2011).
11. Huisken, J., Swoger, J., Del, B.F., Wittbrodt, J. & Stelzer, E.H. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* **305**, 1007–1009 (2004).
12. Gehrig, J. *et al.* Automated high-throughput mapping of promoter-enhancer interactions in zebrafish embryos. *Nat. Methods* **6**, 911–916 (2009).
13. Can, A. *et al.* Attenuation correction in confocal laser microscopes: a novel two-view approach. *J. Microsc.* **211**, 67–79 (2003).
14. Guan, Y.Q., Cai, Y.Y., Zhang, X., Lee, Y.T. & Opas, M. Adaptive correction technique for 3D reconstruction of fluorescence microscopy images. *Microsc. Res. Tech.* **71**, 146–157 (2008).
15. Skibbe, H. *et al.* Fast rotation invariant 3D feature computation utilizing efficient local neighborhood operators. *IEEE Trans. Pattern Anal. Mach. Intell.* published online, doi:10.1109/TPAMI.2011.263 (20 December 2011).
16. Reisert, M., Canterakis, N. & Burkhardt, H. Invariants and covariants in pattern recognition and image analysis in *Pattern Recognition and Machine Vision* (ed. Wang, P.S.-P.) Ch. 5, 71–85 (River, 2010).
17. Schlachter, M. *et al.* Harmonic Filters for 3D multi-channel data: rotation invariant detection of mitoses in colorectal cancer. *IEEE Trans. Med. Imaging* **29**, 1485–1495 (2010).

18. Glocker, B., Komodakis, N., Tziritas, G., Navab, N. & Paragios, N. Dense image registration through MRFs and efficient linear programming. *Med. Image Anal.* **12**, 731–741 (2008).
19. Mueller, T. & Wullimann, M.F. *Atlas of Early Zebrafish Brain Development: a Tool for Molecular Neurogenetics* (Elsevier, 2005).
20. Tay, T.L., Ronneberger, O., Ryu, S., Nitschke, R. & Driever, W. Comprehensive catecholaminergic projectome analysis reveals single-neuron integration of zebrafish ascending and descending dopaminergic systems. *Nat. Commun.* **2**, 171 (2011).
21. Prober, D.A., Rihel, J., Onah, A.A., Sung, R.J. & Schier, A.F. Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish. *J. Neurosci.* **26**, 13400–13410 (2006).
22. Viktorin, G., Chiuchitu, C., Rissler, M., Varga, Z.M. & Westerfield, M. Emx3 is required for the differentiation of dorsal telencephalic neurons. *Dev. Dyn.* **238**, 1984–1998 (2009).
23. Ryu, S. *et al.* Orthopedia homeodomain protein is essential for diencephalic dopaminergic neuron development. *Curr. Biol.* **17**, 873–880 (2007).
24. Thaler, J. *et al.* Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. *Neuron* **23**, 675–687 (1999).
25. Wen, L. *et al.* Visualization of monoaminergic neurons and neurotoxicity of MPTP in live transgenic zebrafish. *Dev. Biol.* **314**, 84–92 (2008).
26. Yee, C.L., Wang, Y., Anderson, S., Ekker, M. & Rubenstein, J.L. Arcuate nucleus expression of NKX2.1 and DLX and lineages expressing these transcription factors in neuropeptide Y(+), proopiomelanocortin(+), and tyrosine hydroxylase(+) neurons in neonatal and adult mice. *J. Comp. Neurol.* **517**, 37–50 (2009).
27. Clark, K.J. *et al. In vivo* protein trapping produces a functional expression codex of the vertebrate proteome. *Nat. Methods* **8**, 506–515 (2011).

ONLINE METHODS

Fish breeding and strains. Zebrafish breeding and maintenance were carried out under standard conditions[28]. To prevent pigmentation, living embryos were raised in egg water containing 0.2 mM 1-phenyl-2-thiourea until they reached the desired stage.

The following wild-type and transgenic lines were used: AB/TL wild-type line; emx3:YFP - *Tg(emx3:YFP)*[22]; vmat2:GFP - *ETvmat2:GFP*[25]; foxd3:GFP - *Tg(foxd3:GFP)*[29]; isl1:GFP - *Tg(isl1:GFP)*[30], neurog1:GFP - *Tg(neurog1:GFP)*[31]; otpb:GFP - *Tg(otpb:GFP)* (M. Manoli and W.D., unpublished data); olig2:EGFP - *Tg(olig2:EGFP)*[32]; hb9:GFP - *Tg(hb9:GFP)*[33] for the *mnx1* locus; hcrt:EGFP - *Tg(-2.0FrHcrt:EGFP)*[21]; neurod:GFP - *Tg(-2.4kbNeurod:GFP)* (E. Zecchin and F. Argenton, unpublished data); dlx5a/6a:GFP - *Tg(1.4dlx5a-dlx6a:GFP)*[34] and nkx2.2a:GFP - *Tg(nkx2.2a:GFP)*[35].

Immunofluorescence and fluorescent *in situ* hybridization. Fixation was performed in 4% paraformaldehyde in PBS overnight at 4 °C. After fixation, the larvae were washed in PBST (PBS containing 0.1% Tween 20) and dehydrated stepwise with increasing concentrations of methanol (25%, 50%, 75% in PBST and 100%). Dehydrated larvae were stored at −20 °C.

Immunofluorescence was carried out as reported[36]. After rehydration, larvae were washed in PBST, blocked for at least 2 h in blocking solution (1% DMSO, 1% Blocking Reagent (Roche), 1% BSA, 5% goat serum in PBST) and incubated overnight at 4 °C with the primary antibody diluted in blocking solution. The following day, larvae were washed in PBTD (PBST, 1% DMSO) several times for 30 min and incubated overnight with the secondary antibody (diluted 1:1,000 in PBTD + 1% Blocking Reagent) in the dark. On the third day, the larvae were washed four times in PBTD and four times in PBST for 15 min. They were then counterstained, transferred to glycerol and imaged as soon as possible.

To detect the GFP in transgenic lines, we used a chicken anti-GFP antibody (Invitrogen; 1:400 dilution) and goat anti-chicken–Alexa 488 (Invitrogen; 1:1,000 dilution).

The following combinations of antibodies were used in wild-type larvae: rabbit anti–tyrosine hydroxylase primary antibody[23] and goat anti-rabbit–Alexa 488 (Invitrogen; 1:1,000 dilution); mouse anti-3A10 primary antibody[37] (Developmental Studies Hybridoma Bank; 1:50 dilution), mouse anti–acetylated tubulin (Sigma-Aldrich; 1:1,000 dilution) and goat anti-mouse–Alexa 488 (Invitrogen; 1:1,000 dilution).

Fluorescent whole-mount *in situ* hybridization was performed as described[38]. Fixed 72-h.p.f. larvae were rehydrated stepwise in decreasing concentrations of methanol (75%, 50% and 25% in PBST) and washed in PBST. They were bleached with 1% H_2O_2 in PBST and subsequently digested with proteinase K (10 μg ml^{-1}) for 45 min. After post-fixation (20 min) and several washing steps with PBST, larvae were prehybridized in hybridization mix (50% formamide, 5× SSC buffer, 5 mg ml^{-1} torula RNA, 50 μg ml^{-1} heparin, 0.1% Tween 20) at 65 °C for at least 2 h. The hybridization step was carried out in hybridization mix containing the specific digoxigenin-labeled RNA antisense probe overnight at 65 °C. The following day the larvae were washed several times at 65 °C (20 min in hybridization mix, twice for 20 (2 × 20) min in 50% formamide in 2× SSCT (2× SSC containing 0.1% Tween 20), 20 min in 25% formamide in 2× SSCT, 2 × 20 min in 2× SSCT, 3 × 30 min in 0.2× SSCT) and then once in TNT buffer (100 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% Tween 20) at room temperature (RT, 20–25 °C). After the washes, the larvae were blocked in TNTB (TNT + 1% Blocking Reagent) for at least 1 h, and incubated overnight with a peroxidase-conjugated anti-digoxigenin antibody at 4 °C (Roche; 1:400 dilution in TNTB). On the third day, the antibody was washed out with TNT (8 × 15 min) and larvae were stained according to the TSA kit instructions (Invitrogen/Molecular Probes). The staining was carried out in the dark for 1 h. Larvae were then washed in TNT three times for 5 min and stored in PBST or PBS until documentation. Antisense RNA probes used are: *th* (encoding tyrosine hydroxylase; RefSeq:NM_131149)[39], *tbr1b* (encoding T-box brain 1b; RefSeq:NM_001115090)[40] and *glyt2* (also known as *slc6a5* and encoding solute carrier family 6 member 5; RefSeq: NM_001009557)[41].

Nuclear staining. To visualize the morphological structures of the larvae, cell nuclei were stained either with TOTO-3 iodide (Invitrogen, 1:2,000 dilution in PBST) or with SYTOX Green (Invitrogen, 1:30,000 dilution in PBST) overnight at RT and then washed in PBST.

Imaging. Larvae were embedded in 80% glycerol in PBS with 1% agarose on glass slides for confocal microscopy. The embedded larvae were imaged with a Zeiss LSM510-i-NLO laser-scanning confocal microscope (Carl Zeiss MicroImaging GmbH). All confocal stacks were taken with a multi-track setting using a Zeiss LD LCI Plan-Apo 25×/0.8 objective, with glycerol as immersion medium. The resolution was set to approximately 1-μm voxel size in each direction and the slices were recorded with 512 × 512 pixels, at 12 bits per pixel. High-quality 3D data sets are obtained by a fusion of multiple stacks per larva recorded at two wavelengths, two laser intensities, at two positions (front part and rear part) from two sides (dorsal and ventral) (**Fig. 2a**). For each position and imaging direction, a total of four channels are imaged, consisting of the expression pattern of interest and a nuclear stain (anatomical reference) using two laser intensities to capture low-fluorescence structures (for example, those deep in the brain) as well as high-fluorescence structures (for example, those on the surface). The larva is mounted between two cover slips and manually turned around for dorsal and ventral recordings. The lasers, filters, and excitation and emission wavelengths used for the different fluorophores are summarized in **Supplementary Table 3**.

High-dynamic-range fusion. The absorption and scattering of light deep into tissue cause extreme differences in fluorescence intensities between surface-structure imaging and deep-structure imaging. We recorded two 12-bit image stacks at different laser intensities. The first intensity was set to avoid saturation of strong fluorescence close to the surface, and the second was set to enable good signals deep within the tissue. The corresponding channels with high and low laser intensities were combined in an HDR data set (16 bits per channel). The required parameters are directly estimated from each pair of images: the gray value of zero intensity is estimated from the main peak in the individual histograms, and the intensity factor is computed from the larger Eigenvector of the co-occurrence intensity distribution. The HDR fusion is then obtained by a maximum-likelihood estimation assuming a Poisson error model (**Supplementary Note 1**).

Stitching and dorsal-ventral alignment. The XuvTools-program[42] fuses ('stitches') the front and rear parts of the dorsal view as well as the front and rear parts of the ventral view (**Fig. 2a**, yellow connectors). This software uses the provided recording parameters (10% overlap) for the initial alignment and then finds the precise alignment (translation only) by a normalized cross correlation of small regions that is centered at points with high local contrast.

A rigid-body registration then aligns the ventral view to the dorsal view (**Fig. 2a**, green connectors). The initial alignment is obtained from the recording parameters (rotation of the sample by 180 degrees around the *y* axis). The precise alignment (translation and rotation) is found by maximizing the locally normalized cross correlation. Because of the extreme variation in the local contrast and intensity of the two images, all standard similarity measures from the medical-image analysis (such as sum of squared differences, globally normalized cross correlation or mutual information) failed on these data sets.

The dorsal image is defined on a cuboid-shaped domain $\Omega \subset \mathbb{R}^3$ and is denoted as $I_{\mathrm{dor}} : \Omega \to \mathbb{R}$; $\mathbf{x} \mapsto I_{\mathrm{dor}}(\mathbf{x})$. The ventral image is defined on the full $\mathbb{R}^3$ (using zero padding) as $I_{\mathrm{ven}}: \mathbb{R}^3 \to \mathbb{R}$. To compute the similarity measure, the domain is partitioned into N small nonoverlapping regions $\Omega_i \subset \Omega$. The rigid transformation is described by a rotation matrix $R_\theta \in \mathbb{R}^{3\times3}$ and the translation $\mathbf{t} \in \mathbb{R}^3$. The rotation is parameterized by three angles $\theta = (\theta_1, \theta_2, \theta_3)$. The corresponding energy function is

$$E[I_{\mathrm{dor}}, I_{\mathrm{ven}}](\theta, \mathbf{t}) = -\sum_{i=1}^{N} \frac{\int_{\Omega_i} (I_{\mathrm{dor}}(\mathbf{x}) - \mu_{\mathrm{dor},i}) \cdot (I_{\mathrm{ven}}(R_\theta \mathbf{x} + \mathbf{t}) - \mu_{\mathrm{ven},i,R_\theta,\mathbf{t}}) d\mathbf{x}}{\sqrt{\int_{\Omega_i} (I_{\mathrm{dor}}(\mathbf{x}) - \mu_{\mathrm{dor},i})^2 d\mathbf{x}} \cdot \sqrt{\int_{\Omega_i} (I_{\mathrm{ven}}(R_\theta \mathbf{x} + \mathbf{t}) - \mu_{\mathrm{ven},i,R_\theta,\mathbf{t}})^2 d\mathbf{x}}}$$

where $\mu_{\mathrm{dor},i}$ and $\mu_{\mathrm{ven},i,R_\theta,\mathbf{t}}$ denote the mean value of the image over the region Ω_i

$$\mu_{\mathrm{dor},i} := \frac{1}{|\Omega_i|} \int_{\Omega_i} I_{\mathrm{dor}}(\mathbf{x}) d\mathbf{x}$$

$$\mu_{\mathrm{ven},i,R_\theta,\mathbf{t}} := \frac{1}{|\Omega_i|} \int_{\Omega_i} I_{\mathrm{ven}}(R_\theta \mathbf{x} + \mathbf{t}) d\mathbf{x}$$

The energy function is minimized by a best-neighbor optimizer that starts with a step size of 16 µm and 16 degrees, reduces the step size by a factor of 2 if no better neighbor is found and terminates at a step size of 0.125.

Elastic registration. The elastic registration of a new larva to the database is carried out in two steps. Each detected landmark position in the new larva corresponds to a landmark position in the standard larva. From these correspondences, a coarse elastic transformation is computed using thin-plate spline interpolation[43]. The resulting transformation is then applied to the new larva (**Fig. 3g**,**h**).

The goal of the subsequent intensity-based elastic fine registration (**Fig. 3i**) is to find a dense deformation field that optimizes both the similarity of the elastically 'deformed' subject image to the reference image and the smoothness of the deformation field. To provide a physical metaphor, this optimization process can be described by the simultaneous action of two force fields on the subject image: the forces of the first field pull each local structure in the subject image toward similar, nearby structures in the reference image. The forces in the second field try to minimize stretching or contraction, which correspond to a simulated stiffness of the subject. The optimal deformation is found when the sum of these forces at every position is 0. This is the case, when both forces are 0 (meaning the local structures are perfectly aligned and no stretching or contraction was necessary) or if both the forces are equilibrated. In the case of equilibrated nonzero forces, a local structure is displaced toward a good matching position, but the surrounding matches and the simulated stiffness of the subject hinder it from going there completely. The relative weight of these two forces is adjusted by a global 'smoothness' parameter that defines the position of the equilibrium.

The search for the optimal deformation field was reformulated as a series of combinatorial optimization problems[18]. This approach uses efficient graph-based solvers, which are much less sensitive to local optima than the standard gradient-descent techniques. Furthermore, these graph-based solvers are often capable of finding the global optimum[44].

To use the framework for microscopic images, we reimplemented the complete system and replaced the similarity measure by a locally normalized cross correlation. This procedure resulted in two advantages. First, a high robustness against local contrast and intensity changes within the images was obtained, and second, we acquired an intuitive way to set the smoothness parameter. The similarity measure cannot get higher than 1.0 (for a perfect match), so the smoothness parameter can easily be adjusted such that no deformations with a stretching factor beyond a certain threshold appear. This stretching factor threshold was set to 2.0 for all registrations in the database, resulting in very accurate registration results in spite of differences in image quality (see **Supplementary Note 4**).

The runtime of the full elastic registration (coarse and fine) including loading, scaling and saving is approximately 3 min on a standard workstation (2.6 GHz hexa-core processor). The full data set size is 1,000 × 500 × 500 voxels. After downscaling by a factor of two and masking out non-brain regions, the processing volume contains approximately 3.6 million voxels.

Elastic registration of data sets with systematic deformations. Several common preparation techniques cause considerable deformation and shrinkage of the samples. But these deformations are often not random. For example, in the case of *in situ* hybridization, most deformations (say, the shrinkage of liquid-filled regions such as the ears) are similar from sample to sample, which allows a systematic correction of these deformations.

Correction was achieved by a two-step registration. The first step compensates for the individual variations through a registration to an intermediate reference larva treated with WISH. The second step then compensates for the systematic variations between WISH-treated larvae and the standard larva by applying an appropriate deformation field that has been computed in advance with manual parameter optimization (**Supplementary Fig. 8**).

Validation of registration accuracy. For the validation of the overall registration accuracy three representative locations were

doi:10.1038/nmeth.2076

considered: (i) the center of the outer plexiform layer of the retina (**Fig. 3j–m**), (ii) the midline between the left and the right part of the subpallium and (iii) the separation surface formed by the constriction at the mid-hindbrain boundary (**Supplementary Figs. 4** and **5**). The accuracy was measured in a 2D slice that is approximately perpendicular to that surface (**Fig. 3j**).

In these 2D slices, the selected surfaces appear as thin dark lines, which were semiautomatically traced in the registered data sets using the NeuronJ plug-in for ImageJ[45]. The resulting lines from all registered larvae are shown in **Figure 3k** and **Supplementary Figures 4a**,**f**,**k** and **5** as overlays on the reference larva. The surface distances were measured relative to representative points on the reference line (**Supplementary Fig. 4b**,**g**,**l**). We obtained a mean distance (registration error) of 0.82 µm in the outer plexiform layer and 0.97 µm in the midline subpallium (**Supplementary Fig. 4e**,**j**) as examples for distinctively visible surfaces. These structures have distinct surfaces opposed to the mid-hindbrain boundary, which is a highly diverse region and often difficult to align correctly without introducing severe deformations to the surrounding tissue. Here we obtained a mean registration error of 3.2 µm (**Supplementary Fig. 4o**).

Compensation of inhomogeneous staining. The image fusion and attenuation correction steps in the ViBE-Z pipeline were designed to reconstruct the attenuation-free intensities of the image by compensating for known physical imaging artifacts. In addition, we also have to consider biological effects and staining artifacts that may influence the mapping of gene expression events to image intensities. A typical staining artifact is an inhomogeneous penetration of the fluorescent markers. We compensate for this effect by using the distribution of stain intensities in the nuclear stain channel. Relating these intensities to the intensities of the most homogeneously stained sample in the database allows us to compute local compensation factors for the new sample. These compensation factors are then applied to the gene expression channel. This compensation works well for the used combination of nuclear staining by TOTO3 and the expression-pattern staining by immunofluorescence. For the WISH data set, however, the predicted compensation factors from the TOTO3 channel did not improve the result in several samples. A potential reason is that the inhomogeneities in sample penetration and the biological tissue variations in these different staining techniques may have different effects on the signal. Therefore this procedure was not applied to WISH data. A second common artifact is a global background intensity that can have multiple biological and staining-related sources. For compensation we subtract a global background intensity that is estimated from the highest peak in the image histogram.

Combination of multiple individuals per stain. The combination of multiple individuals per stain was performed by the voxelwise maximum operation. To ensure the same weight for each individual, the sum over each input image was normalized to 1. After combination, the resulting gray values were rescaled to the range [0,1] by using the maximal gray value in the downscaled image (5 µm voxel size). The downscaling operation makes the result robust against single bright voxels.

Anatomical segmentation. The anatomical segmentation was performed with the Segmentation-Editor plug-in (http://132.187.25.13/home/?category=Download&page=SegmentationEditor) for ImageJ using the nuclear-stain channel of the reference larva. Anatomical delineations and nomenclature follow the atlas of early zebrafish brain development[19]. In general, anatomical borders were delineated based on the nuclear stain. Anatomical borders within the forebrain were refined using the registered expression patterns of *dlx5/6* and *neuroD*. All drawn 3D masks were then smoothed with a small Gaussian filter ($\sigma = 2$ µm) and re-binarized collectively such that at each voxel, the mask with the highest weight is selected. This operation closes small gaps between neighboring masks and creates smooth surfaces in all views. The final segmentation is available as an indexed image in the data set '/anatomy/segment_regions' in each HDF5 file (**Supplementary Data 1–3**).

ViBE-Z database file. All 17 72-h.p.f. gene or transgene expression patterns (each combined from three individuals, see **Supplementary Table 4**) and the anatomical segmentation are available as **Supplementary Data 1**. The database is provided in HDF5 format (hierarchical data format; http://www.hdfgroup.org/HDF5/), which allows users to store any number of multidimensional data sets in a single file and access them by names in a hierarchical directory-like structure, and which provides powerful lossless compression (the ViBE-Z database file is only about 230 MB in size but contains about 4.2 GB of raw data).

In addition, the initial ViBE-Z database files for the 48-h.p.f. larva (**Supplementary Data 2**) and the 96-h.p.f. larva (**Supplementary Data 3**) are provided. The 48-h.p.f. and 96-h.p.f. databases each contain a reference larva, an average brain and an anatomical segmentation.

All data sets have been quantized to 8 bit to keep the file small and to allow convenient and fast operation on standard computers. The original data sets (32-bit floating-point precision) will be provided on the ViBE-Z homepage. The data sets in the atlas file can be loaded into ImageJ using our HDF5 plug-in (provided as **Supplementary Software 1** and **2**). The latest version of the HDF5 plug-in will always be available online at http://lmb.informatik.uni-freiburg.de/resources/opensource/imagej_plugins/hdf5.html. For a display of the anatomical segmentation, we provide a color lookup table as **Supplementary Software 3**.

Batch jobs can be easily programmed in MATLAB (http://www.mathworks.com/), as it natively supports the HDF5 format. For example, loading the AcTub pattern and displaying a maximum-intensity projection of it is done with a few commands:

```
a = hdf5read('ViBE-Z_72hpf_v1.5h','/expression/AcTub');
m = max(a,[ ],3);
image(m'); axis image; colormap gray
```


28. Westerfield, M. *The Zebrafish Book: a Guide for the Laboratory Use of Zebrafish (Danio rerio)* 4th edn (University of Oregon Press, 2000).
29. Gilmour, D.T., Maischein, H.M. & Nusslein-Volhard, C. Migration and function of a glial subtype in the vertebrate peripheral nervous system. *Neuron* **34**, 577–588 (2002).
30. 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.* **20**, 206–218 (2000).
31. Blader, P., Plessy, C. & Strahle, U. Multiple regulatory elements with spatially and temporally distinct activities control neurogenin1 expression in primary neurons of the zebrafish embryo. *Mech. Dev.* **120**, 211–218 (2003).

32. Shin, J., Park, H.C., Topczewska, J.M., Mawdsley, D.J. & Appel, B. Neural cell fate analysis in zebrafish using olig2 BAC transgenics. *Methods Cell Sci.* **25**, 7–14 (2003).
33. Flanagan-Steet, H., Fox, M.A., Meyer, D. & Sanes, J.R. Neuromuscular synapses can form in vivo by incorporation of initially aneural postsynaptic specializations. *Development* **132**, 4471–4481 (2005).
34. Zerucha, T. *et al.* A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. *J. Neurosci.* **20**, 709–721 (2000).
35. Pauls, S., Zecchin, E., Tiso, N., Bortolussi, M. & Argenton, F. Function and regulation of zebrafish nkx2.2a during development of pancreatic islet and ducts. *Dev. Biol.* **304**, 875–890 (2007).
36. Holzschuh, J. *et al.* Noradrenergic neurons in the zebrafish hindbrain are induced by retinoic acid and require tfap2a for expression of the neurotransmitter phenotype. *Development* **130**, 5741–5754 (2003).
37. Furley, A.J. *et al.* The axonal glycoprotein TAG-1 is an immunoglobulin superfamily member with neurite outgrowth-promoting activity. *Cell* **61**, 157–170 (1990).
38. Filippi, A. *et al.* Expression and function of nr4a2, lmx1b, and pitx3 in zebrafish dopaminergic and noradrenergic neuronal development. *BMC Dev. Biol.* **7**, 135 (2007).
39. Holzschuh, J., Ryu, S., Aberger, F. & Driever, W. Dopamine transporter expression distinguishes dopaminergic neurons from other catecholaminergic neurons in the developing zebrafish embryo. *Mech. Dev.* **101**, 237–243 (2001).
40. Yonei-Tamura, S., Tamura, K., Tsukui, T. & Izpisúa Belmonte, J.C. Spatially and temporally-restricted expression of two T-box genes during zebrafish embryogenesis. *Mech. Dev.* **80**, 219–221 (1999).
41. Higashijima, S.-I., Mandel, G. & Fetcho, J.R. Distribution of prospective glutamatergic, glycinergic, and GABAergic neurons in embryonic and larval zebrafish. *J. Comp. Neurol.* **480**, 1–18 (2004).
42. Emmenlauer, M. *et al.* XuvTools: free, fast and reliable stitching of large 3D datasets. *J. Microsc.* **233**, 42–60 (2009).
43. Bookstein, F.L. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567–585 (1989).
44. Komodakis, N., Tziritas, G. & Paragios, N. Performance vs computational efficiency for optimizing single and dynamic mrfs: setting the state of the art with primal-dual strategies. *Comput. Vis. Image Underst.* **112**, 14–29 (2008).
45. Meijering, E. *et al.* Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry A* **58**, 167–176 (2004).

doi:10.1038/nmeth.2076

Erratum: ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains

Olaf Ronneberger, Kun Liu, Meta Rath, Dominik Rueβ, Thomas Mueller, Henrik Skibbe, Benjamin Drayer, Thorsten Schmidt, Alida Filippi, Roland Nitschke, Thomas Brox, Hans Burkhardt & Wolfgang Driever
Nat. Methods **9**, 735–742 (2012); **published online 17 June 2012; corrected after print 9 July 2012.**

In the version of this article initially published, two items in the Online Methods section were incorrect. The MATLAB code in the 'ViBE-Z database file' section contained an extraneous semicolon, which appeared in the HTML only and has been corrected. In the section 'Stitching and dorsal-ventral alignment', two formulas had a 'mapsto' symbol instead of an arrow. These errors have been corrected in the HTML and PDF versions of the article.