ZEBRAFISH
Volume 00, Number 00, 2013

DOI: 10.1089/zeb.2013.0886

Review Article

Automated Processing of Zebrafish Imaging Data: A Survey

Ralf Mikut,[1] Thomas Dickmeis,[1] Wolfgang Driever,[2] Pierre Geurts,[3] Fred A. Hamprecht,[4] Bernhard X. Kausler,[4] María J. Ledesma-Carbayo,[5] Raphaël Marée,[3] Karol Mikula,[6] Periklis Pantazis,[7] Olaf Ronneberger,[2] Andres Santos,[5] Rainer Stotzka,[1] Uwe Strähle,[1,4] and Nadine Peyriéras[8]

Abstract

Due to the relative transparency of its embryos and larvae, the zebrafish is an ideal model organism for bioimaging approaches in vertebrates. Novel microscope technologies allow the imaging of developmental processes in unprecedented detail, and they enable the use of complex image-based read-outs for high-throughput/high-content screening. Such applications can easily generate Terabytes of image data, the handling and analysis of which becomes a major bottleneck in extracting the targeted information. Here, we describe the current state of the art in computational image analysis in the zebrafish system. We discuss the challenges encountered when handling high-content image data, especially with regard to data quality, annotation, and storage. We survey methods for preprocessing image data for further analysis, and describe selected examples of automated image analysis, including the tracking of cells during embryogenesis, heartbeat detection, identification of dead embryos, recognition of tissues and anatomical landmarks, and quantification of behavioral patterns of adult fish. We review recent examples for applications using such methods, such as the comprehensive analysis of cell lineages during early development, the generation of a three-dimensional brain atlas of zebrafish larvae, and high-throughput drug screens based on movement patterns. Finally, we identify future challenges for the zebrafish image analysis community, notably those concerning the compatibility of algorithms and data formats for the assembly of modular analysis pipelines.

Introduction

Over the last 20 years, the zebrafish (*Danio rerio*) has become one of the most important model organisms for *in vivo* studies of vertebrates in biology, both in fundamental research and for biomedical applications.[1,2] The genetic similarity between zebrafish and mammals, its small size, fast development, cheap husbandry, and its relevance for replacing mammals in animal testing make the zebrafish the model of choice in a large range of research projects. Crucially, the transparency of the embryos makes them amenable to deep *in vivo* imaging. Further reasons for the attractiveness of the model include the almost complete genome sequence (www.sanger.ac.uk/Projects/D rerio/) and an ever-growing number of transgenic and mutant fish lines, allowing detailed *in vivo* studies of gene regulation and function.

These zebrafish characteristics are exploited to address important questions in genetics, developmental biology, drug discovery, toxicology, and biomedical research. Notably, zebrafish models exist for a broad range of human diseases, for example, cardiovascular diseases,[3] cancer,[4] or movement disorders.[5] Imaging approaches include three-dimensional time lapse (3D+t) imaging of embryos by confocal or multi-photon laser scanning microscopy, which generates different signals for image analysis, including two-photon-excited fluorescence, second harmonic generation (SHG), and third harmonic generation (THG).[6] Another technique used in 3D+t imaging is selective plane illumination microscopy

[1]Karlsruhe Institute of Technology, Karlsruhe, Germany.
[2]University of Freiburg and BIOSS Centre for Biological Signalling Studies, Freiburg, Germany.
[3]GIGA & Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium.
[4]Heidelberg Collaboratory for Image Processing (HCI), University of Heidelberg, Heidelberg, Germany.
[5]Biomedical Image Technologies, Universidad Politecnica de Madrid & CIBER-BBN, Madrid, Spain.
[6]Slovak University of Technology, Bratislava, Slovakia.
[7]Department of Biosystems Science and Engineering (D-BSSE), ETH Zürich, Zürich, Switzerland.
[8]CNRS Gif-sur-Yvette, Gif-sur-Yvette, France.

(SPIM),[7] which is also called light sheet fluorescence microscopy (LSFM). In addition, video series are recorded to observe developmental or behavioral changes over longer periods of time in larvae and adults. Stitched high-resolution images, in which many high-resolution images with a small field of view of the same object are fused, are employed to observe features at cellular or subcellular resolution. Finally, high-throughput studies require an analysis of the effects of thousands of potential drugs or toxicants. With the use of automated microscopes and robotic handling, high-throughput screening (HTS) projects generate thousands of images, leading to Terabytes of data. Consequently, there is an increasing demand for automated image processing to generate quantitative results and to avoid time-consuming manual analyses. Most often, the required image analysis is far beyond the scope of commercially available solutions, such as the built-in software of microscopes.

Therefore, the development of dedicated image processing methods has become a serious bottleneck in the full exploitation of the information contained in the acquired image sets. Many custom-made and nongeneric solutions have been developed to answer specific questions. However, not all solutions have been published, and transfer of the available methods to other projects is sometimes difficult due to missing standardization and strong restrictions on input formats, software tool deployment, and so on. Nevertheless, these solutions have an enormous potential to support the analysis of current and future image-based phenotypic studies and to avoid parallel developments that tend to "reinvent the wheel."

To promote a better integration of image analysis software with microscopy imaging strategies, in particular in the context of large-scale screening projects, it is crucial to comprehensively survey existing image analysis solutions and to rate their potential for application to other problems. Furthermore, one needs a better understanding of the requirements of both biologists and computer scientists in image analysis that should then guide how precisely data are acquired and analyzed. This would define, for example, a minimal raw data quality that is necessary for image analysis, or set acceptable error margins for biological feature detection. Clarifying such issues would greatly benefit the design of future screens, of the image analysis solutions, and of their interfaces. Finally, a long-term goal would be the design of standardized workflows for data acquisition, data handling, image processing, and data analysis, integrating many existing generic and zebrafish-specific modules.

The aim of this survey paper is to support this networking and sharing process by giving an overview of the state of the art and by opening a discussion about future steps toward these goals in both the zebrafish and the image processing communities. First, Requirements for automated processing of zebrafish imaging data section of this paper describes requirements for automatic image processing of zebrafish data, especially with regard to high-throughput applications. In State-of-the-art: methods and tools section, a survey of existing methods, tools, and general standardization activities in biological imaging is given, followed by an overview of tools in Available software and computation tools section, and of recent applications in developmental biology, physiology, pharmacology, and toxicology in Applications based on automated image processing section. Finally, we have Discussion: open questions and future trends section.

Requirements for Automated Processing of Zebrafish Imaging Data

Automated image analysis is not limited to any type of microscopy or any kind of assay. Image processing is going to revolutionize the way we perform and exploit microscopy imaging in many ways. It will allow enormous gains in time, reproducibility, and objectivity of the results; enable the full exploitation of the information in the raw data; and make them amenable to further statistical or modeling studies. Simple two-dimensional (2D) data acquired in epifluorescence can be subjected to image analysis as well as 3D+t data from multiphoton or SPIM microscopy or high-throughput 3D data from phenotypic drug screening. Thus, the requirements of the zebrafish community for automated processing of imaging data are not confined to any particular application. However, high-throughput assays cannot even be designed without strict standardization—often implying robotization—and image processing. In this section, we, therefore, focus on the requirements of the community to perform -scale screening assays.

Pipelining high-throughput assays

The concept of HTS means automating all the steps of the assays, building complete workflows comprising fish breeding, specimen handling, mounting, treatment with substances (e.g., toxicants or drug candidates), or other conditions (e.g., light or temperature pulses), automated microscopy imaging, data management, image processing, and evaluation of the results (Fig. 1).

The type of assay and workflow to be designed depends on the biological question; for example, toxicological effect on cells, embryos, larvae, adult fish, drug-specific phenotypic effects at the cellular or subcellular level, and mutants exhibiting behavioral changes as larvae or adult fish. Some typical readouts for such questions could be as follows:

- place preference, swimming speed,
- detection of embryonic lethality,
- morphological changes in embryos/larvae,
- changes in reporter gene expression in embryos/larvae,
- the number of labeled cells in specific tissues, and
- movement of cells and tissues in 3D+t imaging.

Given this diversity of assays and readouts, the corresponding requirements in terms of specimen manipulation, imaging, data management, and image processing are equally diverse: Thus, data acquisition can employ 2D, 3D, or 3D+t imaging on either fixed or live specimens, from one-cell stage to adults, cell cultures, or tissue explants, observed at different levels of organization ranging from the molecular level to the whole organism interacting with its environment. Once the context of the experiment has been chosen, the design of the imaging protocol needs to be adapted to the expected results and to the requirements for image processing.

The spatiotemporal resolution of the acquired images, the field of view, the signal-to-noise ratio (SNR), and the survival or stability of specimens under the assay conditions are crucial parameters. Thus, one main point is to choose imaging conditions according to the size of the smallest structures that

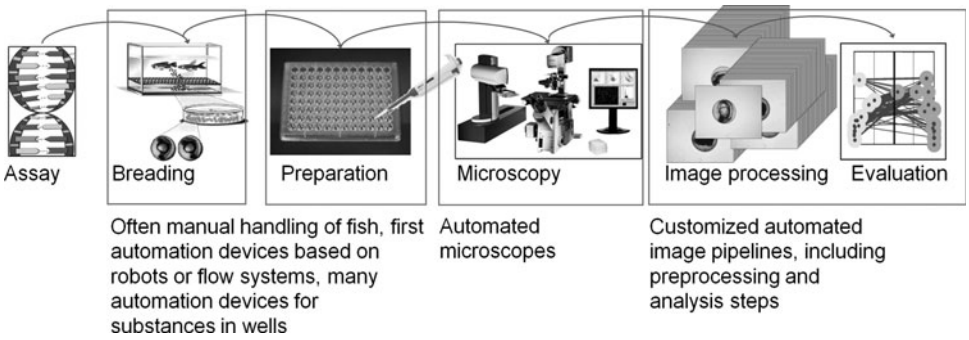

FIG. 1. A typical workflow for HTS with the zebrafish as a model organism (modified from Alshut *et al.*[8]).

need to be separated. The resolution required in the x-, y-, and z-direction needs to be defined, as well as the size of the volume that needs to be explored. Sometimes, a single z-plane may be sufficient, but at least a few z-slices are needed if one wants to generate a reconstructed focused 2D image of a specimen via postprocessing of focus planes. Many z-slices are needed for a full reconstruction of 3D structures with laser scanning microscopy or SPIM. In addition, the SNR and the dynamic range of the image are relevant to define the resolution of gray values. The influence of other disturbances (e.g., external light sources) should be negligible. Furthermore, the temporal resolution of image acquisition needs to fit the scale of the events that need to be followed. For instance, short time intervals are needed for a quasi-static imaging for cell tracking, fish tracking, or heartbeat detection; while longer intervals are sufficient to quantify temporal changes of defined morphological parameters. Sometimes, difficult compromises regarding various trade-offs between the parameters have to be found in order to reach the aim of the assay. For example, in 3D + t imaging, increasing the image resolution will usually reduce the achievable temporal resolution.

In total, the following parameters have to be defined:

- G: number of bytes that describe gray values per pixel (usually defined by the cameras or microscopes, where 1 byte corresponds to $2^8 = 256$ gray values),
- X * Y: image size in pixels (usually defined by the cameras or microscopes),
- Z: number of slices in z-direction,
- C: number of acquired brightfield, fluorescent, or other signal channels,
- T: number of time points,
- D: number of different treatments combined in an experiment (e.g., to investigate the effects of different drugs or toxicants for single or multiple dilutions on wildtype fish, mutants, or transgenic lines), and
- N: number of fish for each treatment or combination of treatments.

As a first approximation, such an experiment generates a dataset with a total size of P acquired pixels:

$$P = G * X * Y * Z * C * T * D * N \quad (1)$$

(assuming identical values of G, X, Y, Z, C, T, and N for each treatment combination and assuming one fish per image).

Automatic analysis is especially useful for experiments with large T, D, and N. Simple screens typically generate only a few Gigabytes (e.g., $G = 1$ resulting in 256 gray values, $X = Y = 1024$, $C = 3$ channels, $Z = 1$, $T = 1$, D = 100 drugs, $N = 8$ fish per drug; $P = 2.5$ GB) that can be easily handled on a local hard disk. In contrast, larger 3D + t screens ($Z \gg 1$, $T \gg 1$) can produce data sizes of many Terabytes of data, for example, $G = 2$ resulting in 65536 gray values, $X = Y = 4096$, $C = 2$, $Z = 500$ slices, $T = 300$, $D = 5$ drugs, $N = 8$ fish per drug; $P = 402$ TB. Such screens have stronger demands in terms of data management, as discussed in Data management requirements section.

The possibility to perform automated image analysis relies on a number of *a priori* conditions. Images should be acquired in a standard image format that can be directly processed by many image processing tools or should be easily converted into other standard formats. An experienced observer should be able to manually extract the targeted biological information from the acquired images. The methodology of analysis developed by this experienced observer should be transferable into an automatic algorithmic strategy. Image processing algorithms are expected to extract the desired biological information with acceptable error rates. In principle, these algorithms should offer a compromise between accuracy, development time, resource efficiency with regard to memory usage and computing time, robustness against variations in images due to disturbances or noise, standardization aspects, and generalization abilities. The algorithms should run in a reasonable computing time on commercially available computers.

Metadata requirements

All images should be described by so-called metadata, explaining the spatiotemporal position of the image, acquisition details, and all the experimental conditions that are relevant for the study in a computer-readable form (e.g., coded in text or extensible markup language [XML] files, filenames, and directory names). These metadata should be sufficient to repeat the experiment.

Data without metadata are useless, particularly if the data have to be stored over a long period of time or are processed through various processing steps. The role of metadata becomes even more important with the reinforcement of quality control rules that ask for data availability over more than a decade. The needs for common and sustainable data formats, including metadata, are well accepted in the community of biologists, but existing metadata standards and exchange protocols such as Open Archive Initiative Protocol for Metadata Handling (OAI-PMH)[9] or for other types of experiments (see the overview by Taylor *et al.*[10]) have not yet been disseminated.

The complete protocol, including acquisition conditions of all images, should be clearly defined to guarantee reproducible experiments and to avoid image analysis bias.[11] It should include, for example, information about the number of fish, fish strains, anesthetization, mounting conditions, age, state (hatched, dechorionated or in chorion), imaging parameters, microscopy setup parameters, light source characteristics (including power intensity), and so on. Defining the appropriate annotation framework, linking it to the data, as well as processing and analyzing it are issues that deserve careful attention and should be placed high on the agenda of the zebrafish imaging community.

Data management requirements

Managing and handling the data deluge produced by automated microscopes requires totally new methods and technologies. For example, in only 10 h, an SPIM observing the development of a zebrafish embryo may produce 7 TB of raw data, resulting in an average data rate of 200 MB/s before compression. Such data rates can be easily handled in state-of-the-art data acquisition computers that are equipped with a RAID-10 storage system.[12] However, storage of the results produced in 5 days will require a total of 105 TB disk space in this data acquisition computer, resulting in high costs and a high failure probability. Therefore, close collaborations with data centers are mandatory to handle the raw data. Data centers will support the data, producing scientists in all fields of large-scale data management and processing, from data transfer to long-term archiving of the data (Fig. 2).

Complex data processing on the data acquisition computer is often not feasible, as the amount of data streams limits the overall throughput. Special data-intensive computing environments that are able to process Terabytes of data in minutes are required but often not available locally. Thus, data processing and analysis workflows should run on the infrastructure of available data centers. Modern data centers can be accessed from distributed distant sites, allowing workflows to start automatically after a dataset is integrated into the system.[13] Transferring the "preprocessed" data, which are now fitting into well-accepted standard data and metadata formats, to such data centers requires data lines and software systems that are capable of reaching the required data rates.

Easy and fast access to the raw data itself, to the processed data, and to the metadata is essential for further scientific evaluation of what has been produced. Teams of scientists distributed over many locations need to exchange their data, either raw or processed, as well as their scientific results. Such data sharing and networking of communities within their digital campus will likely be a crucial part of the scientific infrastructure in the future.

State-of-the-Art: Methods and Tools

Concepts for automated processing of biological images

After describing general requirements for the use of automated image analysis in high-throughput screens and other data-intensive imaging applications, we now turn to a more detailed description of the methods and tools that are available for automated image processing. The major issue that we want to address here is the adaptation of the image data to its processing pipeline and vice versa. Adequate formats and processes will require a close interdisciplinary collaboration between biologists and computer scientists relying on shared expertise and practice.

In the next few sections, image processing methods are subdivided according to the anatomical level of analysis, dimensionality, and type of data to be processed.

"Level" describes if the analyzed objects are

- whole organisms (adult fish, larvae or embryos),
- organs or tissues, or
- cells.

"Dimension" is related to the captured images in the dataset described by Equation (1). Most methods and applications can be categorized by

- 2D: two dimensional with X and $Y >> 0$,
- 2D+t: two dimensional with an additional time recording in form of video images with X, Y, and $T >> 0$ (where T is the total number of time points, and t indicates the temporal dimension of the data),

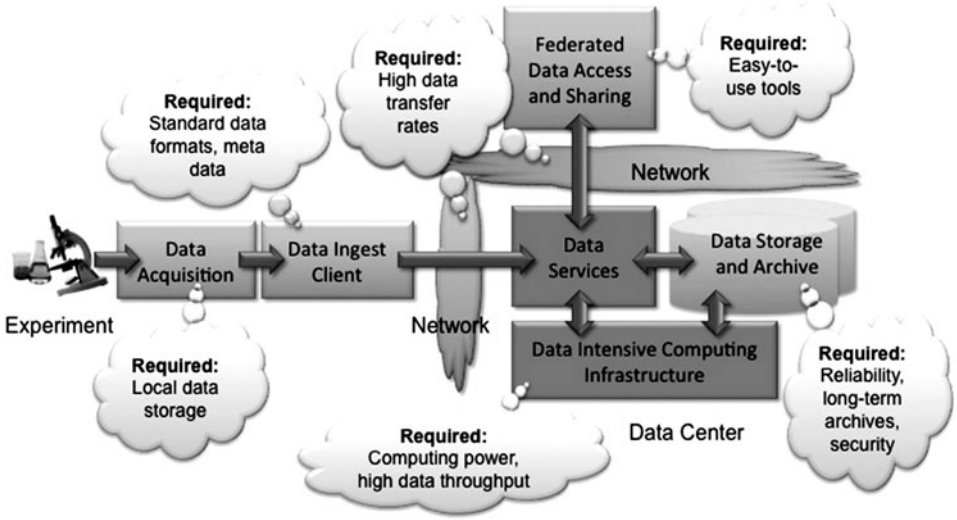

FIG. 2. Data workflow in a large high-throughput screen.

- 3D: three dimensional describing spatial configurations in detail with X, Y, and $Z >> 0$,
- 3D + t: time recordings of spatial configurations with X, Y, Z, and $T >> 0$.

In most cases, applications with only a few spatial slices (typically $Z = 2 \ldots 10$) are mapped by preprocessing methods, for example, by generating an extended focus, to 2D images. Applications with a few fluorescence (or other signal) channels (typically $C = 2 \ldots 3$) or time points (typically $T = 2 \ldots 3$) are analyzed separately by such 2D methods, or by 3D methods if $Z >> 0$. In this context, also data of simple bioluminescence/fluorescence readers or of cell counters can be interpreted as a kind of degenerated images with the dimension 2D + t, 2D, or even 1D + t. Here, time series of cells or larvae passing through a bioluminescence/fluorescence sensor are analyzed.

An overview about method-oriented versus application-oriented papers using these categorizations and further aspects, including fish age and number, algorithm and application focus, as well as used tools, is given in Table 1. This list can also be found in the Excel table zip_tools.xls in the additional material. Regularly updated versions of this list will be available at (http://sourceforge.net/projects/zebrafishimage/).

Based on the "type of data" processed, image processing methods can be divided into two types:

- Preprocessing methods deliver "better" raw data. Typical tasks, with increasing complexity, are noise reduction, correction of attenuation (the gradual loss in light intensity when passing through tissues), correction of inhomogeneous illumination, stitching of mosaic images, extended focus algorithms to fuse z-slices with different focus planes to a focused 2D image, and so on. Such methods are usually application independent and can be found in many image processing toolboxes.
- Analysis methods extract biologically relevant information from the images. These methods include different operations to find object boundaries, such as for cells or tissues, to compute features such as brightness and shape parameters, to fuse segments (or patterns) from different fluorescent channels (or other signals) or from different specimens into one reference view over time, and to track segments (i.e., objects) over time. Such operations can be application independent (many segmentation operations) or adapted to fish-specific characteristics (tissue detection, landmark-based fusion of data from different specimens etc.).

Image processing algorithms consist of pipelines of many single image operations, including preprocessing and analysis methods with problem-adapted parameters. We now discuss a few examples of different methods that are applied to specific zebrafish imaging problems.

Preprocessing methods

The goal of preprocessing methods is to provide "better" raw data quality, which means usually the compensation of known limitations of the image acquisition process. It should be noted that "better" here does not mean "looking nicer," but rather "better suited for the automated extraction of quantitative information." Three main principles can be used for preprocessing:

(1) Previous knowledge about the imaging process may reveal typical patterns of out-of-focus signal contributing to a focus plane or of deformations of the signal shape. An example for preprocessing methods correcting such problems are deconvolution techniques. These techniques reduce the blur that is introduced from the microscope optics, using a mathematical model of the imaging process.
(2) Assumptions about the structures in the image. Denoising techniques, for example, can reduce noise by assuming that certain pixel patterns or structures are more likely than others in a noise-free image. For instance, Luengo-Oroz *et al*.[52] proposed a multidimensional (3D + time) filter that uses the inherent redundancy of the temporal dimension to improve the membrane delineation in all directions (Fig. 3A).
(3) Fusion of multiple recordings of the same sample by applying assumptions about imaging process and structures in the image. Several methods apply this principle.

High dynamic range (HDR) techniques increase the dynamic range of an image by fusion of two or more recordings with different illumination times or laser intensities. The dynamic range of an imaging system is described by dividing the largest possible gray value (before overexposure) by the lowest possible gray value above zero or above the noise level. In a study by Ronneberger *et al.*,[63] this technique was used to improve signal quality in confocal stack recordings extending over the whole depth (400 μm) of the larval brain (see Fig. 4 for an example).

Stitching[76] allows one to increase the field of view without decreasing the resolution by fusing multiple recordings taken at different positions (tiles, see frontal and rear parts in Fig. 5A). An example for this from neuroscience is the tracking of axons extending over an anatomical area that covers several fields of view.[69]

Finally, Multi-View Fusion increases the resolution in the axial direction (e.g., fusion of multiple views in light sheet microscopy.[70,77,78] For example, Rubio-Guivernau *et al*.[64] proposed a wavelet-based method to combine the information from multiple views in a single volume (Fig. 3B). Fusion can also compensate for the light attenuation in thick samples by integrating two recordings that are taken from opposite sides (two view attenuation correction, see Fig. 5A, dorsal and ventral parts).[63]

As an example, a combination of the three techniques (HDR, stitching, and two-view attenuation correction) was used to record the full brain of 2–4-day old zebrafish larvae at cellular resolution with a standard confocal microscope (Fig. 5).[63]

Methods for image data analysis

Fish tracking, behavior patterns, and social behavior. Various approaches have been proposed to automatically extract movement of adult fish or hatched larvae as well as other behavioral parameters using automated video analysis in different experimental settings. The targeted parameters were, for example, velocity, total distance traveled, or inter-fish distance for social behavior analysis, and had to be extracted from data acquired, for example, in wells or Petri dishes (2D + t), in tanks with one camera (2D + t), or with two cameras for a 3D reconstruction (3D + t).[19] However, some of

Table 1. Categorization of Method- and Application-Oriented Papers

Paper	Application or method-oriented	Imaging modality	Level	Dimension	Fish age and screen size	No. of fish	Algorithms: depth of description (brief or in detail) and outcome	Application	Tools used
Alshut *et al.*[15]	M	Screening WFM	Whole organism	2D	34 hpf	384	B Phenotype detection	Toxicology	MATLAB, Gait-CAD
Alshut *et al.*[8]	M	Screening WFM	Whole organism	2D	48 hpf	576	B Phenotype detection	Toxicology	MATLAB, Gait-CAD
Berghmans *et al.*[16]	A	Camera	Whole organism	2D + t	6 dpf	>1000	D Behavior patterns	Drug discovery	EthoVision
Bhat and Liebling[17]	M	Microscope	Tissue	2D + t	48 hpf	n.s.	D Heartbeat detection	n.s.	MATLAB, ImageJ
Bourgine *et al.*[18]	M	Confocal microscope	Cellular	3D + t	n.s.	n.s.	D Cell nuclei tracking, cell membrane detection	Development	Own code
Cachat *et al.*[19]	M/A	Camera	Whole organism	3D + t	Adult	625	D Behavior patterns, fish tracking	Toxicology	EthoVision
Cao *et al.*[20]	M	n.s.	Whole organism	2D	30 hpf–5dpf	492	B Phenotype detection	Genetics	Own code
Cario *et al.*[21]	M	Camera	Whole organism	2D + t	Adult	>100	D Behavior patterns, fish tracking	n.s.	MATLAB, LSRTrack
Carvalho *et al.*[22]	A	Confocal microscope, camera	Whole organism	1D	6 dpf	2996	D Phenotype detection, tissue detection, image based pipeline automation	Drug discovery	Copas
Castro *et al.*[23]	M	2PEF Microscope	Cellular/Whole organism	3D	6 hpf	12	D Cell patterns	Development	Own code
Chan *et al.*[24]	M	Microscope	Tissue	2D + t	52 hpf	25	B Heartbeat detection	n.s.	Own code
Chang *et al.*[25]	M	Cameras and confocal microscope	Whole organism	2D	n.s.	100–200	D Image based pipeline automation	n.s.	MATLAB
Chen *et al.*[26]	M	Microscope	Tissue	2D	n.s.	14	D Phenotype detection	Toxicology	MATLAB
Creton[27]	M	Camera	Whole organism	2D + t	6 dpf	>100	D Behavior patterns, fish tracking, social behavior	n.s.	ImageJ, Metamorph
d'Alencon *et al.*[28]	A	Screening WFM	Cellular	2D + t	56 hpf	~5000	B Cell patterns	Inflammation, drug discovery	Labview
Dempsey *et al.*[29]	M	Confocal microscope	Cellular	3D + t	Adult; 8–14hpf	n.s.	D Cell patterns, cell nuclei tracking	Development, regeneration	Imaris, MATLAB
Fangerau *et al.*[30]	M	SPIM	Cellular	3D + t	3–5 hpf, 15–16 hpf	<10	D Behavior patterns, cell patterns, cell nuclei tracking	Development	Scifer
Fink *et al.*[31]	M	Camera	Tissue	2D + t	48–72 hpf	n.s.	D Heartbeat detection, tissue detection	Genetics	MATLAB
Gehrig *et al.*[32]	M/A	Screening WFM	Tissue	2D	32 hpf	17,793	D Phenotype detection, tissue detection	Genetics	MATLAB, Zebrafishminer
Graf *et al.*[33]	M	Camera	Whole organism	2D	4 hpf	650	D Image based pipeline automation	n.s.	Own code
Green *et al.*[34]	A	Camera	Whole organism	2D + t	Adult	171	D Behavior patterns, fish tracking, social behavior	Side effects of drugs	EthoVision

(*continued*)

Table 1. (Continued)

Paper	Application or method-oriented	Imaging modality	Level	Dimension	Fish age and screen size	No. of fish	Algorithms: depth of description (brief or in detail) and outcome	Application	Tools used
Grossman *et al.*[35]	A	Camera	Whole organism	2D+t	Adult	530	D Behavior patterns, fish tracking	Toxicology	EthoVision
Irons *et al.*[36]	A	Camera	Whole organism	2D+t	6 dpf, adult	100…1000	B Behavior patterns, fish tracking	Toxicology	EthoVision
Jeanray *et al.*[37]	M	Microscope	Whole organism	2D	3 dpf	240	B Phenotype detection	Toxicology	CYTOMINE
Kamali *et al.*[38]	M	Confocal microscope	Tissue	3D	7 dpf	11	D Cell patterns	n.s.	Own code, MATLAB(?)
Kanungo *et al.*[39]	A	Microscope	Cellular	2D	28 hpf	240	D Phenotype detection, cell patterns	Toxicology	MetaXpress
Kato *et al.*[40]	M	Camera	Whole organism	2D+t	3 dpf-adult	15…20	D Behavior patterns	n.s.	Own code
Kausler *et al.*[41]	M	SPIM	Cellular	3D+t	2–5 hpf	1	D Cell nuclei tracking	Development	Own code
Keller *et al.*[7]	M/A	SPIM	Cellular	3D+t	1–20 hpf	n.s., < 10(?)	D Cell nuclei tracking	Development	MATLAB
Kokel *et al.*[42]	A	Microscope	Whole organism	2D+t	28 hpf	~ 70,000	D Behavior patterns	Drug discovery	Metamorph
Kokel *et al.*[43]	A	Microscope	Whole organism	2D+t	28 hpf	479	D Behavior patterns	Neuroscience	Metamorph, MATLAB, Gait-CAD
Kriva *et al.*[44]	M	Confocal microscope	Cellular	3D	3–7 hpf	1	D Cell nuclei detection	Development	Own code
Liu *et al.*[45]	M	Microscope	Cellular, Whole organism	2D	24 hpf	n.s., ~100	D Phenotype detection, cell patterns	Genetics	ZFIQ
Liu *et al.*[46]	M	Microscope	Cellular	2D	24 hpf	10…100	D Cell patterns	n.s.	ZFIQ
Liu *et al.*[47]	M	Camera	Whole organism	2D	48 hpf	>1000	D Phenotype detection	Toxicology	Caliph&Emir
Lou *et al.*[48]	M	SPIM	Cellular	3D+t	2–4 hpf	1	D Cell nuclei tracking	Development	DELTR
Lou [49]	M	SPIM	Cellular	3D+t	2–4 hpf	1	D Cell nuclei tracking	Development	DELTR
Lou and Hamprecht[50]	M	SPIM	Cellular	3D+t	2–4 hpf	1	D Cell nuclei tracking	Development	BOT
Luengo-Oroz *et al.*[51]	M	2PEF Microscope, THG, SHG	Cellular	3D+t	0–3 hpf	6	B Cell nuclei tracking, cell membrane detection	Development	MATLAB
Luengo-Oroz *et al.*[52]	M	2PEF Microscope	Cellular	3D+t	n.s.	n.s.	B Denoising, cell nuclei tracking, cell membrane detection	Development	ITK, MATLAB
Mandrell *et al.*[53]	M	Camera	Whole organism	2D	6 hpf	1600	D Image based pipeline automation	n.s.	Own code
Meijer *et al.*[54]	A	Confocal microscope	Cellular	2D+t	16 hpf–3dpf	n.s.	B Cell patterns, cell nuclei tracking	Genetics, inflammation	Huygens, ImageJ, Image Pro Plus
Mikula *et al.*[55]	M	2PEF Microscope	Cellular	3D	9 hpf	1	D Cell patterns, cell membrane detection	Development	Own code
Ocorr *et al.*[56]	M	Camera	Tissue	2D+t	48–72 hpf	n.s.	D Heartbeat detection	Genetics	MATLAB
Ohn and Liebling[57]	M	Microscope	Tissue	2D+t	52 hpf	10	D Heartbeat detection	n.s.	Own code
Olivier *et al.*[58]	M	2PEF Microscope, SHG, THG	Cellular	3D+t	1–4 hpf	6	D Cell nuclei tracking, cell membrane detection	Development	MATLAB

(*continued*)

Table 1. (Continued)

Paper	Application or method-oriented	Imaging modality	Level	Dimension	Fish age and screen size	No. of fish	Algorithms: depth of description (brief or in detail) and outcome	Application	Tools used
Pardo-Martin *et al.*[59]	M	Confocal microscope	Whole organism	2D	24–72 hpf	>500	D Image based pipeline automation	n.s.	MATLAB
Peravali *et al.*[60]	M	Screening WFM	Tissue	3D	48 hpf	~1000	D Cell patterns, tissue detection, image based pipeline automation	n.s.	MATLAB, Fiji
Pfriem *et al.*[61]	M	Camera	Whole organism	2D	24–120 hpf	n.s.	D Image based pipeline automation	n.s.	Labview
Rihel *et al.*[62]	A	Camera	Whole organism	2D+t	4–7 dpf	>60,000	B Behavior patterns, fish tracking	Drug discovery	Videotrack, Zebrabox, MATLAB
Ronneberger *et al.*[63]	M/A	Confocal microscope	Tissue	3D	48–96 hpf	85	D Expression analysis, tissue detection	Neuroscience, development	ImageJ, ViBE-Z
Rubio-Guivernau *et al.*[64]	M	SPIM	Whole	3D	n.s.	n.s.	D Multiview Fusion	n.s	Own code, MATLAB
Saydmohammed *et al.*[65]	A	Microscope	Whole organism	2D	29 hpf	>1000	B Phenotype detection	Drug discovery, toxicology	Definiens, algorithms from Vogt *et al.*[66]
Selderslaghs *et al.*[14]	A	Microscope, camera	Whole organism	2D+t	24–192 hpf	100…1000	D Behavior patterns, fish tracking	Toxicology	EthoVision
Spomer *et al.*[67]	M	Camera	Tissue	2D+t	60–72 hpf	576	D Heartbeat detection, tissue detection, image based pipeline automation	n.s.	Labview
Stern *et al.*[68]	M	Microscope	Whole organism	2D	Adult	~100	D Phenotype detection	Toxicology	CYTOMINE
Tay *et al.*[69]	A	Confocal microscope	Tissue	3D	4 dpf	310	D Phenotype detection	Neuroscience	ImageJ
Temerinac-Ott *et al.*[70]	M	Confocal microscope, SPIM	Whole organism	3D	24 hpf, 72 hpf	5	D Preprocessing	n.s	MATLAB
Vogt *et al.*[66]	M	Microscope	Whole organism	2D	26–48 hpf	>100	D Phenotype detection	Drug discovery	Definiens
Walker *et al.*[71]	A	Plate reader	Whole organism	2D	1–31 dpf	>1000	B Phenotype detection	Genetics	Microplate reader
Weger *et al.*[72]	A	Plate reader	Whole organism	1D+t	1–5 dpf	7296	D Phenotype detection	Drug discovery, toxicology	Microplate reader, R
Wittmann *et al.*[73]	A	Screening WFM	Cellular	2D+t	3 dpf	768	B Phenotype detection, cell patterns	Toxicology, drug discovery, inflammation	Labview, MATLAB, Python
Xu *et al.*[74]	M	Microscope	Whole organism	2D	3–5 dpf	18	D Phenotype detection	Toxicology	Own code (?)
Zanella *et al.*[75]	M	Confocal microscope	Cellular	3D	3–7 hpf	n.s.	D Cell membrane detection	Development	Own code

A, application-oriented paper; M, method-oriented paper; M/A, method-oriented paper with strong application aspects; 2PEF, two-photon-excited fluorescence; SHG, second harmonic generation; THG, third harmonic generation; SPIM, selective plane illumination microscope; 1D+t, one dimensional "image" and time; 2D, two dimensional image; 3D, three dimensional image; hpf, hours postfertilization; dpf, days postfertilization; B, brief; D, in detail; screening WFM, screening widefield microscope; n.s., not specified; ?, not described in detail.

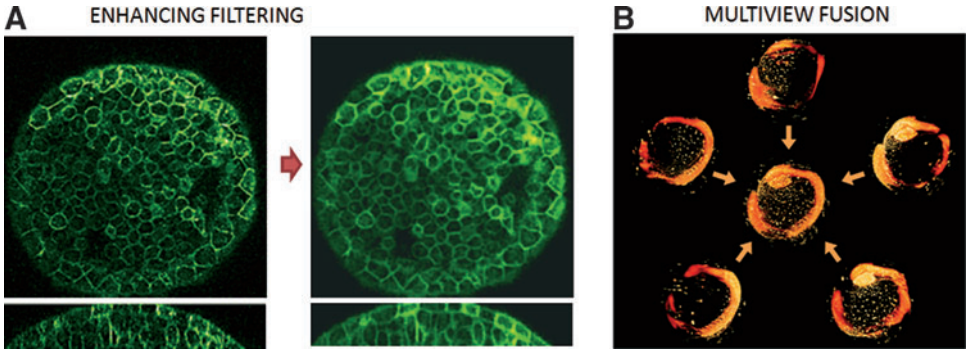

FIG. 3. **(A)** Cross-sections of a zebrafish embryo image filtered following the approach by Luengo-Oroz[52]. *Left*: Cross-sections of the original image. *Right*: Cross-sections of the image filtered by an enhancing (3D+time) multidirectional method. **(B)** Rendering of the fusion of five light sheet fluorescence microscopy views acquired for a zebrafish embryo using the wavelet approach by Rubio-Guivernau *et al.*[64] Renderings of the individual views are shown in a smaller scale around the fusion. Color images available online at www.liebertpub.com/zeb

these analyses are based on commercial instruments whose image analysis algorithms are not published, making them potentially difficult to adapt to specific needs. To circumvent such problems, an open source solution for academic use for the video analysis of zebrafish larvae in multiwell plates was proposed by Cario.[21]

Approaches for adult fish or hatched larvae are not easily transferable to quantify the movements of embryos in their chorion, because movements in the chorion are restricted to rotations. Thus, alternative approaches were proposed that compute gray-value changes between two subsequent video frames in whole wells[14] or lines inside wells.[42]

Heartbeat detection. Image-based heartbeat detection of zebrafish is a key task in the analysis of side-effects of drugs and in the definition of toxicological endpoints. All approaches targeting this parameter defined one or more regions of interest (ROI), for example, heart, atrium, or ventricle. In these ROIs, changes in heart position or blood flow based on 2D+t data are analyzed. Here, different levels of detail were discussed, including a simple yes-no-decision ("beating" or not), a quantification of heartbeat frequency[67] and an additional detection of details such as systolic and diastolic diameters, percent of fractional shortening, or velocity of the contraction wave.[31,56,57] A main problem for live imaging in the zebrafish is the current requirement for immobilization (achieved for example by mechanical constraint in agarose gels[57] and anesthesia[67]). Since immobilization and anesthesia may affect heart rate frequency in these tests, future efforts may be necessary to be able to quantify the zebrafish heartbeat without significant perturbations compared with the normal zebrafish behavior or waking state.

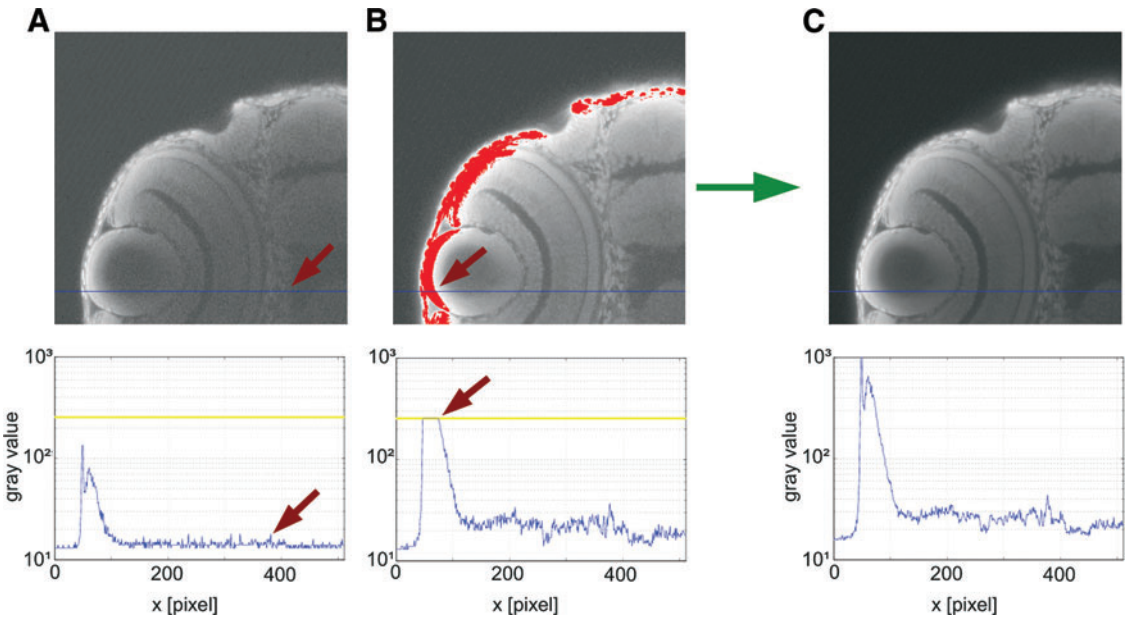

FIG. 4. High dynamic range (HDR) fusion. *Upper panels*: Raw and processed images. *Lower panels*: *Gray value* profile extracted along the *blue line* (corresponding positions marked by *red arrows* in **(A)** and **(B)**). **(A)** Recording with low laser intensity. No over-exposure, but partially bad signal-to-noise-ratio (SNR). **(B)** Recording with high laser intensity. Partial over-exposure (highlighted in *red*), but good SNR. **(C)** After HDR fusion: no over-exposure, good SNR. Color images available online at www.liebertpub.com/zeb

Phenotype assessment by detection of cells and of tissue patterns. Phenotype assessment by robust pattern recognition is crucial to come up with metrics for the characterization of how the zebrafish system responds to perturbations/stimuli at all levels of observation. These patterns may include morphology of the embryo or of certain tissues, as well as changes in reporter gene expression.

Stern *et al.*[68] proposed an approach to automatically detect specific features of interest in 2D microscopy images with the aim of performing automatic morphometric measurements in the context of zebrafish developmental and toxicological studies. A supervised learning approach was followed, in which images were first manually annotated by experts using the CYTOMINE (Cytology and histology image analysis) platform to localize features of interest. The images and their annotations were then automatically exploited for the training of models, which should assess the existence of similar features in new, unseen images. The approach first extracts subwindows (or patches) around points of interest and at other, randomly chosen positions within images. These patches are then described by various visual features. Next, either a classification or a regression model (using extremely randomized decision trees) is built to recognize the points of interest. The model is subsequently applied to new images to predict the localization of similar points and to extract morphometric measurements (i.e., distances and angles). Importantly, the same learning-based approach can be applied to different imaging modalities (e.g., alcian blue staining of cartilage, alizarin red staining of bone).

Applying similar approaches, different algorithms for the automatic classification of embryonic defects in 2D were proposed.[8,15,37,79] All of them rely on supervised machine learning, that is, experts first collected training examples of embryonic defects that are then used to train classification models. Liu used six image descriptors for color and texture combined with a support vector machine to recognize three basic phenotypes (i.e., hatched, unhatched, and dead).[79] Jeanray *et al.*[37] computed pixel-based image descriptors and extremely randomized trees to recognize basic phenotypes and more subtle ones such as pericardial edema and curved tails. Alshut *et al.*[8,15] extracted embryos from images of a microtiter plate well, excluded invalid wells, and detected dead embryos based on the identification of dark regions (Fig. 6).

Automated pattern recognition was also employed for the spatial detection of reporter gene expression driven by various enhancer–promoter combinations in multiple tissues.[32] This was achieved with the development of a high-throughput pipeline for image acquisition and analysis in thousands of fish embryos. The analysis method first automatically detects the embryo outlines, then performs orientation and warping of experimental embryos onto a 2D reference shape, and finally quantifies fluorescence signals (pixel intensities) in presegmented domains of the embryo corresponding to different tissues, such as heart or neural tube.

Phenotype detection can also be performed with regard to the specific pattern of a group of cells. This includes, for example, counting, localization, and pattern detection of a few cells in specific tissues or in specific locations. This is relevant, for example, to analyze patterns of leukocytes in inflammation arrays.[73] In principle, the precise definition of these patterns is application dependent and can be defined in rule-based scripts such as, for example, in MATLAB.

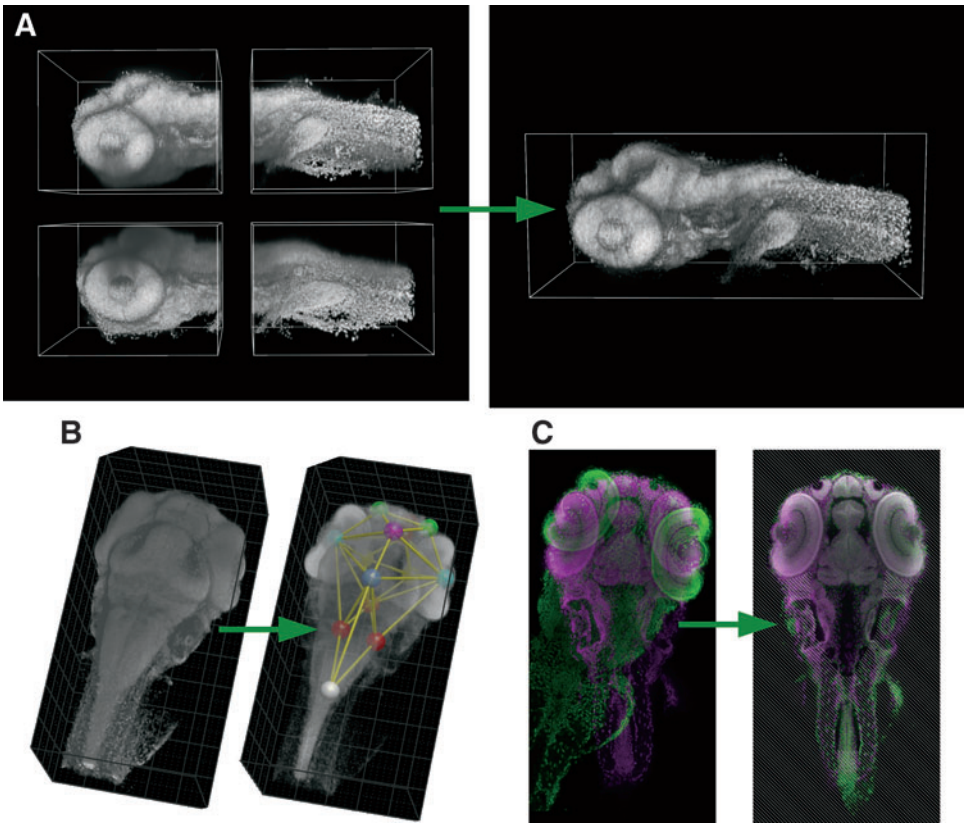

FIG. 5. Image analysis in ViBE-Z: **(A)** Stitching and multi-view fusion with attenuation correction creates a high-quality data set of a 72 hpf zebrafish larva. **(B)** Automated landmark detection. **(C)** Landmark-initialized elastic registration of subject (*green*) to the reference brain (*magenta*). Color images available online at www.liebertpub.com/zeb

In yet another set of applications, landmark recognition, pattern segmentation, and registration (mapping to a reference brain) are combined to build 3D atlases of protein and gene expression.

Ronneberger *et al.*[63] developed the Virtual Brain Explorer for Zebrafish (ViBE-Z) to enable a highly precise mapping of 3D gene expression patterns to stage-specific reference brains, which is currently implemented for 2-, 3-, and 4 day-old larvae. After preprocessing of the image data (Fig. 5A), ViBE-Z uses a trainable landmark detection algorithm on a reference staining to identify certain landmarks (Fig. 5B). For the anatomical reference staining, two types of stains have been used so far: dye labeling of all cell nuclei or anti-acetylated tubulin antibody stain of axon tracts, with each providing sufficient 3D information to accomplish landmark detection. Then, ViBE-Z performs a landmark-initialized elastic registration of the analyzed specimens onto the reference larva (Fig. 5C). The expression pattern that is to be mapped is transformed accordingly and stored in the database. The whole system is available through a web interface, allowing any user to transform newly recorded expression patterns to this reference.

Cell nuclei and membrane detection, cell tracking and lineage reconstruction. Modern 3D microscopy enables the recording of 3D + t images of developing zebrafish embryos at the cellular level, with a temporal resolution that is high enough to track labeled cell nuclei and to identify cell divisions, which allows for the reconstruction of the clonal history of the cells in the embryo. In addition to cell nuclei, cell membranes can be labeled with a second dye and recorded in another channel, which (1) enables the analysis of morphological changes of the cells during development and (2) provides boundary conditions for the labeled nuclei, increasing the fidelity of nuclear segmentation and tracking.

Whole-organism labeling leads to a complex task of image processing that is further complicated by the limitations in SNRs in fluorescence imaging. In addition, temporal resolution in the most commonly used microscopy setups (e.g., confocal or two-photon microscopy) is rather low, making checks for fidelity in segmentation and tracking very demanding. Recent overviews about this special field are given by Hockendorf *et al.*[80] with a focus on zebrafish embryogenesis and by Meijering *et al.*[81] with a focus on generic 2D + t and 3D + t tracking methods and related software tools.

The first step in automated cell tracking is a robust 3D detection of cell nuclei using cell shape information. Furthermore, nuclei detection at the time of mitosis is essential for the identification of cell divisions and subsequent tracking of the daughter cells.[18,49,50,55] Capturing all the features of shape changes of the nucleus at the time of cell division requires a temporal resolution on the scale of minutes, as mitosis proceeds rather rapidly. Reconstructing a full or even partial cell lineage from fluorescently labeled nuclei data is a first important goal.[41,48,58] The deployment of the cell lineage in space and time is the basis for major biological insights, including the identification of the polyclonal origin of organs, the spatial dispersion of cell clones, the rate of proliferation along the lineage, and regularities in the orientation of division planes.[30]

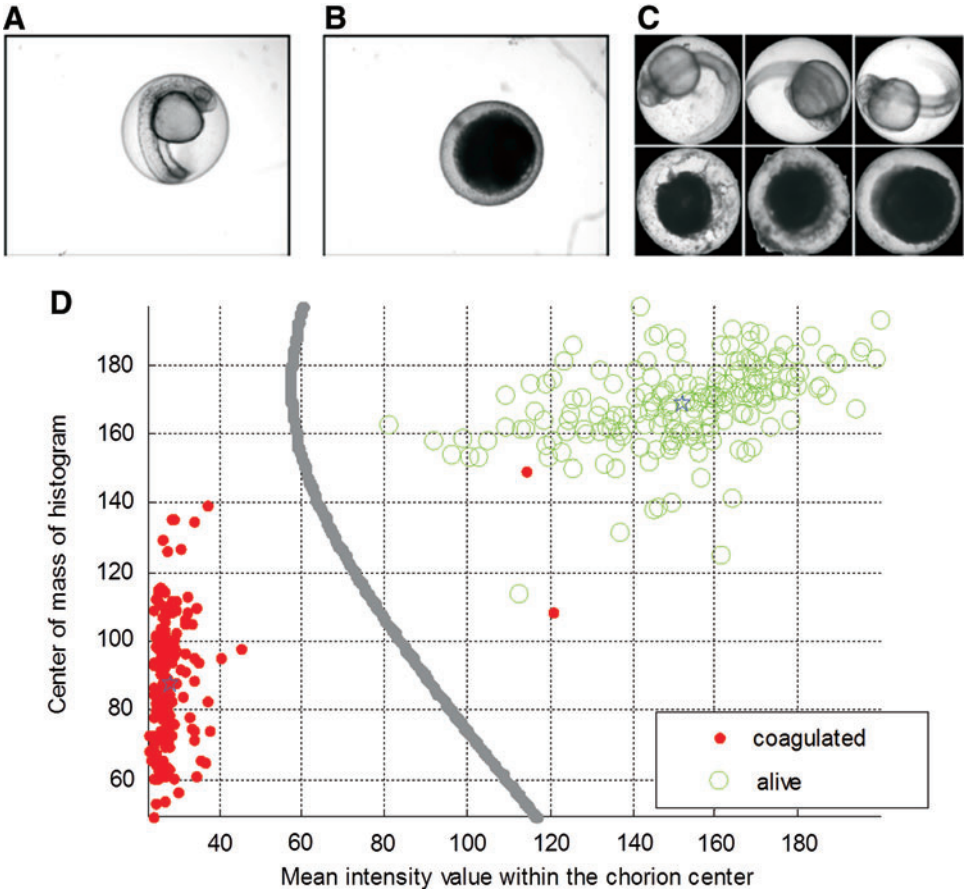

FIG. 6. Example for a phenotype recognition screen including **(A)** a sample image of a living embryo in a microtiter plate well, **(B)** a sample image of a dead embryo, **(C)** three examples for living and dead (necrotic) embryos in different orientations and after extraction from the well image, **(D)** the result of image processing and evaluation based on features automatically extracted from the images. The y-axis gives values for the center of mass of the *gray* value histogram, and the x-axis indicates the mean intensity value in the chorion center. These two values allow an almost error-free classification. Each plotted symbol indicates a necrotic ("coagulated," red points) or living ("alive," *green circles*) embryo as evaluated by manual annotation. The *gray line* gives the cut-off line separating fields for automated classification of necrotic and live embryos.[15] Color images available online at www.liebertpub.com/zeb

An analysis pipeline specifically geared toward the reconstruction of early Zebrafish embryogenesis up to the 1024 cells stage was presented by Luengo-Oroz *et al.*[51] It relies on recordings from multiharmonic microscopy, which allows imaging mitotic spindles (SHG) and cell contours (THG) in an unstained specimen (Fig. 7A). The pipeline produces a full cell lineage after integrating the mitosis detection from the SHG channel with the membrane segmentation from THG channel using a viscous watershed algorithm (Fig. 7B, C).

Another approach for the analysis and reconstruction of zebrafish embryogenesis at later developmental stages (several thousands of cells) was given by Bourgine.[18] The authors developed and tested methods of nonlinear diffusion filtering, nuclei center detection, nuclei and membrane image segmentation, whole embryo segmentation, and mitosis detection using data produced by the BioEmergences platform (http://bioemergences.iscpif.fr), which then served to design the image processing workflow of the platform. In addition to lineage tracing, image processing with the platform revealed cell cycle lengthening during early brain development (Fig. 8).

The development of robust cell tracking algorithms with minimal error rates remains a challenge. An original contribution to the field[82] uses 3D+t image segmentation and extraction of cell trajectories as centerlines of segmented 3D+t tubular structures. The analysis by Luengo-Oroz *et al.*[52] also exploits temporal coherence processing and simultaneously the spatiotemporal dimensions using 3D+t morphological operators. The authors applied this technique to cell tracking, identifying the cell trajectory as a single 4D (3D+t) object, and to the topological description of spatiotemporal gene expression.

Attempts to achieve exhaustive cell tracking from ubiquitously labeled nuclei lead to limited accuracy. It is now clear that, to improve the tracking outcome, image complexity needs to be decreased when cells are densely packed in a tissue. A number of strategies have been described that produce mosaic staining, including transplantation of labeled cells, or random genetic recombination such as in the so-called "brainbow" transgenic lines.[83] Alternatively, Photoconvertible Optical Tracking (PhOTO) transgenic lines enable genetically encoded global and targeted cell labeling for lineage tracing applications throughout the lifetime of the zebrafish. These transgenic lines allow targeted Photoconvertible Optical Tracking Of Nuclei (PhOTO-N) and Membranes (PhOTO-M) with high SNR ratio and are especially well suited for monitoring slow-dividing populations of cells (e.g., stem cells) during regenerative processes.[29]

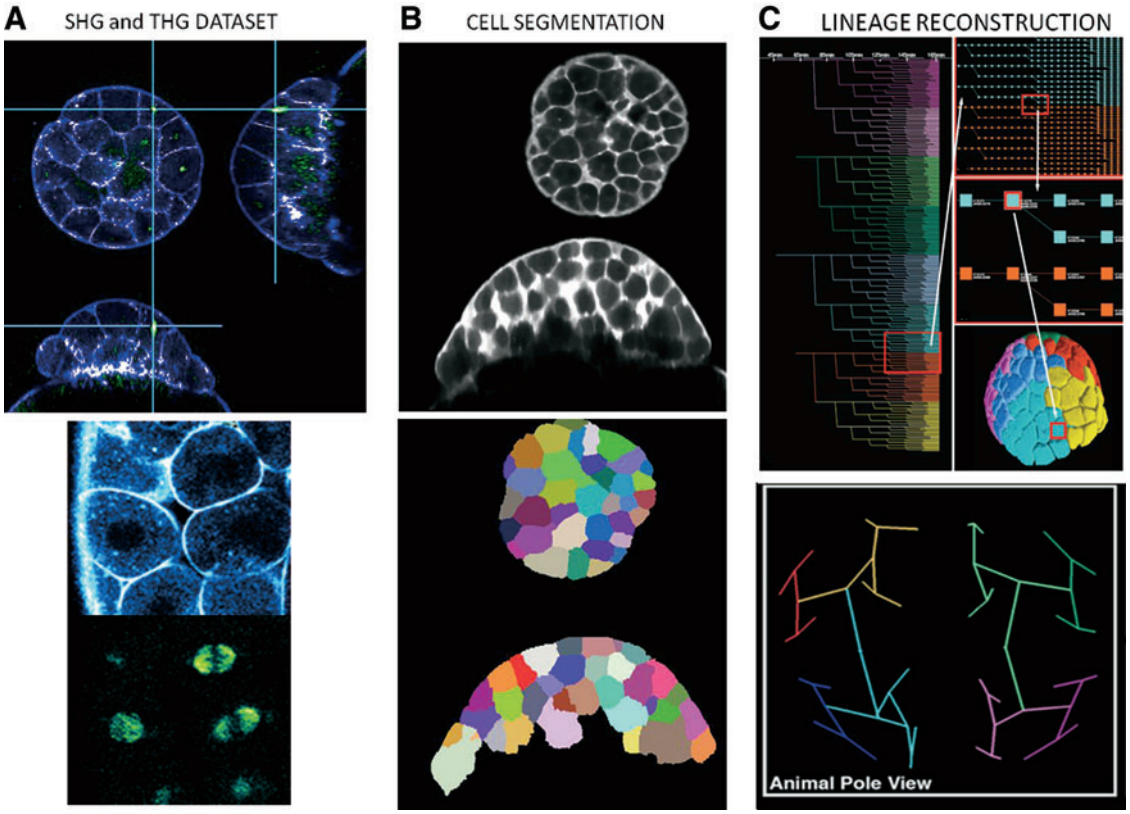

FIG. 7. Reconstruction of zebrafish early embryonic development from multiharmonic imaging data. **(A)** *Top*: three orthogonal views of an original dataset combining third harmonic generation (THG; *blue*) and second harmonic generation (SHG; *green*). *Bottom*: THG and SHG details. **(B)** Cell segmentation after application of a viscous watershed algorithm. *Top*: viscous filtering applied to the original SHG image. *Bottom*: resultant cell segmentation. **(C)** The SHG channel is used to identify cell mitosis, and the integration with the cell segmentation enables cell linage reconstruction (*top*) and the representation of the spatial deployment of the cell lineage tree (*bottom).* **(A, B)** are partly from Luengo-Oroz *et al.*[51] and **(C)** from Olivier *et al.*,[58] reprinted with permission. Color images available online at www.liebertpub.com/zeb

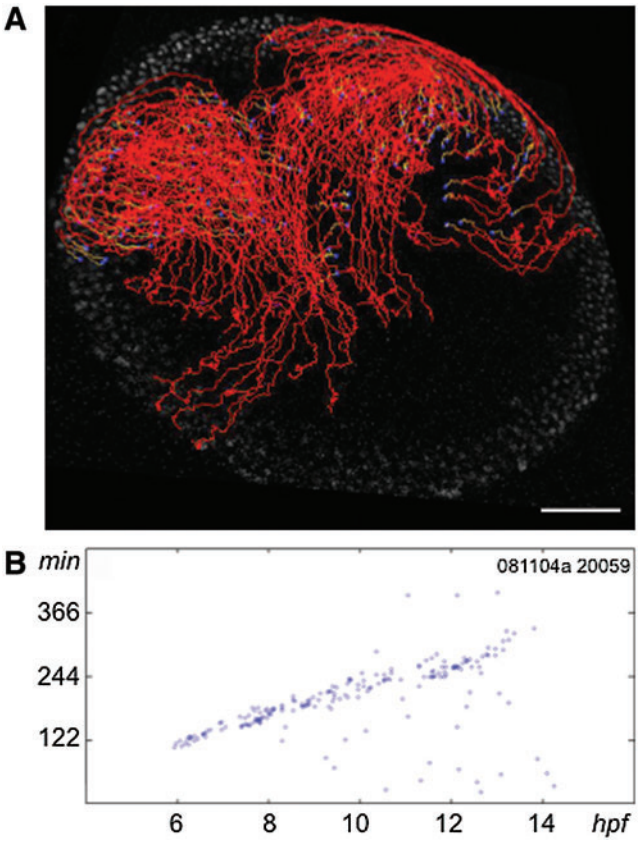

FIG. 8. Automated cell lineage reconstruction and cell cycle length analysis in the presumptive zebrafish brain. **(A)** Animal pole view of the zebrafish brain by early somitogenesis. Anterior is *bottom right* and posterior *top left*. Cell trajectories in the forebrain and midbrain are assessed through the processing of a 3D+time image data set with a mosaic staining obtained through the transplantation of cells with red and green nuclei into a host embryo with green nuclei. The complete cell lineage of the "red nuclei" population has been reconstructed. Reconstructed and raw data (in *white*) are superimposed with the Mov-IT visualization interface. *Red lines* mark cell trajectories throughout the whole spatiotemporal sequence. *Blue cubes* indicate the approximate center of nuclei. Scale bar is 100 μm **(B)**. Similar to data as in **(A)**, each dot indicates a cell division observed at a given developmental time (in hours postfertilization at 26°C), and is plotted as a function of the time elapsed since its birth through its mother's division. The cell cycle lengthens linearly (in min) from 2 to 5 h. The data set and tracking identification numbers are indicated *top right*. This information is a part of the metadata of the experiment (BioEmergences platform, unpublished). Color images available online at www.liebertpub.com/zeb

With the same objective of improving cell and tissue labeling, another line of research explores strategies that are alternative to fluorescent staining. As the zebrafish develops, increased light scattering due to tissue depth, tissue autofluorescence, and light absorption by pigment cells can make fluorescence imaging challenging, especially when performing deep-tissue imaging experiments. SHG nanoprobes may circumvent some of these problems. SHG nanoprobes are nontoxic nanomaterials that take advantage of the nonlinear optical phenomenon of SHG to produce contrast. Recent publications[84,85] have used these nanomaterials for high-contrast labeling *in vivo* during early zebrafish development. Their nonphototoxic and nonbleaching signal makes them prime candidates for long-term cell labeling and visualization in space and time.

Image processing for fully automated zebrafish handling pipelines. Large-scale screening projects require, in addition to imaging data management, the automation of complete screening pipelines, including mechanical handling of fish, substances, and image acquisition. Relatively simple real-time image processing routines are applied to support, for example, automated sorting or selective image acquisition. Examples include the development of flow-based systems for high-throughput imaging and laser microsurgery of larvae "on the fly,"[59] workflows for the automatic injection of bacteria into the yolk,[22] image-based sorting robots,[33,61] a screening platform with automated manipulation of zebrafish including orienting and positioning regions of interest within the microscope's field of view,[25] and automated zooming into ROI for heartbeat detection.[67]

Available Software and Computation Tools

Computer scientists involved in image processing research often rely on scientific computing environments (e.g., MATLAB, SciPy, or OpenCV) for prototyping and evaluation of their algorithms. However, the challenges of bioimage analysis and the need for better software applicability motivated the development of novel tools that could be subdivided into generic tools (not specialized to zebrafish) and specific zebrafish tools.

A complete list of these tools, including WWW links and short descriptions, can be found in the Excel table zip_tools.xls in the additional material. An updated version of this list will be later available at (http://sourceforge.net/projects/zebrafishimage/).

Generic tools

We provide next a nonexhaustive list (in alphabetical order) of generic bioimage analysis tools (for more detailed descriptions, see the survey paper of Eliceiri *et al.*[86]). Although not primarily developed for zebrafish imaging data, these tools might be re-used or tailored to answer specific zebrafish biological questions; see, for example:

- BioImageXD[87]: a 3D visualization and analysis tool using processing pipelines,
- Bisque[88]: a database for the exchange and exploration of biological image data,
- CellProfiler[89]: a software to measure phenotypes in cell and noncell images,
- CYTOMINE[90]: a rich Internet application that allows data visualization remotely. It enables collaborative annotations, data sharing, and analyses of large (>Gigabyte) biomedical images on the web. It is based on generic design principles and machine learning methods that could be used for various tasks (including assessing pathologies *in silico*, and zebrafish developmental and toxicological studies),

- Definiens (Definiens AG): a commercial image processing package for a rule-based design of pipelines,
- EthoVision (Noldus): a commercial package for the tracking of animals with many applications to zebrafish,
- Gait-CAD[91]: a MATLAB-based toolbox for image processing, tracking, and integrated data mining,
- Icy[92]: a platform for image analysis,
- ImageJ: the pioneer in bioimage analysis. Newer variants: ImageJ2/Fiji,[93]
- ITK/VTK: an open-source segmentation and registration toolkit for multidimensional imaging data,
- ilastik[94]: an interactive image classification and segmentation toolkit using machine learning,
- Knime[95]: a workflow system for data analysis and image processing,
- Labview (National Instruments Corp.): a commercial lab automation tool with a focus on real-time processing, including modules for image processing,
- MATLAB (Mathworks Inc.): a commercial mathematical tool with script-based programming and a large variety of additional libraries, including image processing,
- OMERO[96]: a client–server software for visualization, management, and analysis of biological microscope images.

After extraction of features and time series from images and videos, further evaluation and processing steps can be done using statistics tools or data mining tools.[97]

Specific tools for application to the zebrafish model

Another group of tools has been developed specifically for zebrafish image analysis tasks. It includes stand-alone software tools and web services, add-ons for generic tools, and image processing algorithms that are integrated into complete screening systems:

- DanioVision (Noldus, Inc.): an integrated commercial system for the tracking of zebrafish larvae,
- DeltR[48]: an automated pipeline for analysis of time-resolved LSFM images of zebrafish embryogenesis,
- IN Cell Investigator Zebrafish Analysis Plug-In (GE Healthcare Life Sciences): an add-on for the commercial IN Cell system containing preconfigured analysis modules for >50 assays and applications,
- LSRTrack[21]: a MATLAB add-one for tracking of zebrafish larvae (free for academic use),
- ViBE-Z[63]: a free web-based image analysis framework for virtual colocalization studies in larval zebrafish brains in 3D,
- Zebrabox, Zebralab, Zebrafish Cubicle, and Zebratower (Viewpoint): complete commercial systems for zebrafish 2D+t tracking analysis,
- Zebrafish High-Content Screening Automation Leica HCS LSI (Leica): a commercial confocal microscope including integrated image processing software for zebrafish,
- Zebrafishminer: an open source Gait-CAD add-on for tissue detection in zebrafish with algorithms from Gehrig,[32] and
- ZFIQ[98]: a software package that integrates image analysis routines for zebrafish cell detection, zebrafish cell quantification, and zebrafish neuron detection (free for academic use).

A concept of platform services: centralized image processing for distributed image acquisition

The platform BioEmergences (www.bioemergences.eu) proposes the scientific community services for the automated reconstruction of multiscale dynamics in the morphogenesis of model organisms, thereby exploring a "digital campus" concept. It forms a part of the France-BioImaging (http://france-bioimaging.org/) infrastructure, which is the French counterpart of Euro-BioImaging (www.eurobioimaging.eu/). BioEmergences is integrated into the Open Mole architecture (www.openmole.org), original algorithms developed by the platform partners, to produce pipelines that are optimized for specific applications. A first successful application of the BioEmergences Platform was the reconstruction of the zebrafish embryonic cell lineage tree from two-photon laser scanning microscopy,[51,58] discussed in greater detail earlier (Cell nuclei and membrane detection, cell tracking and lineage reconstruction section).

Applications Based on Automated Image Processing

The previous sections show that many methods and tools have been developed for the automated image processing with applications to zebrafish data. However, most of the publications in the field are method-oriented papers that are using zebrafish datasets for proof of concepts to evaluate new methods and to highlight future applications, rather than focusing on new biological insights. Many examples for such proof-of-concept papers are mentioned in the previous section. However, a few method-oriented papers also give some additional biological insights, for example by enabling a sort of "quantitative phenomenology" of complex biological processes such as gastrulation movements[7] or the geometry and timing of the egg cleavages.[58]

Most application papers are based on phenotypic analysis (with a focus on behavioral parameters, discussed in Phenotypic analysis: behavioral parameters section, or on simple morphometric parameters or patterns of reporter gene expression, discussed in Phenotypic analysis: simple morphometric parameters or patterns of reporter gene expression section) or on the construction of 3D atlases of gene and protein expression (Constructing 3D atlases of gene and protein expression section). Integration of anatomical information, gene expression patterns, and functional brain data section discusses the integration of such analyses with anatomical and functional brain data.

Phenotypic analysis: behavioral parameters

Drug screens based on behavioral responses to light stimuli, individual fish movements, or social behaviors have been reported that use automated assessment of relevant parameters.

The effect of 14,000 compounds on the Photo Motor Response of zebrafish embryos has been quantified by 2D+t image analysis of the mean movement in their chorion of 8–10 embryos placed in a single well.[42] This screen identified, for

example, novel acetylcholinesterase (AChE) inhibitors. AChE was suggested as the target of these compounds based on clustering of the behavioral profiles obtained from the image analysis, in which the unknown substances co-clustered with known AChE inhibitors. Using a similar approach, the rest–wake cycle effects over 3 days of 5648 compounds were assessed by quantifying the locomotor activity of zebrafish larvae.[62] This screen implicated ether-a-go-go-related gene (ERG) potassium channels and immunomodulators in the regulation of rest and locomotor activity.

In other studies, movement patterns of a single zebrafish larva (a few days old) placed in a well of a 96-well plate were analyzed with commercial software. This enabled the assessment of effects of 14 anti-epileptic drugs,[16] of the concentration-dependent neuroactive effects of ethanol, d-amphetamine, and cocaine,[36] and of the age- and concentration-dependent effects of chlorpyrifos.[14]

The analysis of behavioral patterns becomes more complicated with the simultaneous observation of adult fish in tanks. Classical studies of the social behavior of fish use one camera in 2D+t approaches, with which side effects of different concentrations of the hallucinogenic agents mescaline and phencyclidine[34] or the effects of lysergic acid diethylamide (LSD)[35] were observed. The reconstruction of 3D+t movements can be achieved by simultaneously operating two cameras. This was done to analyze the complete swim path of fish in different genetic (*leopard* strain) or environmental conditions (test of seven different substances).[19] The fish movements were measured and analyzed by hierarchical clustering, revealing two clusters of substances with anxiogenic and anxiolytic effects, respectively.

An overview about further activities in neurobiological image processing in zebrafish with a focus on behavioral parameters can be found in the survey paper of Xia.[99]

Phenotypic analysis: simple morphometric parameters or patterns of reporter gene expression

In addition to the image processing approaches that were customized for the zebrafish data previously discussed, many standard approaches using commercial devices such as plate readers or cell counters have been proposed. Here, simple features are detected (e.g., mean signal intensity in the image, number of fluorescent cells in a region), followed by further analysis and determination of statistically relevant differences. The main advantage of these approaches is their availability in the market, but the level of complexity that can be investigated is usually limited.

Large-scale phenotypic screens can use simple readouts. Automated Reporter Quantification *in vivo* (ARQiv) provides a quantitative data acquisition approach that is amenable to reporter-based assays in zebrafish embryos, larvae, or juveniles. The method has been validated to quantify loss and regeneration of fluorescently tagged cell types, relative activity of a transgenic Notch-signaling reporter, and metabolism of the animals in the well by measuring the accumulation of reactive oxygen species.[71] In addition to the typically used zebrafish transgenic for fluorescent reporter genes, animals carrying bioluminescent reporter constructs can provide a sensitive readout alternative for large-scale compound screens, as recently shown for a reporter for glucocorticoid signaling.[72] Kanungo *et al*.[39,100] quantified the toxic effects of different concentrations of ethanol on axon length. A drug screen with the zebrafish tuberculosis model of *Mycobacterium marinum* infection was performed with a fully automated pipeline, including automated yolk injection and a quantification based on a COPAS flow cytometry system to determine the total level of red fluorescence, representative of bacterial load.[22] Another approach based on target gene expression patterns[65] analyzed a library of 1040 annotated bioactive agents on Fibroblast Growth Factor signaling *in vivo*. Here, a 2D pattern detection algorithm was applied to assess brightness and extent of reporter gene expression in specific regions.[66] On a similar theme, a high-throughput pipeline was developed to record embryo domain-specific reporter expression to map the interactions between cis-regulatory modules and core promoters.[32] Automated microscopy coupled with custom-built embryo detection and segmentation software allowed the spatial description of reporter activity for 202 enhancer-promoter combinations.

An indepth understanding of the processes underlying inflammation requires the *in vivo* observation, visualization, and analysis of cell behaviors. A few approaches analyze such processes by automated image processing and tracking of leukocytes[28,54,73] and neuromast positions[28,73] with 2D+t approaches, for example, to analyze pro-inflammatory or anti-inflammatory effects of 320 drugs.[73]

Constructing 3D atlases of gene and protein expression

Gene and protein expression atlases are valuable resources that are used to quantify and understand multiscale processes of morphogenesis in time and space. The automated reconstruction of prototypic 3D atlases for zebrafish development, either at the scale of the whole embryo or for specific organs, requires the development of dedicated methods. This type of work has been pioneered by *Drosophila*.[101,102] Comparable strategies have been recently developed for the zebrafish model, for example, for the whole blastula and early gastrula[23] or for the brain at early larval stages.[63] Mapping expression data of analyzed specimen into an anatomical reference model enables the establishment of 3D atlas databases with a unified spatial representation for many different expression patterns. To reach this goal, several preconditions and algorithmic steps are required: (1) Imaging: 3D datasets need to be of standardized high quality and resolution (see Preprocessing methods section above). (2) Registration: A common anatomical reference staining needs to be recorded for each specimen to enable registration to the anatomical reference model; landmarks have to be recognized in an automated fashion; and registration strategies need to be developed, which usually start with a coarse landmark-based registration and are subsequently optimized by fine elastic registration that relies on local information provided by the anatomical reference stain (see description of the ViBE-Z system above, Phenotype assessment by detection of cells and of tissue patterns section and Fig. 5). (3) Last but not least, algorithms for threshold determination of signal versus background need to be determined. Except for the very early stereotypic cell cleavages of the early embryo, identification of individual positive cells will not be possible due to stochastic aspects of cell positioning within anatomical elements of zebrafish embryos. However, using high-quality data and precise registration, ViBE-Z has achieved cellular resolution with regard to

mapping individual nuclei to even small anatomical brain domains at one to five micrometer resolution.[63] An initial database (atlas) with 15 aligned gene expression patterns (each averaged over approximately three individuals) along with a manual segmentation of morphological structures is provided for virtual colocalization studies (Fig. 9).

In addition, the development of dedicated interactive visualization tools is an issue to fully exploit the atlas resources. The future of these atlases will certainly depend on the effort that is made to help them grow through the addition of new data by the community. The design of the ViBE-Z system[63] is geared into this direction: ViBE-Z accepts 3D confocal stacks generated by the community, and the web interface performs landmark detection, fine elastic registration, and provides the 3D registration to the anatomical reference for download in a format that can be further processed by investigators with ImageJ or other programs. It will be interesting to see whether this or similar systems will be accepted by the community as data standards and for upload into databases. It is necessary that 3D expression atlases will be tightly integrated into the extensive and standardized annotation system provided in ZFIN. The registration into anatomical standard spaces, however, also makes annotation-free searches of large datasets possible by using defined volume elements as queries.[63] Tools for this and other types of searches need to be developed and integrated into interactive search engines.

Integration of anatomical information, gene expression patterns, and functional brain data

The recent development of optogenetic tools has enabled the functional interrogation of brain function in zebrafish larvae. Defined neuronal groups can be activated by depolarization or inhibited by hyperpolarization using light-switchable ion channels and ion pumps in whole zebrafish larvae. This can be combined with genetically encoded fluorescent calcium sensors to visualize circuit activity in normal behaving or optogenetically manipulated larvae. The transparent nature of the larval brain enables "holistic" imaging of calcium activity throughout the whole brain, opening unique opportunities for new insights into circuit function.[103] New challenges for image processing emerge: Circuit activity data should be merged with anatomical information to provide information on the neurons that show activity. Furthermore, data from multiple larval brains should be integrated to identify contributions to defined circuits. Ahrens *et al.*[103] have provided algorithms that enable this type of analysis. Ideally, the spatial coordinates of each individual experimental brain could be recorded before calcium imaging, and registration to anatomical reference models could be used to predict the location of regions that may be subjected to optogenetic manipulation. Such advances in optical imaging and automated image processing will significantly enhance our understanding of neural circuits.

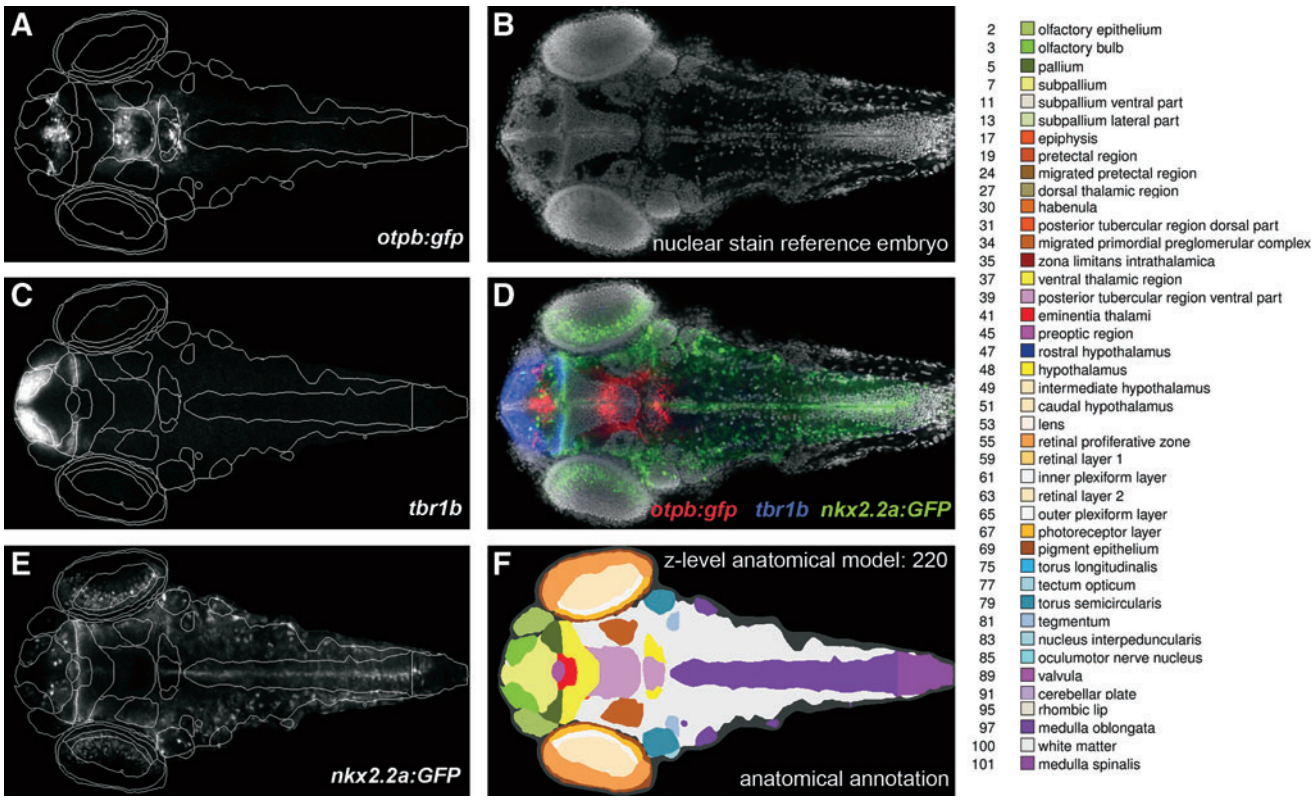

FIG. 9. The Virtual Brain Explorer for Zebrafish (ViBE-Z)[63] provides a web-based interface to register 3D datasets recorded by a standardized procedure to an anatomical reference model. **(A, C, E)** shows expression data channels of a single longitudinal dorsoventral focal plane from three different 3 day old larval brains (anterior to the *left*; *otpb:gfp* and *nkx2.2a*;GFP immunofluorescence and *tbr1b* whole mount *in situ* hybridization) along with lines delimiting anatomical brain regions. **(B)** shows the fluorescent nuclear staining pattern used for registration to the reference larvae. **(D)** shows superimposition of the three expression data channels and the nuclear stain in the same dorsoventral plane. **(F)** shows color coded anatomical domains with color code given on the right. Color images available online at www.liebertpub.com/zeb

Discussion: Open Questions and Future Trends

Automated processing is changing the standards of quantitative analysis of imaging data in the zebrafish model. Automated image analysis is of particular interest for high-content screening. However, method and software integration into standardized, modular, and open processing pipelines, including data storage, image processing, data analysis, visualization, and web-based publication of results, still needs to be developed.

We identify a number of open questions for a newly born interdisciplinary community:

How Can Existing Zebrafish-Specific Image Processing Algorithms, Tools, and Data Be Made Available for the Community?

As a first step, the publication of existing zebrafish-specific modules and a clear description of necessary interfaces for the corresponding input and output data would help the community get a better overview of existing methods and tools. In addition, this list should be enriched with links to papers describing successful applications, including user rating and comments. This paper is a first step in this direction, and it provides an Excel table hosted at Sourceforge (http://sourceforge.net/projects/zebrafishimage/) that will regularly be updated with a short description of open source and commercial tools. A second step is to host open source components in online code repositories (e.g., Sourceforge, GitHub) to avoid the loss of algorithms when developers leave and group websites get outdated.

How Can Algorithms and Tools be Modularized, Generalized, and Bundled into Compatible Pipelines?

It is unlikely that the variety of different standards, image repositories, algorithms, and tools will vanish in the near future. Strong communities exist for many of these tools (such as ITK, ImageJ/Fiji, and MATLAB), and at the moment, none of them have a clear general advantage or disadvantage. In addition, software development requires specific knowledge and training efforts, making developers usually hesitate to change their toolbox preferences. This means that heterogeneous tool environments will have to also be handled in the future. Similar effects can be expected for standards and image repositories.

Initiatives such as the Open Bio Imaging Alliance (www.openbioimage.org/) try to influence this process and provide an overview of existing software resources with the corresponding knowledge. The aim of this initiative is to "federate the harmonious community-based development of interoperable software and promote good practices, including the careful validation of tools." Therefore, the development of algorithms and tools for image processing in the zebrafish should take the following into consideration:

(1) The use of routines would be greatly facilitated if implemented as plugins for larger generic tools. Previously cited generic tools such as CellProfiler, CYTOMINE, Icy, and ImageJ provide API or plug-in functionalities to extend their capabilities.
(2) A condition for interoperability is a clear definition and description of standards and compatibilities between different algorithms and tools. With this effort, output images of one tool can be imported into other tools. This requires import and export routines to translate between different formats. One example is the import routines for more than one hundred bioimage formats in OMERO,[96] but robust bidirectional tools are missing.
(3) To avoid manual work, it will also be necessary to establish zebrafish-specific guidelines for the specification of metadata corresponding to all types of assays, imaging setups, image types, algorithms, and tools. In addition, all relevant metadata should be saved in an open and machine-readable format. These guidelines should be implemented as extensions to data formats of other more general initiatives and integrated into existing zebrafish databases, mainly into the Zebrafish Model Organism Database of the Zebrafish Information Network ZFIN (http://zfin.org).

How Should Screens Be Designed to Make Image Processing Easier?

Another important question is the planning of screens that can be easily processed. Here, a first requirement is the standardization of all the parameters in the assay (fish lines, age of the specimen, etc.) and of the image acquisition parameters (e.g., microscope parameters such as resolution in space and time, number of voxels, etc.). Image processing experts should be consulted in an early phase of the planning of the experiments to choose all these parameters and assay conditions according to the requirements of image processing routines. Otherwise, normalization and artifact corrections will probably be necessary that are very costly in time, and some artifacts might even compromise the image analysis results.[11]

A more comprehensive list of such recommendations can be found in.[104] As discussed earlier, the extension of these more general standards to the needs of the zebrafish community would be necessary.

How Can Datasets with Terabytes of Data Be Handled?

It is a known and successful paradigm in the analysis of large datasets to "move the knowledge extraction algorithms to the data rather than the data to the algorithms."[105] It means that data should be transferred to repositories, and these repositories should be able to process the data using distributed file systems and computer clusters such as, for example, Hadoop.[106] One option is the "going public" of existing large screening centers for customized screens of external partners and the integration of next-generation data repositories into these centers. Steps in this direction can be found, for example, in the connection between the European Zebrafish Resource Center and the Large-Scale Data Facility of the Karlsruhe Institute of Technology.[13] The program "Large-Scale Data management and Analysis" of the German Helmholtz Association (www.helmholtz-lsdma.de/) established a Data Life Cycle Lab "Key Technologies," which specifically focuses on data management of experiments requiring imaging with high data rates.

The Research Data Alliance (http://rd-alliance.org/), currently in the founding phase, is a world-wide initiative that will provide policies as well as new technologies for federated data access and sharing. Other initiatives such as the founding

of an UNITWIN UNESCO (http://unitwin-cs.org/) chair for gathering universities and academic research institution around the world within a Digital Campus for Complex Systems have similar goals. In addition, unsolved zebrafish image processing problems (e.g., robust tissue detection in zebrafish phenotypes with severe malformations, fusion of multiple fluorescence channels with delay times between channels in cell tracking screens,…) should be identified and published as challenges for international image processing conferences; see, for example, the Particle Tracking Challenge at the IEEE International Symposium on Biomedical Imaging (ISBI).[107] Organizing such challenges require a clear definition of the goals of an image processing task, of the metrics employed, as well as of ground-truth information (e.g, segmentation results) and metadata in standard formats.

Acknowledgments

The authors acknowledge all participants of the Satellite Workshop "Future image processing platforms" of the 2nd European Zebrafish Principal Investigator Meeting in Karlsruhe,[108] Germany, March 21st–24th, 2012, and the members of the Zebrafish Image Processing (ZIP) Group for many fruitful discussions. This work was supported by the European Union (EU COST Action BM0804, EU IP ZF-Health for WD, NP and US), the grant APVV-0184-10 (KM), the CYTOMINE grant (DGO6 WIST3 1017072, RMa), the grant ANR-10-INBS-04-05 (NP), the GIGA (with financial support from the Wallonia and the European Regional Development fund, RMa), and the Helmholtz Association (RMi, TD, RS, and US). PG is a research associate of the Belgian Science Research Foundation (FRS-FNRS). The work of WD and OR was funded by the Excellence Initiative of the German Federal and State Governments (Centre for Biological Signalling Studies EXC 294). MJL-C and AS acknowledge partial funding from Spain's Ministry of Science and Innovation (TEC2010-21619-C04-03 & TEC2011-28972-C02-02) and European Regional Development Funds. The work of PP was funded by the Swiss National Center of Competence in Research (NCCR) "Nanoscale Science". RS was supported by Helmholtz Portfolio Extension "Large Scale Data Management and Analysis" with the contribution of the Data Life Cycle Lab "Key Technologies" and the "Data Services Integration Team".

Disclosure Statement

No competing financial interests exist.

References

1. Detrich III HW, Westerfield M, Zon LI. The Zebrafish: Disease Models and Chemical Screens, third edition. Academic Press, Waltham, MA, 2011.
2. Yang LX, Ho NY, Alshut R, Legradi J, Weiss C, Reischl M, *et al.* Zebrafish embryos as models for embryotoxic and teratological effects of chemicals. Reprod Toxicol 2009;28:245–253.
3. Bakkers J. Zebrafish as a model to study cardiac development and human cardiac disease. Cardiovasc Res 2011;91:279–288.
4. Mione MC, Trede NS. The zebrafish as a model for cancer. Dis Model Mech 2010;3:517–23.
5. Flinn L, Bretaud S, Lo C, Ingham PW, Bandmann O. Zebrafish as a new animal model for movement disorders. J Neurochem 2008;106:1991–1997.
6. Luengo-Oroz MA, Ledesma-Carbayo MJ, Peyrieras N, Santos A. Image analysis for understanding embryo development: a bridge from microscopy to biological insights. Curr Opin Genet Dev 2011;21:630–637.
7. Keller PJ, Schmidt AD, Wittbrodt J, Stelzer EH. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science 2008;322:1065–1069.
8. Alshut R, Legradi J, Liebel U, Yang L, van Wezel J, Strähle U, *et al.* Methods for automated high-throughput toxicity testing using zebrafish embryos. Lect Notes Artif Intell 2010;6359:219–226.
9. Devarakonda R, Palanisamy G, Green JM, Wilson BE. Data sharing and retrieval using OAI-PMH. Earth Sci Inform 2011;4:1–5.
10. Taylor CF, Field D, Sansone SA, Aerts J, Apweiler R, Ashburner M, *et al.* Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nat Biotechnol 2008;26:889–896.
11. Shamir L. Assessing the efficacy of low-level image content descriptors for computer-based fluorescence microscopy image analysis. J Microsc 2011;243:284–292.
12. Layton JB. Intro to Nested-RAID: RAID-01 and RAID-10. Linux Magazine, 2011, January 6.
13. Jejkal T, Hartmann V, Stotzka R, Otte J, Garcia A, van Wezel J, *et al.* LAMBDA—The LSDF Execution Framework for Data Intensive Applications. 20th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), 2012. Available at: http://dx.doi.org/10.1109/PDP.2012.69
14. Selderslaghs IW, Hooyberghs J, De Coen W, Witters HE. Locomotor activity in zebrafish embryos: a new method to assess developmental neurotoxicity. Neurotoxicol Teratol 2010;32:460–471.
15. Alshut R, Legradi J, Mikut R, Strähle U, Reischl M. Robust Identification of Coagulated Zebrafish Eggs using Image Processing and Classification Techniques. 19 Workshop Computational Intelligence, pp. 9–21, 2009. Available at: http://digbib.ubka.uni-karlsruhe.de/volltexte/1000013551
16. Berghmans S, Hunt J, Roach A, Goldsmith P. Zebrafish offer the potential for a primary screen to identify a wide variety of potential anticonvulsants. Epilepsy Res 2007;75:18–28.
17. Bhat S, Liebling M. Cardiac Tissue and Erythrocyte Separation in Bright-field Microscopy Images of the Embryonic Zebrafish Heart for Motion Estimation. IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), 2009. Available at: http://dx.doi.org/10.1109/ISBI.2009.5193155
18. Bourgine P, Cunderlik R, Drblikova-Stasova O, Mikula K, Peyrieras N, Remesikova M, *et al.* 4d embryogenesis image analysis using Pde methods of image processing. Kybernetika 2010;46:226–259.
19. Cachat J, Stewart A, Utterback E, Hart P, Gaikwad S, Wong K, *et al.* Three-dimensional neurophenotyping of adult zebrafish behavior. PLoS One 2011;6:e17597.
20. Cao Y, Semanchik N, Lee SH, Somlo S, Barbano PE, Coifman R, *et al.* Chemical modifier screen identifies HDAC inhibitors as suppressors of PKD models. Proc Natl Acad Sci U S A 2009;106:21819–21824.

21. Cario CL, Farrell TC, Milanese C, Burton EA. Automated measurement of zebrafish larval movement. J Physiol 2011;589:3703–3708.
22. Carvalho R, de Sonneville J, Stockhammer OW, Savage ND, Veneman WJ, Ottenhoff TH, *et al.* A high-throughput screen for tuberculosis progression. PLoS One 2011;6: e16779.
23. Castro C, Luengo-Oroz M, Desnoulez S, Duloquin L, Fernandez-de-Manuel L, Montagna S, *et al.* An Automatic Quantification and Registration Strategy to Create a Gene Expression Atlas of Zebrafish Embryogenesis. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2009. Available at: http://dx.doi.org/10.1109/IEMBS.2009.5332436
24. Chan PK, Lin CC, Cheng SH. Noninvasive technique for measurement of heartbeat regularity in zebrafish (*Danio rerio*) embryos. BMC Biotechnol 2009;9:11.
25. Chang TY, Pardo-Martin C, Allalou A, Wahlby C, Yanik MF. Fully automated cellular-resolution vertebrate screening platform with parallel animal processing. Lab Chip 2012;12:711–716.
26. Chen S, Zhu Y, Xia W, Xia S, Xu X. Automated analysis of zebrafish images for phenotypic changes in drug discovery. J Neurosci Methods 2011;200:229–236.
27. Creton R. Automated analysis of behavior in zebrafish larvae. Behav Brain Res 2009;203:127–136.
28. d'Alencon CA, Pena OA, Wittmann C, Gallardo VE, Jones RA, Loosli F, *et al.* A high-throughput chemically induced inflammation assay in zebrafish. BMC Biol 2010;8:151.
29. Dempsey WP, Fraser SE, Pantazis P. PhOTO zebrafish: a transgenic resource for *in vivo* lineage tracing during development and regeneration. PLoS One 2012;7: e32888.
30. Fangerau J, Hockendorf B, Wittbrodt J, Leitte H. Similarity Analysis of Cell Movements in Video Microscopy. IEEE Symposium on Biological Data Visualization (BioVis), 2012. Available at: http://dx.doi.org/10.1109/BioVis.2012.6378595
31. Fink M, Callol-Massot C, Chu A, Ruiz-Lozano P, Izpisua Belmonte JC, Giles W, *et al.* A new method for detection and quantification of heartbeat parameters in *Drosophila*, zebrafish, and embryonic mouse hearts. Biotechniques 2009;46:101–113.
32. Gehrig J, Reischl M, Kalmar E, Ferg M, Hadzhiev Y, Zaucker A, *et al.* Automated high-throughput mapping of promoter-enhancer interactions in zebrafish embryos. Nat Methods 2009;6:911–916.
33. Graf SF, Hotzel S, Liebel U, Stemmer A, Knapp HF. Image-based fluidic sorting system for automated zebrafish egg sorting into multiwell plates. J Lab Autom 2011;16:105–111.
34. Green J, Collins C, Kyzar EJ, Pham M, Roth A, Gaikwad S, *et al.* Automated high-throughput neurophenotyping of zebrafish social behavior. J Neurosci Methods 2012;210: 266–271.
35. Grossman L, Utterback E, Stewart A, Gaikwad S, Chung KM, Suciu C, *et al.* Characterization of behavioral and endocrine effects of LSD on zebrafish. Behav Brain Res 2010;214:277–284.
36. Irons TD, MacPhail RC, Hunter DL, Padilla S. Acute neuroactive drug exposures alter locomotor activity in larval zebrafish. Neurotoxicol Teratol 2010;32:84–90.
37. Jeanray N, Maree R, Pruvot B, Stern O, Geurts P, Wehenkel L, *et al.* Phenotype classification of zebrafish embryos by supervised learning. Toxicol Lett 2012;211:S152.
38. Kamali M, Day LJ, Brooks DH, Zhou X, O'Malley DM. Automated identification of neurons in 3D confocal datasets from zebrafish brainstem. J Microsc 2009;233:114–131.
39. Kanungo J, Lantz S, Paule MG. *In vivo* imaging and quantitative analysis of changes in axon length using transgenic zebrafish embryos. Neurotoxicol Teratol 2011;33:618–623.
40. Kato S, Nakagawa T, Ohkawa M, Muramoto K, Oyama O, Watanabe A, *et al.* A computer image processing system for quantification of zebrafish behavior. J Neurosci Methods 2004;134:1–7.
41. Kausler BX, Schiegg M, Andres B, Lindner M, Leitte H, Hufnagel L, *et al.* A Discrete Chain Graph Model for 3D+t Cell Tracking with High Misdetection Robustness, 2012. Available at: http://dx.doi.org/10.1007/978-3-642-33712-3_11
42. Kokel D, Bryan J, Laggner C, White R, Cheung CY, Mateus R, *et al.* Rapid behavior-based identification of neuroactive small molecules in the zebrafish. Nat Chem Biol 2010;6:231–237.
43. Kokel D, Dunn T, Ahrens M, Alshut R, Cheung CY, Saint-Amant L, *et al.* Identification of non-visual photomotor response cells in the vertebrate hindbrain. J Neurosci 2013;33:3834–3841.
44. Kriva Z, Mikula K, Peyrieras N, Rizzi B, Sarti A, Stasova O. 3D early embryogenesis image filtering by nonlinear partial differential equations. Med Image Anal 2010;14:510–526.
45. Liu T, Lu J, Wang Y, Campbell WA, Huang L, Zhu J, *et al.* Computerized image analysis for quantitative neuronal phenotyping in zebrafish. J Neurosci Methods 2006;153: 190–202.
46. Liu T, Li G, Nie J, Tarokh A, Zhou X, Guo L, *et al.* An automated method for cell detection in zebrafish. Neuroinformatics 2008;6:5–21.
47. Liu R, Lin S, Rallo R, Zhao Y, Damoiseaux R, Xia T, *et al.* Automated phenotype recognition for zebrafish embryo based *in vivo* high throughput toxicity screening of engineered nano-materials. PLoS One 2012;7:e35014.
48. Lou X, Kaster F, Lindner M, Kausler B, Kothe U, Hockendorf B, *et al.* Deltr: digital embryo lineage tree reconstructor. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2011. Available at: http://dx.doi.org/10.1109/ISBI.2011.5872698
49. Lou X. Biomedical Data Analysis with Prior Knowledge: Modeling and Learning: University of Heidelberg, 2011. Available at: http://nbn-resolving.de/urn:nbn:de:bsz:16-opus-122945
50. Lou X, Hamprecht FA. Structured Learning for Cell Tracking. Twenty-Fifth Annual Conference on Neural Information Processing Systems (NIPS 2011), 2011. Available at: http://books.nips.cc/papers/files/nips24/NIPS2011_0766.pdf
51. Luengo-Oroz MA, Rubio-Guivernau JL, Faure E, Savy T, Duloquin L, Olivier N, *et al.* Methodology for reconstructing early zebrafish development from *in vivo* multiphoton microscopy. IEEE Transact Image Proc 2012;21:2335–2340.
52. Luengo-Oroz MA, Pastor D, Castro C, Faure E, Savy T, Lombardot B, *et al.* 3D+t Morphological processing: applications to embryogenesis image analysis. IEEE Transact Image Proc 2012;21:3518–3530.
53. Mandrell D, Truong L, Jephson C, Sarker MR, Moore A, Lang C, *et al.* Automated zebrafish chorion removal and single embryo placement: optimizing throughput of zebrafish developmental toxicity screens. J Lab Autom 2012;17:66–74.

54. Meijer AH, van der Sar AM, Cunha C, Lamers GE, Laplante MA, Kikuta H, *et al.* Identification and real-time imaging of a myc-expressing neutrophil population involved in inflammation and mycobacterial granuloma formation in zebrafish. Dev Comp Immunol 2008;32:36–49.
55. Mikula K, Peyrieras N, Remesikova M, Stasova O. Segmentation of 3D cell membrane images by PDE methods and its applications. Comput Biol Med 2011;41:326–339.
56. Ocorr K, Fink M, Cammarato A, Bernstein S, Bodmer R. Semi-automated optical heartbeat analysis of small hearts. Journal of visualized experiments: JoVE 2009;31.
57. Ohn J, Liebling M. *In vivo*, High-throughput Imaging for Functional Characterization of the Embryonic Zebrafish Heart. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2011. Available at: http://dx.doi.org/10.1109/ISBI.2011.5872696
58. Olivier N, Luengo-Oroz MA, Duloquin L, Faure E, Savy T, Veilleux I, *et al.* Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. Science 2010;329:967–971.
59. Pardo-Martin C, Chang TY, Koo BK, Gilleland CL, Wasserman SC, Yanik MF. High-throughput *in vivo* vertebrate screening. Nat Methods 2010;7:634–636.
60. Peravali R, Gehrig J, Giselbrecht S, Lutjohann DS, Hadzhiev Y, Muller F, *et al.* Automated feature detection and imaging for high-resolution screening of zebrafish embryos. Biotechniques 2011;50:319–324.
61. Pfriem A, Pylatiuk C, Alshut R, Ziegener B, Schulz S, Bretthauer G. A modular, low-cost robot for zebrafish handling. Engineering in Medicine and Biology Society (EMBC), 2012. Available at: http://dx.doi.org/10.1109/EMBC.2012.6346097
62. Rihel J, Prober DA, Arvanites A, Lam K, Zimmerman S, Jang S, *et al.* Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. Science 2010;327:348–351.
63. Ronneberger O, Liu K, Rath M, Ruebeta D, Mueller T, Skibbe H, *et al.* ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains. Nat Methods 2012;9:735–742.
64. Rubio-Guivernau JL, Gurchenkov V, Luengo-Oroz MA, Duloquin L, Bourgine P, Santos A, *et al.* Wavelet-based image fusion in multi-view three-dimensional microscopy. Bioinformatics 2012;28:238–245.
65. Saydmohammed M, Vollmer LL, Onuoha EO, Vogt A, Tsang M. A high-content screening assay in transgenic zebrafish identifies two novel activators of fgf signaling. Birth Defects Res Part C 2011;93:281–287.
66. Vogt A, Cholewinski A, Shen X, Nelson SG, Lazo JS, Tsang M, *et al.* Automated image-based phenotypic analysis in zebrafish embryos. Dev Dyn 2009;238:656–663.
67. Spomer W, Pfriem A, Alshut R, Just S, Pylatiuk C. High-throughput screening of zebrafish embryos using automated heart detection and imaging. J Lab Automat 2012;17:435–442.
68. Stern O, Marée R, Aceto J, Jeanray N, Muller M, Wehenkel L, *et al.* Automatic localization of interest points in zebrafish images with tree-based methods. Pattern Recog Bioinform 2011:179–190.
69. Tay TL, 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 2011;2:171.
70. Temerinac-Ott M, Ronneberger O, Ochs P, Driever W, Brox T, Burkhardt H. Multiview deblurring for 3-D images from light-sheet-based fluorescence microscopy. IEEE Transact Image Proc 2012;21:1863–1873.
71. Walker SL, Ariga J, Mathias JR, Coothankandaswamy V, Xie X, Distel M, *et al.* Automated reporter quantification *in vivo*: high-throughput screening method for reporter-based assays in zebrafish. PLoS One 2012;7:e29916.
72. Weger BD, Weger M, Nusser M, Brenner-Weiss G, Dickmeis T. A chemical screening system for glucocorticoid stress hormone signaling in an intact vertebrate. ACS Chem Biol 2012;7:1178–1183.
73. Wittmann C, Reischl M, Shah AH, Mikut R, Liebel U, Grabher C. Facilitating drug discovery: an automated high-content inflammation assay in zebrafish. Journal of visualized experiments: JoVE 2012;65:e4203.
74. Xu X, Xu X, Huang X, Xia W, Xia S. A high-throughput analysis method to detect regions of interest and quantify zebrafish embryo images. J Biomol Screen 2010;15:1152–1159.
75. Zanella C, Campana M, Rizzi B, Melani C, Sanguinetti G, Bourgine P, *et al.* Cells segmentation from 3-D confocal images of early zebrafish embryogenesis. IEEE Transact Image Proc 2010;19:770–781.
76. Emmenlauer M, Ronneberger O, Ponti A, Schwarb P, Griffa A, Filippi A, *et al.* XuvTools: free, fast and reliable stitching of large 3D datasets. J Microsc 2009;233:42–60.
77. Huisken J, Swoger J, Del Bene F, Wittbrodt J, Stelzer EHK. Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science 2004;305:1007.
78. Preibisch S, Saalfeld S, Schindelin J, Tomancak P. Software for bead-based registration of selective plane illumination microscopy data. Nat Methods 2010;7:418–419.
79. Liu Y, Yu Y. Interactive image segmentation based on level sets of probabilities. IEEE Transact Visualization Comput Graphics 2011;18:1.
80. Hockendorf B, Thumberger T, Wittbrodt J. Quantitative analysis of embryogenesis: a perspective for light sheet microscopy. Dev Cell 2012;23:1111–1120.
81. Meijering E, Dzyubachyk O, Smal I. Methods for cell and particle tracking. Methods Enzymol 2012;504:183–200.
82. Mikula K, Peyrieras N, Remesiková M, Smisek M. 4D numerical schemes for cell image segmentation and tracking. Finite Volumes for Complex Applications VI Problems & Perspectives, Springer Proceedings in Mathematics Volume 4, 2011, pp. 693–701.
83. Pan YA, Livet J, Sanes JR, Lichtman JW, Schier AF. Multicolor brainbow imaging in zebrafish. Cold Spring Harb Protoc 2011;2011:pdb prot5546.
84. Culic-Viskota J, Dempsey WP, Fraser SE, Pantazis P. Surface functionalization of barium titanate SHG nanoprobes for *in vivo* imaging in zebrafish. Nat Protoc 2012;7:1618–1633.
85. Pantazis P, Maloney J, Wu D, Fraser SE. Second harmonic generating (SHG) nanoprobes for *in vivo* imaging. Proc Natl Acad Sci U S A 2010;107:14535–14540.
86. Eliceiri KW, Berthold MR, Goldberg IG, Ibanez L, Manjunath BS, Martone ME, *et al.* Biological imaging software tools. Nat Methods 2012;9:697–710.
87. Kankaanpaa P, Paavolainen L, Tiitta S, Karjalainen M, Paivarinne J, Nieminen J, *et al.* BioImageXD: an open, general-purpose and high-throughput image-processing platform. Nat Methods 2012;9:683–689.

88. Kvilekval K, Fedorov D, Obara B, Singh A, Manjunath BS. Bisque: a platform for bioimage analysis and management. Bioinformatics 2010;26:544–552.
89. Carpenter A, Jones T, Lamprecht M, Clarke C, Kang I, Friman O, *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 2006;7:R100.
90. Marée R, Stevens B, Rollus L, Rocks N, Moles-Lopez X, Salmon I, *et al.* A Rich Internet Application for Remote Visualization and Collaborative Annotation of Digital Slide Images in Histology and Cytology. 12th European Congress on Telepathology and 5th International Congress on Virtual Microscopy, 2012.
91. Stegmaier J, Alshut R, Reischl M, Mikut R. Information fusion of image analysis, video object tracking, and data mining of biological images using the open source MATLAB toolbox Gait-CAD. Biomed Tech 2012;57:458–461.
92. de Chaumont F, Dallongeville S, Chenouard N, Herve N, Pop S, Provoost T, *et al.* Icy: an open bioimage informatics platform for extended reproducible research. Nat Methods 2012;9:690–696.
93. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, *et al.* Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9:676–682.
94. Sommer C, Straehle C, Kothe U, Hamprecht FA. ilastik: Interactive Learning and Segmentation Toolkit. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2011. Available at: http://dx.doi.org/10.1109/ISBI.2011.5872394
95. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, *et al.* KNIME: the konstanz information miner. In: Data Analysis, Machine Learning and Applications. Preisach, C, Burkhardt, H, Schmidt-Thieme, L, Decker R, (eds), Springer, 2008, pp. 319–326.
96. Allan C, Burel JM, Moore J, Blackburn C, Linkert M, Loynton S, *et al.* OMERO: flexible, model-driven data management for experimental biology. Nat Methods 2012; 9:245–253.
97. Mikut R, Reischl M. Data mining tools. Wires Data Min Knowl 2011;1:431–443.
98. Liu T, Nie J, Li G, Guo L, Wong ST. ZFIQ: a software package for zebrafish biology. Bioinformatics 2008;24:438–439.
99. Xia S, Zhu Y, Xu X, Xia W. Computational techniques in zebrafish image processing and analysis. J Neurosci Methods 2013;213:6–13.
100. Kanungo J, Cuevas E, Ali SF, Paule MG. Ketamine induces motor neuron toxicity and alters neurogenic and proneural gene expression in zebrafish. J Appl Toxicol 2013;33: 410–417.
101. Fowlkes CC, Hendriks CL, Keranen SV, Weber GH, Rubel O, Huang MY, *et al.* A quantitative spatiotemporal atlas of gene expression in the *Drosophila* blastoderm. Cell 2008;133: 364–374.
102. Peng H, Chung P, Long F, Qu L, Jenett A, Seeds AM, *et al.* BrainAligner: 3D registration atlases of *Drosophila* brains. Nat Methods 2011;8:493–500.
103. Ahrens MB, Li JM, Orger MB, Robson DN, Schier AF, Engert F, *et al.* Brain-wide neuronal dynamics during motor adaptation in zebrafish. Nature 2012;485:471–477.
104. Linkert M, Rueden CT, Allan C, Burel JM, Moore W, Patterson A, *et al.* Metadata matters: access to image data in the real world. J Cell Biol 2010;189:777–782.
105. Lynnes C, Mack R. KDD services at the Goddard earth sciences distributed active archive center. In: Data Mining for Scientific and Engineering Applications. Grossman, RL, Kamath, C, Kegelmeyer, P, Kumar, V, Namburu R, (eds), Springer, 2001, pp. 165–181.
106. Shvachko K, Kuang H, Radia S, Chansler R. The Hadoop Distributed File System. IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST), 2010. Available at: http://dx.doi.org/10.1109/MSST.2010.5496972
107. Meijering E, de Chaumont F, Smal I, Chenouard N, Maška M, Olivo-Marin JC. The 2012 Particle Tracking Challenge: An Objective Comparison of Particle Tracking Methods. BioImage Informatics Conference (BII 2012), Dresden, 2012.
108. Cavodeassi F, Bene FD, Fürthauer M, Grabher C, Herzog W, Lehtonen S, *et al.* Report of the 2nd European Zebrafish Principal Investigator Meeting in Karlsruhe, Germany, March 21–24, 2012. Zebrafish 2013;10:119–123.

Address correspondence to:
Ralf Mikut, PhD
Karlsruhe Institute of Technology
P.O. Box 3640
Karlsruhe 76021
Germany

E-mail: ralf.mikut@kit.edu