

Automated prior knowledge-based quantification of neuronal patterns in the spinal cord of zebrafish

Johannes Stegmaier^{1,*†}, Maryam Shahid^{2,3,†}, Masanari Takamiya², Lixin Yang², Sepand Rastegar², Markus Reischl¹, Uwe Strähle^{2,3,†} and Ralf Mikut^{1,†}

¹Institute for Applied Computer Science (IAI), Karlsruhe Institute of Technology, Karlsruhe, Germany, ²Institute for Toxicology and Genetics (ITG), Karlsruhe Institute of Technology, Karlsruhe, Germany and ³Faculty of Biosciences, Ruprecht-Karls-University of Heidelberg, Heidelberg, Germany

Associate Editor: Martin Bishop

ABSTRACT

Motivation: To reliably assess the effects of unknown chemicals on the development of fluorescently labeled sensory-, moto- and interneuron populations in the spinal cord of zebrafish, automated data analysis is essential.

Results: For the evaluation of a high-throughput screen of a large chemical library, we developed a new method for the automated extraction of quantitative information from green fluorescent protein (eGFP) and red fluorescent protein (RFP) labeled spinal cord neurons in double-transgenic zebrafish embryos. The methodology comprises region of interest detection, intensity profiling with reference comparison and neuron distribution histograms. All methods were validated on a manually evaluated pilot study using a Notch inhibitor dose-response experiment. The automated evaluation showed superior performance to manual investigation regarding time consumption, information detail and reproducibility.

Availability and implementation: Being part of GNU General Public Licence (GNU-GPL) licensed open-source MATLAB toolbox GaitCAD, an implementation of the presented methods is publicly available for download at <http://sourceforge.net/projects/zebrafishimage/>.

Contact: johannes.stegmaier@kit.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on July 26, 2013; revised on September 23, 2013; accepted on October 15, 2013

1 INTRODUCTION

Research in the life sciences is producing an ever increasing amount of experimental data in basically all areas. Particularly the fields of toxicology, chemical genetics and classical genetics, where possibly tens of thousands of treatments of embryos have to be analyzed, benefit from lately established high-throughput acquisition techniques that enable large-scale experiments, which have been inconceivable in the past (Gehrig *et al.*, 2009). In recent years, zebrafish embryos have emerged as attractive models for the screening of small molecules with the aim to identify new entities with a specific effect on cellular and developmental processes (Flinn *et al.*, 2008). The zebrafish provides

a complex cellular environment that cannot be mimicked in conventional pharmacological screens with *in vitro* systems. Furthermore, a multitude of properties such as fast reproduction and *ex utero* development, easy husbandry and handling, availability of many stable transgenic lines and the transparency in early developmental stages make the zebrafish an ideal model organism for high-throughput screening (Detrich *et al.*, 2011). In many cases, however, the facilitated and highly automated possibilities on the experimentalists' side unravel a variety of new challenges and problems in subsequent data analysis steps. As meaningful information contained in the large amount of obtained data can hardly be extracted manually anymore, specialized software is needed to support or optimally take over the whole data analysis in an automated pipeline (Eliceiri *et al.*, 2012; Liu *et al.*, 2008b; Stegmaier *et al.*, 2012).

Besides various articles describing manually/semiautomatically evaluated experiments (Kanungo *et al.*, 2011; Yang *et al.*, 2010b), multiple approaches for the automated evaluation of zebrafish-based screens have already been proposed in the literature. A comprehensive survey can be found in Mikut *et al.* (2013). Methods for the automated analysis of gene expression based on image co-localization and morphing of regional intensity information to a reference embryo were described by Gehrig *et al.* (2009) and Ronneberger *et al.* (2012). Various publications tackled the problem of automated phenotype recognition (Cao *et al.*, 2009; Jeanray *et al.*, 2012; Liu *et al.*, 2012; Vogt *et al.*, 2009) by means of multi-resolution segmentation, randomized sub-window- and descriptor-based classification, respectively. Moreover, to detect and quantify nuclei in images, several blob detection methods exist, e.g. based on circular Hough transform, diffused gradient vector field and Laplacian-of-Gaussian (LoG) blob detection, as described by Al-Kofahi *et al.* (2010) and Liu *et al.* (2006), (2008a). Approaches for the automated region of interest (ROI) detection via template matching were presented by Peravali *et al.* (2011) or Chen *et al.* (2011). The approaches described by Chen *et al.* (2011) and Liu *et al.* (2006, 2008a) are most closely related to our analysis. However, it is not yet possible to quantify the diverse neuron distributions of the spinal cord in zebrafish.

The aim of this work was to derive an automated analysis pipeline that could be used to analyze the effect of chemicals on position, number and fluorescence intensity of the three different neuronal cell types in the spinal cord of zebrafish, namely

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two and last two authors should be regarded as Joint First Authors.

Rohon–Beard sensory neurons, interneurons and motoneurons (Supplementary Fig. S1). Therefore, we performed a pilot study using a double-transgenic zebrafish with Rohon–Beard cells and motoneurons expressing RFP and interneurons expressing eGFP in the spinal cord. The developed algorithms should reliably extract both quantitative and qualitative information and should be applicable for high-throughput screening.

In Section 2, we briefly present the experimental design of the pilot study used for algorithmic development and introduce the newly developed image analysis pipeline, which is comprised of ROI extraction, intensity profile analysis of different neuron types as well as quantification of sensory- and motoneurons. The validation of the methods and biological conclusions can be found in Section 3. Finally, in Section 4, we discuss the benefits and limitations of our methods.

2 METHODS

2.1 Experimental design

To establish a new automated analysis method for zebrafish-related screens, we exploited the fluorescent properties of spinal cord neurons of different sizes, location and signal intensity. We used a double-transgenic zebrafish line to tailor algorithms that are specifically suited to extract information on number and position of neurons in the spinal cord region of laterally oriented zebrafish. The purpose of using this particular transgenic combination is to enable the investigation of the formation of the three key neuronal types of the spinal cord (sensory-, inter- and motoneurons) in one embryo.

2.1.1 Fish maintenance and transgenic lines The hybrid two-color transgenic line was generated by crossing a *dbx1b* (developing brain homeobox 1b) transgenic zebrafish driving eGFP expression in dI6, V0 and V1 interneurons (Yang *et al.*, unpublished) with a *Neurog1* (Neurogenin 1) transgenic zebrafish driving RFP expression in Rohon–Beard cells and motoneurons (Blader *et al.*, 2003 and Supplementary Fig. S1). Double homozygous parents were outcrossed with wild-type AB zebrafish. This resulted in inheritance of each of the transgenes by 100% offspring in a hemizygous state assuring equal intensity of transgene expression in each individual. All zebrafish husbandry and experimental procedures were performed in accordance with the German animal protection standards and were approved by the Government of Baden–Württemberg, Regierungspräsidium, Karlsruhe, Germany (AktENZEICHEN 35-9185.64). The zebrafish line was maintained as described by Westerfield (2007). The fish were kept and raised at the Institute of Toxicology and Genetics, Karlsruhe Institute of Technology, Germany.

2.1.2 Chemicals The LY-411575 compound, a potent gamma-secretase inhibitor that blocks Notch signaling through preventing the release of the Notch intracellular domain from the plasma membrane, was used to disturb normal neurogenesis (Fauq *et al.*, 2007). In a previous study, Shin *et al.* (2007) demonstrated that the loss of Notch signaling results in the formation of excess primary motoneurons and a deficit of Kolmer–Agfuer primary cells (KA' interneurons) as early as 20–25 hours post fertilization (hpf). However, according to our observation, at this stage Notch inhibitor causes the trunk to bend severely. Therefore, the used exposure time window between 24–48 hpf in our experiment was selected to affect the neurogenesis of late-born neuronal population, minimizing this malformation that could interfere with the results. All the working concentrations of LY-411575 were made in 0.003% 1-phenyl 2-thiourea (PTU) in fish water. PTU inhibits pigment formation by blocking tyrosinase-dependent steps in melanin synthesis so that

pigmentation does not interfere with the detection of the fluorescent signal (Karlsson *et al.*, 2001).

2.1.3 Experimental Conditions Embryos were collected and washed thoroughly with buffered egg water [reverse osmosis water containing sea salt (pH 6.73): Ca 0.8 mg/l; 60 µg/ml; Instant Ocean, Red Sea, Houston, TX, USA, K 0.6 mg/l; Mg 2 mg/l; Na 16 mg/l; S 2 mg/l (Yang *et al.*, 2010a)]. The embryos were manually dechorionated at 24 hpf. In one 96-well plate, three embryos per well were distributed with 150 µl of 0.003% PTU (prepared with egg water) as a medium. In a dose-response experiment for LY-411575 (100 µM stock concentration in 0.003% PTU), embryos were placed in 11 different concentrations between 0.5–10 µM. Control embryos were incubated in embryo medium containing 0.04% dimethyl sulfoxide. Three biological repeats were performed and each embryo was imaged twice, yielding to at least six image sets per concentration. Better focused images were manually selected to have three high-quality images of three embryos per concentration and repeat, ending up with images of 108 embryos.

2.1.4 Microscopy parameters To ensure an informative readout, a specific set of microscopy parameters was necessary to visualize the three types of neurons. Images were taken at 48 hpf using a Leica compound microscope DM5000 at 10× magnification (HC PL FLUOTAR 10× NA 0.30 objective). RFP is expressed by both motoneurons and sensory neurons, however, at different intensities. For technical reasons, it was not possible to acquire images with different exposure parameters through the same filter channel. Hence, we used two different filter systems (RFP, CY3) with overlapping sensitivities to image sensory and motoneurons (Supplementary Section S4). This allowed us to optimize exposure times such that each neuron type was optimally visible in one of the channels (RFP: 40, GFP: 200, CY3: 410 ms). The remaining parameters, i.e. gamma, saturation and gain, were identical for all channels and set to 0.99, 1.5 and 4, respectively. Images were acquired with an intensity range of 8 bit at a resolution of 1392 × 1040 pixels (px) with a physical pixel size of 1 px ≅ 0.64 µm.

2.1.5 Algorithmic parameters Owing to the linear shape observed in all images of this pilot study, rotation correction was turned off and a radius of $r = 70$ was used to extract the cropped images, i.e. the cropped images had a resolution of 1392 × 141 px (Section 2.3). We used $\sigma_{\min} = \{2, 3\}$, $\sigma_{\max} = \{3, 4\}$ and $\sigma_{\text{step}} = \{1, 1\}$ for detection of the motoneurons (CY3) and sensory neurons (RFP), respectively (Section 2.6 and Supplementary Fig. S2). Note that these parameters are directly depending on the physical pixel size and have to be adapted with respect to the investigated image material.

2.1.6 Validation experiments Statistical values of extracted features were calculated on all embryos with identical treatment (nine images per concentration) for all 108 available images. To check for significant differences of extracted features to the control group, we used a two-sided Wilcoxon rank sum test ($\alpha = 5\%$). The quality of extracted regression curves of all images was manually assessed. Moreover, to validate the seed detection method, Rohon–Beard cells were independently counted by two experts on a blinded subset of 36 randomly selected images with three images per concentration. Manual counting was not performed for motoneurons due to the high ambiguity in out of focus regions. This subset was also used for qualitative assessment of the treatment effect by the trained human into the categories *no effect*, *slight effect* and *strong effect*. Using the manually identified hypothetical class associations, we trained a Bayes classifier with ANOVA-based (ANalysis Of VAriance) selection of one feature. The classifier performance was validated with leave-one-out cross validation.

2.2 Prior knowledge-driven algorithmic design

Available prior knowledge such as geometric properties of the specimen, spatial localization of different objects of interest and the size of stained nuclei can be used to design, fine-tune and improve developed algorithms for specific problems. In our pilot screen, we made use of three different channels for imaging, which allowed us to enhance the fluorescent signal of specific objects of interest by adjusting the exposure times separately for each channel. Automatic acquisition of an overlay image within a few seconds preserved focus and guaranteed a perfect co-localization of the images. The sensory- and motoneurons were visible in both the RFP and the CY3 channel. Owing to the different expression levels, we used a shorter exposure time for the RFP channel to achieve a good imaging of the sensory neurons and a longer exposure time for capturing the CY3 signal to visualize the less prominent motoneuron response. In later analysis steps, the corresponding overexposed Rohon–Beard cells in the CY3 channel were not considered and the motoneurons visible in the RFP channel were only considered in the CY3 channel. The prior information is given in Table 1.

Based on these properties, a new image analysis pipeline was developed to identify alterations in spatial appearance of the different neuron types and to possibly detect different numbers of present nuclei. The spinal cord center, i.e. an imaginary line running through the center of the interneuron band, was used as a spatial reference for all detected features. The pipeline concept derived from the available prior knowledge is depicted in Figure 1, and the individual steps are introduced in the following sections.

Table 1. Prior knowledge used for the algorithmic design

Type	Channel	Description
Interneurons (dl6, V0, V1)	GFP	Thick band of cells between sensory- and motoneurons, individual cells not detectable, cubic polynomial to model the spinal cord
Sensory neurons	RFP, CY3	Mostly well-separated nuclei on dorsal side of the spinal cord, overexposed in CY3, good signal in RFP channel, diameter: 7 – 10µm
Motoneurons	RFP, CY3	Densely packed nuclei on the ventral side of the spinal cord, weak signal in RFP channel, good signal in CY3 channel, diameter: 5 – 7µm

Note: Individual channels were perfectly co-localized and exhibited the fluorescent signal of stained nuclei from three different neuron types with channel-specific exposure times and parameters as described in Section 2.1.4.

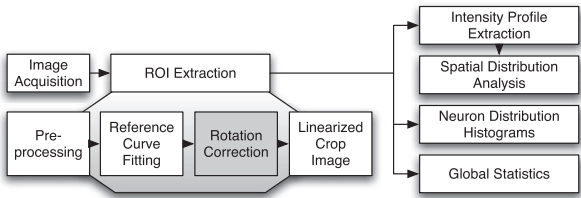


Fig. 1. Schematic overview of the processing pipeline. Components shaded in gray can be optionally enabled if required by the image material

2.3 Region of interest extraction

Initially, a reproducible comparable image region had to be cropped from the original images **I** with a resolution of $I_x \times I_y$. Based on prior knowledge about the presence of a band of interneurons that appeared as a dense stripe running along the spinal cord in the GFP images (Fig. 2A), we tried to model the backbone of the investigated zebrafish by a polynomial regression approximation through this band. Therefore, we applied an intensity adjustment that saturated the lower and the upper 1%-quantiles of the intensity values, followed by a Gaussian smoothing, to preserve only the low frequency content of the images. For the image material considered here, a standard deviation of $\sigma = 10$ worked well in practice.

Subsequently, a binarized image (Fig. 2B) of the smoothed GFP image was created using Otsu’s method (Otsu, 1979). To further improve the segmented regions and get a dense connected binary object, a morphological closing operation was performed. In the case of a reduced fluorescent signal of interneurons, however, the auto-fluorescent signal of the yolk sac became more prominent than the desired band of interneurons. To make the regression curve extraction more reliable in these cases, we used prior knowledge about the expected position of the band of interneurons in the laterally oriented embryos, to assign the connected components in the binary image to the respective tissues. We averaged all rows of pixels in the horizontal direction of the image as shown in Equation (2) but for the whole uncropped image and identified intensity peaks in this representation. If two dominant, sufficiently separated peaks were detected, the minimum between the peaks was used as a cutoff point and pixel values on the ventral side of this minimum, i.e. the pixels associated with the yolk sac were set to zero. In the dichotomized image we used the image space coordinates of foreground pixels to identify a regression model that minimized the sum of squared distances, i.e. we solved:

$$\hat{\theta} = \operatorname{argmin}_{\theta \in \mathbb{R}^n} \sum_{(i_x, i_y) \in P_f} (i_y - f(i_x, \theta))^2 \tag{1}$$

In Equation (1), f represents an arbitrary regression function with unknown parameter vector θ , $\hat{\theta}$ is the minimizing parameter vector and P_f is the set of foreground pixel coordinates that are used to perform the regression. In the special case considered for this work, i.e. laterally oriented zebrafish embryos, a third order polynomial defined as $f(x, \theta) = \theta_1^3 + \theta_2^2 + \theta_3 + \theta_4$, with parameter vector $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)^T$, sufficiently described the possible deformations of the spinal cord region. Choosing an appropriate model for the regression was crucial to attain comparable results. The individual processing steps for the ROI extraction are depicted in Figure 2. A linearized version of the images was achieved by extracting a cropped image region with respect to the regression curve, i.e. a band of a predefined size r that matched the spinal cord region below and above the regression curve was used to form

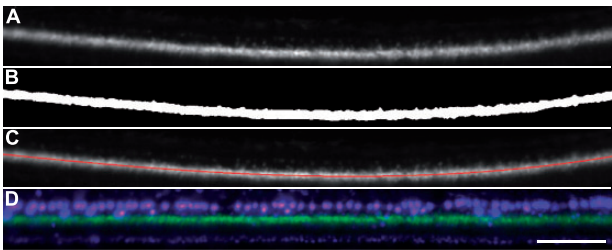


Fig. 2. Steps to extract comparable regions of interest from the microscopy images. Cropped original image (A), thresholded image using Otsu’s method (B), regression curve overlaid on the original image (C), linearized crop image for RFP, GFP and CY3 channels (D). Scale bar = 100µm

a new image. This yielded a linearized and cropped version of the spinal cord region \mathbf{I}_c as shown in Figure 2D with size $I_x \times (2 + 1)$. In the case of a strong curvature of the embryo, however, it made sense to optionally use a rotation correction before extracting the band of interest along the regression curve (Supplementary Section S1).

2.4 Intensity profile extraction

A main goal was to automatically extract information about the distribution of different neuron types located in the spinal cord. One approach to attain this information was the use of averaged intensity values along each horizontal row of pixels in the cropped image:

$$\mathbf{p}_l[i_y] = \frac{1}{I_x} \sum_{i_x=1}^{I_x} \mathbf{I}_c[i_x, i_y], \quad (2)$$

for all rows $i_y = 1, \dots, 2r + 1$ and all channels $l \in \{\text{RFP}, \text{GFP}, \text{CY3}\}$. This vector \mathbf{p}_l , which we call intensity profile, can be visualized in a common 2D plot (Supplementary Fig. S3). Besides the individual visualization, the vertical alignment of multiple intensity profiles provides an instant overview of possible treatment effects (Fig. 6).

The intensity profiles were calculated individually for all three channels, yielding information about the different neuron types. The detection of local extrema in the intensity profiles was used to identify local maximum intensity information as well as the spatial location where most neurons of a certain type resided. We calculated the area under the curves to get an impression about the overall presence of the fluorescent signal according to

$$\text{ip}_l = \sum \mathbf{p}_l[i_y]. \quad (3)$$

Based on prior knowledge of the spatial location of different neurons, the range for the summation was adjusted accordingly. The sensory neurons, also known as Rohon–Beard cells, were located on the dorsal side of the spinal cord and were visible in the RFP channel ($[1, r + 1]$), the interneurons expressed their fluorescent signal mainly in the GFP channel and were centered above the regression curve ($[0.5r + 1, 1.5r + 1]$). As the fluorescent signal of the RFP expression was visible in the CY3 channel, motoneurons could be observed on the ventral side of the spinal cord ($[r + 1, 2r + 1]$) in this channel using a prolonged exposure time (Section 2.1.4).

2.5 Spatial distribution analysis

The intensity profiles of available control embryos were used to construct a reference intensity profile $\mathbf{p}_{l,\text{ctrl}}$ that represented the average of the original intensity profiles of all control embryos with the associated standard deviation series $\sigma_{l,\text{ctrl}}$ as described by Wolf *et al.* (2006) (Supplementary Section S2). Using these series of statistical properties of the control embryos, it was possible to draw conclusions about the difference between treated embryos to the respective non-treated reference. Therefore, we calculated arithmetic mean of the normalized distance \mathbf{nd}_l , i.e. the variance normalized mean deviation of an arbitrary intensity profile \mathbf{p}_l from the averaged control reference $\mathbf{p}_{l,\text{ctrl}}$:

$$\mu_{l,\text{nd}} = \frac{1}{I_y} \sum_{i_y=1}^{I_y} \mathbf{nd}_l[i_y] = \frac{1}{I_y} \sum_{i_y=1}^{I_y} \frac{|\mathbf{p}_l[i_y] - \bar{\mathbf{p}}_{l,\text{ctrl}}[i_y]|}{\sigma_{l,\text{ctrl}}[i_y]} \quad (4)$$

This feature directly reflects the degree of deviation from the control reference by means of n -fold variance. If the mean value is below one, for instance, on average all sample points of the investigated intensity profile lie within a band of one standard deviation around the respective values of the reference curve. Zero indicates perfect agreement between treatment and reference.

2.6 Neuron distribution histograms

To quantify the possible impact of investigated toxicants on the presence of different neuron types, stained nuclei had to be counted automatically. In this work, we made use of an automated blob detection algorithm similar to the one described in Al-Kofahi *et al.* (2010). Based on a LoG filtering of the images at various scales, spherical objects could be identified in the image:

$$\text{LoG}(\mathbf{x}, \sigma) = \sigma^2 \sum_{i=1}^2 \frac{\partial^2 G(\mathbf{x}, \sigma)}{\partial x_i^2}, \quad (5)$$

with $\mathbf{x} = (x_1, x_2)^\top$ and $G(\mathbf{x}, \sigma)$ being a Gaussian-filtered image with standard deviation σ . To be able to detect differently sized spherical objects, we used a LoG scale-space maximum projection with a discrete step size:

$$\text{LoGMP}(\mathbf{x}, \sigma_{\min}, \sigma_{\max}) = \max_{\sigma_{\min} \leq \sigma \leq \sigma_{\max}} \text{LoG}(\mathbf{x}, \sigma) \quad (6)$$

Owing to the fact that the neurons of the same type have similar sizes, however, we skipped the automatic scale selection performed in Al-Kofahi *et al.* (2010) and only used a set of preselected scales. A LoG-based blob detector recognizes circular objects with radius $r = \sqrt{2} \cdot \sigma$ (Al-Kofahi *et al.*, 2010). By manual inspection of a few representative images, prior knowledge about the appropriate set of scales could be gathered and was used to incline the algorithm to search only for the desired object size (see Table 1). Using the LoG scale-space maximum projection, points of interest were directly identified in the local eight neighborhood of each investigated pixel in the maximum projection instead of using a 3D proximity within neighboring scales as performed, e.g. in Lowe (2004). The intermediate steps and the results of the automated seed detection applied on the RFP channel of one imaged embryo are shown in Supplementary Figure S4. To visualize the distribution of nuclei in the spinal cord region of the embryo, a histogram of nuclei was calculated. Therefore, detected nuclei were binned according to their deviation from the regression curve. Owing to the fact that sensory neurons are only visible on the dorsal side of the regression curve and motoneurons are only observed ventrally, we omitted detected nuclei on the respective other sides to reduce the amount of false positives. An exemplary histogram can be found in Supplementary Figure S5. By locating the mode and maxima of the respective histogram, the spatial occurrence and the quantity of different neuron populations could be estimated.

2.7 Extracted features

We extracted several features that sufficiently characterized the images. The most straightforward features were global statistical properties of the linearized cropped images, such as mean, variance, min, max and median that were attained from the intensity values of each channel ($\bar{\mathbf{g}}_l, \bar{\mathbf{g}}_{l,\sigma}, \bar{\mathbf{g}}_{l,\min}, \bar{\mathbf{g}}_{l,\max}, \bar{\mathbf{g}}_{l,\text{med}}$). Using information about the spatial appearance of the three neuronal populations with respect to the regression curve, the regions of interest could be further specified. The global statistical values for these reduced image areas are equivalent to the integrated intensity profiles ($\text{ip}_l, \text{ip}_{l,\sigma}, \text{ip}_{l,\min}, \text{ip}_{l,\max}, \text{ip}_{l,\text{med}}$). Next to the global intensity-based features, we extracted properties of the intensity profiles as described in Section 2.4. The main characteristics of the intensity profiles were defined by the location and the shape of their peaks. By again exploiting knowledge about the spatial distribution of the three neuron types, peak locations, peak maxima and the peak width could be estimated ($\text{plp}_l, \text{pvp}_l, \text{pwp}_l$). Besides the single features, the intensity profiles could be used to monitor treatment effects. Using an averaged intensity profile of all control embryos with associated local variance ($\mathbf{p}_{l,\text{ctrl}}, \sigma_{l,\text{ctrl}}$), it was possible to define a norm corridor that specified an untreated embryo as shown in Section 2.5. Embryos whose intensity profiles clearly lay outside of this norm corridor should definitely be considered for further investigations (nd). In addition, the seed detection

method described in Section 2.6 was used to calculate an approximation to the number of nuclei for Rohon–Beard cells and motoneurons (nn). Determining the mode of the calculated nuclei histograms could additionally specify locations with high nucleus density (mn). Depending on the biological question to be answered, the list of extracted features could be arbitrarily extended. For the validation of the presented methodology, we used 11 reasonable features (Table 2).

2.8 Implementation details

All described algorithms were implemented in the MATLAB programming language and are available to the public as an extension package for the open-source MATLAB toolbox Gait-CAD (Stegmaier et al., 2012). The results were calculated on a desktop PC equipped with an Intel®Core™i7-2600 CPU at 3.4 GHz with 32 GB of memory installed. For the current amount of data, it was sufficient to process all experiments locally. However, for upcoming high-throughput screens the platform-independent MATLAB implementation and the associated algorithms should also be executable in a distributed computing environment and usage of large-scale storage systems.

3 RESULTS

3.1 Validation of the methods

Manual investigation of the acquired image material offers limited reproducibility and is time consuming even for trained experts. Initial tests unveiled intra-expert and inter-expert variations of the extracted values, leading to the conclusion that even a trained evaluator is not able to satisfactorily reproduce results. Two experienced human experts required ~4 min to extract the specified features for only one image comprising RFP, GFP and CY3 channel. For an exemplary experiment, consisting of 36 images in this case, already 2.5 h of manual investigations were necessary. In contrast to that the software-based evaluation was perfectly reproducible and took <6 s per image on a single processor of a common desktop computer. This corresponds to a 40-fold reduction of evaluation time, allowing to scale up experiments in a high-throughput manner. The quality of identified regression curves and consequently the linearization of the cropped images was manually investigated for all images of the pilot screen. The algorithm successfully extracted and linearized the spinal cord regions in 98% of the images, which corresponded to two misdetections in 108 images. To evaluate the performance of the Rohon–Beard cell detection method described in Section 2.6, manual counting was performed on a subset of 36 images, i.e. three images per concentration, from the previously described experiments. Using the seed detection method

Table 2. Summary of extracted features used for the analysis

Name	Source	Channels	Range	Description
nn	I_c	RFP, CY3	$[0, \infty]$	Number of nuclei ^a
pvp	p_i	RFP, GFP, CY3	$[0, 255]$	Profile peak values ^a
ip	p_i	RFP, GFP, CY3	$[0, 255r]$	Integrated Fluorescence
nd	nd_i	RFP, GFP, CY3	$[0, \infty]$	Norm deviation

^aIndicates that the information extraction for these features is only reasonable on one side of the regression curve.

described in Section 2.6, we detect $90.0 \pm 3.3\%$ of the manually counted Rohon–Beard cells in 36 separately evaluated test images. Furthermore, properties such as regional intensity, spatial positioning of intensity peaks with respect to the spinal cord center and regional properties of neuron populations were qualitatively assessed by a human expert to rate the automatically provided results of the software and to perform a coarse classification into the classes *no effect*, *slight effect* and *strong effect*. The manually performed qualitative assessment of the effects by direct visual comparison of the images unveiled strong decrease interneuron intensity for concentrations $>0.5 \mu\text{M}$ of LY-411575 and a slightly increased number of motoneurons. On this basis, a hypothetical classification of the different treatment concentrations was performed. Identified classes were *no effect* for control embryos, *slight effect* for $0.5 \mu\text{M}$ of LY-411575 and a *strong effect* for higher concentrations using the 36 test images. We achieve 98% correct classifications for this task on 108 investigated images using the leave-one-out cross validation with a Bayes classifier that was trained with one ANOVA selected feature and the hypothetical class associations. Using ANOVA for feature selection, pvp_{GFP} turned out to be the most discriminative value for the observed effect.

3.2 Biological results

The aim of the performed exposure experiment was to validate the automated quantification and characterization of the three main neuron types in the spinal cord of zebrafish embryos. Table 3 summarizes a selection of features that were extracted from 108 acquired images resulting from a 3-fold replicated dose-response experiment (Supplementary Figs. S6 and S7). Using the Wilcoxon rank sum test, significant differences of treatment and control were identified (Supplementary Table S2). Based on visual inspection of the 36 test images, we suspected a constant number of Rohon–Beard cells and the corresponding RFP intensity, decreased interneuron intensity in the GFP channel and a

Table 3. Identified values for a subset of the most prominent features, averaged over all embryos per different concentrations of LY-411575. GFP intensity is significantly decreased for increased treatment concentrations

Concentration (μM)	ip_{GFP}	pvp_{GFP}	nd_{GFP}
Ctrl	49.8 ± 2.2	112 ± 7.3	0.75 ± 0.4
0.5	42.2 ± 1.8^a	74.4 ± 9.1^a	1.53 ± 0.3^a
1.0	32.0 ± 1.4^a	49.0 ± 3.3^a	2.36 ± 0.4^a
2.0	27.7 ± 1.2^a	38.1 ± 3.0^a	2.93 ± 0.2^a
3.0	27.9 ± 0.7^a	38.5 ± 1.2^a	2.85 ± 0.3^a
4.0	27.9 ± 1.5^a	37.9 ± 2.7^a	3.00 ± 0.3^a
5.0	25.5 ± 0.7^a	32.7 ± 1.3^a	3.27 ± 0.2^a
6.0	25.2 ± 1.0^a	31.0 ± 2.2^a	3.29 ± 0.3^a
7.0	25.4 ± 0.9^a	30.6 ± 1.9^a	3.16 ± 0.2^a
8.0	25.6 ± 1.0^a	31.3 ± 1.3^a	3.12 ± 0.4^a
9.0	25.0 ± 0.8^a	30.2 ± 1.4^a	3.29 ± 0.2^a
10.0	25.3 ± 0.8^a	30.7 ± 1.6^a	3.18 ± 0.2^a

^aSignificant differences to the control embryos were identified by the Wilcoxon rank sum test.

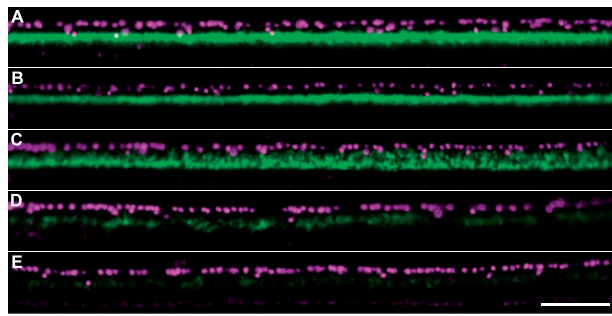


Fig. 3. Overlay of RFP and GFP images obtained for control embryos (A) and embryos treated with LY-411575 at concentrations of 1.0nM (B), 0.5μM (C), 1.0μM (D) and 10.0μM (E). Scale bar = 100μm

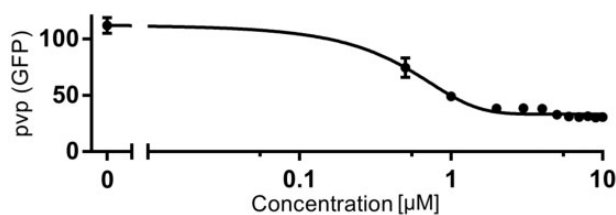


Fig. 4. The dose-response curve shows an exponential decrease of the GFP peak intensity (pvp_{GFP}) for increased LY-411575 concentration. Similar behavior is observable in most intensity-based features of the GFP channel

slightly increased number of motoneurons and increased intensity for increasing concentrations. Figure 3 shows a selection of linearized original images at various concentrations of LY-411575.

A significant reduction of interneuron intensity compared with the control embryos could be observed for all GFP-related features (e.g. ip_{GFP} , pvp_{GFP} or nd_{GFP}). As shown in Figure 4, the GFP peak intensity value exponentially decreased for increased treatment concentrations. Additionally, a tendency to an increased number of motoneurons could be observed in the CY3 channel for LY-411575 concentrations $\geq 4.0\mu M$ (e.g. pvp_{CY3} , nd_{CY3}). A more detailed view on the extracted information was attained by an overlay plot of all intensity profiles (Fig. 5 and Supplementary Fig. S10). The shaded area in Figure 5 indicates the norm corridor, i.e. profiles that clearly lie outside this area should be investigated in more detail. It could clearly be observed that LY-411575 concentrations higher than $0.5\mu M$ caused decreased interneuron intensity. In the CY3 channel, a rather subtle and unspecific alteration of motoneuron appearance was detected. However, besides significant norm distance deviations of nd_{RFP} for LY-411575 concentrations $\geq 0.5\mu M$, no significant effect was observed for the remaining features of the RFP channel (ip_{RFP} , pvp_{RFP} and nn_{RFP}). This corroborated the manual evaluation of the RFP images that neither the number nor the intensity of sensory neurons was affected by the treatment. The decreased GFP intensity was caused by a decreased number of interneurons. As the counting of interneurons was not possible using the considered image material, this was confirmed using 3D image stacks (CTRL and

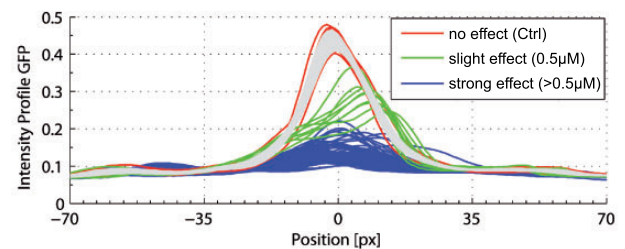


Fig. 5. Overlay of 108 intensity profiles for the GFP channel. Curves are colored according to the classification. A significant reduction of the interneuron intensity can be observed. A slight effect can be seen for a concentration $0.5\mu M$ and a strong effect for higher concentrations of LY-411575. The norm corridor is foreshadowed by the gray surface

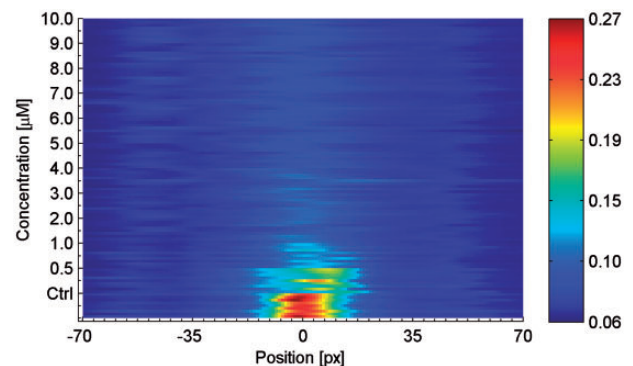


Fig. 6. Image plot of 108 aligned color-coded intensity profiles for the GFP channel. Curves are sorted by increasing LY-411575 concentration. A ventrally oriented shift of the interneuron peak location can be observed for a treatment of $0.5\mu M$

$10\mu M$). The images were acquired using confocal microscopy and were manually evaluated by two trained experts (Supplementary Section S3).

Figure 6 displays a vertical alignment of all 108 intensity profiles of the GFP channel images. This visualization provided a convenient way to compare all intensity profiles in one comprehensive image. More clearly than in the previous depictions, a ventrally oriented shift of the GFP peak location could be observed for a concentration of $0.5\mu M$. An additional experiment with LY-411575 concentrations from 1nM to $10\mu M$ confirmed the dose-response estimation (Supplementary Figs. S7, S8 and S9). A comprehensive biological analysis of the observed effects is up to prospective work, as the focus of this article is put on the demonstration of the used methods.

4 DISCUSSION

Although diverse zebrafish-related implementations already exist, up to the best of our knowledge none of the available methods and software packages was directly applicable to the challenges observed in our experiment. To analyze the double-transgenic zebrafish line, information fusion of content found in different fluorescent channels was required to attain profound insights into treatment effects on neural development. The most similar methodology to ours with a focus on automated

quantification of the spinal cord region of zebrafish was described by Chen *et al.* (2011) and Liu *et al.* (2006, 2008b). The embryo alignment method described by Chen *et al.*, however, was only applicable if the spinal cord exhibited a clear linear shape. We overcame the linear shape restriction by use of third order polynomial regression to extract a linearized and rotation-corrected crop image of the spinal cord region. In contrast to the blob detection described by Liu *et al.* which had a high error rate for sensory neuron detection, our method uses a LoG scale-space based blob detection method to identify approximate centroids of stained nuclei in different fluorescence channels (Al-Kofahi *et al.*, 2010). Additionally, we present a method to quantify intensity profiles that reflect distributions of different neuron types present in the spinal cord region. This method proved to be highly robust against misdetections and provided an instant impression of phenotype alterations for sensory neurons, interneurons and motoneurons on chemical exposure. We exploited the potential of spatial distribution methods to analyze and compare these intensity profiles between different embryos and treatments. The proposed methods were validated on a manually assembled ground truth from a dose-response experiment with a Notch receptor inhibitor.

The spinal cord detection robustly identified regression curves in 98% of the 108 images that could successfully be used as a reference for further processing. The erroneous regressions, however, were still usable as the spinal cord was matched for the most part and only subtle deviations in border regions of the images occurred. The problems were mostly caused by undesired imaging effects such as dirt, dust or overexposed regions. In some of these cases, the binarization of the images had difficulties in extracting the correct spinal cord region. This issue might be solved by an explicit shape model that describes the silhouette of the desired region.

Although the described method reliably extracted sensory neurons from the given images, it should be noted that it performed sub-optimally for quantification of motoneurons in the CY3 channel. Despite the inaccurate counts, however, the method could still be useful to estimate qualitative tendencies of the motoneuron distributions. To attain reliable information about motoneuron appearance, the intensity profile information described in Section 2.4 is recommended. The limited capabilities of the algorithm to differentiate between individual motoneurons were mainly caused by insufficient image resolution and could vanish as soon as the magnification level is increased. Additionally, more sophisticated nucleus detection methods such as described by Liu *et al.* (2008a) or even a fusion of nuclei detected by different algorithms could be performed. Assuming a perfectly linearized crop image, intensity profile-based methods and values were both informative and robust against misdetections. Using the norm corridor of control embryo intensity profiles, a deviation of treated embryos from the norm could be easily detected either by intensity profile visualizations as depicted in Figures 5 and 6 or quantitatively by using measures such as the norm deviation described in Equation (4). The classification results showed that the features were well suited to distinguish between different treatment effects observed in the considered experiment.

Owing to the high versatility of the extracted features, investigated embryos or rather the possible treatment effects could be

quantitatively characterized. Especially in experiments with a high number of samples, the demonstrated automation methods drastically facilitate and accelerate the evaluation of experiments. Besides the decreased time spent for analysis, the results are highly reproducible and offer quantitative assessment of features of interest. Furthermore, the automated analysis enables the extraction of image properties that are beyond the capabilities of visual inspection by humans, such as slight intensity variations or pixel accurate feature localization. The described methods are specialized for a distinct type of biological investigation. However, because of the modular design of all algorithms implemented in MATLAB, an adaption to similar problems and the reuse of implemented methods should be straightforward. In upcoming work, the developed tools will be used to automate the analysis of large screens containing a variety of chemical compounds with unknown effects on the neural development of zebrafish.

Funding: EU IP ZF-Health; (FP7-HEALTH-2007-B2); NeuroXsys; the Interreg network for synthetic biology in the Upper Rhine valley (NSB-Upper Rhine); the BMBF funded network EraSysBio; and the Helmholtz Association.

Conflict of Interest: none declared.

REFERENCES

- Al-Kofahi, Y. *et al.* (2010) Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Trans. Biomed. Eng.*, **57**, 841–852.
- Blader, P. *et al.* (2003) Multiple regulatory elements with spatially and temporally distinct activities control neurogenin1 expression in primary neurons of the zebrafish embryo. *Mech. Dev.*, **120**, 211–218.
- Cao, Y. *et al.* (2009) Chemical modifier screen identifies HDAC inhibitors as suppressors of PKD models. *Proc. Natl Acad. Sci. USA*, **106**, 21819–21824.
- Chen, S. *et al.* (2011) Automated analysis of zebrafish images for phenotypic changes in drug discovery. *J. Neurosci. Methods*, **200**, 229–236.
- Detrich, H.W. *et al.* (2011) *The Zebrafish: Disease Models and Chemical Screens*. 3rd edn. Academic Press, Waltham, MA, USA.
- Eliceiri, K. *et al.* (2012) Biological imaging software tools. *Nat. Methods*, **9**, 697–710.
- Fauq, A.H. *et al.* (2007) A multigram chemical synthesis of the γ -secretase inhibitor ly411575 and its diastereoisomers. *Bioorg. Med. Chem. Lett.*, **17**, 6392–6395.
- Flinn, L. *et al.* (2008) Zebrafish as a new animal model for movement disorders. *J. Neurochem.*, **106**, 1991–1997.
- Gehrig, J. *et al.* (2009) Automated high throughput mapping of promoter-enhancer interactions in zebrafish embryos. *Nat. Methods*, **6**, 911–916.
- Jeanray, N. *et al.* (2012) Phenotype classification of zebrafish embryos by supervised learning. *Toxicol. Lett.*, **211**, S152.
- Kanungo, J. *et al.* (2011) Ketamine induces motor neuron toxicity and alters neurogenic and proneural gene expression in zebrafish. *J. Appl. Toxicol.*, **33**, 410–417.
- Karlsson, J. *et al.* (2001) Generating transparent zebrafish: a refined method to improve detection of gene expression during embryonic development. *Mar. Biotechnol.*, **3**, 522–527.
- Liu, R. *et al.* (2012) Automated phenotype recognition for zebrafish embryo based *in vivo* high throughput toxicity screening of engineered nano-materials. *PLoS One*, **7**, e35014.
- Liu, T. *et al.* (2006) Computerized image analysis for quantitative neuronal phenotyping in zebrafish. *J. Neurosci. Methods*, **153**, 190–202.
- Liu, T. *et al.* (2008a) An automated method for cell detection in zebrafish. *Neuroinformatics*, **6**, 5–21.
- Liu, T. *et al.* (2008b) Zfiq: a software package for zebrafish biology. *Bioinformatics*, **24**, 438–439.
- Lowe, D.G. (2004) Distinctive image features from Scale-Invariant keypoints. *Int. J. Comp. Vision*, **60**, 91–110.
- Mikut, R. *et al.* (2013) Automated processing of zebrafish imaging data - a survey. *Zebrafish*, **10**, 401–421.

- Otsu,N. (1979) A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.*, **9**, 62–66.
- Peravali,R. *et al.* (2011) Automated feature detection and imaging for high-resolution screening of zebrafish embryos. *BioTechniques*, **50**, 319–324.
- Ronneberger,O. *et al.* (2012) ViBE-Z: A framework for 3D virtual colocalization analysis in zebrafish larval brains. *Nat. Methods*, **9**, 735–742.
- Shin,J. *et al.* (2007) Notch signaling regulates neural precursor allocation and binary neuronal fate decisions in zebrafish. *Development*, **134**, 1911–1920.
- Stegmaier,J. *et al.* (2012) Information fusion of image analysis, video object tracking, and data mining of biological images using the open source MATLAB toolbox Gait-CAD. *Biomed. Eng.*, **57** (Suppl.1), 458–461.
- Vogt,A. *et al.* (2009) Automated image-based phenotypic analysis in zebrafish embryos. *Dev. Dyn.*, **238**, 656–663.
- Westerfield,M. (2007) *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio rerio)*. 5th edn. University of Oregon Press, Eugene, OR.
- Wolf,S. *et al.* (2006) Automated feature assessment in instrumented gait analysis. *Gait Posture*, **23**, 331–338.
- Yang,L. *et al.* (2010a) Methyl mercury suppresses the formation of the tail primordium in developing zebrafish embryos. *Toxicol. Sci.*, **115**, 379–390.
- Yang,L. *et al.* (2010b) Regulatory interactions specifying kolmer-agduhr interneurons. *Dev.*, **137**, 2713–2722.