

Small molecule screening by imaging Ulrike S Eggert* and Timothy J Mitchison

Useful small molecule tools can be discovered in imaging screens that measure phenotype in single cells or small organisms. Recent examples include identification of small molecule inhibitors of processes such as cell migration, cytokinesis, mitotic spindle length determination, melanogenesis, aggresome formation, membrane transport and nuclear export. Imaging screens are currently limited by challenges in the areas of image analysis and target identification. We discuss the use of model organisms such as zebrafish in screens and review different methods of target identification. The emerging field of automated image analysis is also introduced.

Addresses

Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA

^{*} Current address: Dana-Farber Cancer Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 01225, USA.

Corresponding author: Eggert, Ulrike S (ulrike_eggert@hms.harvard.edu)

Current Opinion in Chemical Biology 2006, 10:232-237

This review comes from a themed issue on Combinatorial chemistry and molecular diversity Edited by Scott K Silverman and Paul J Hergenrother

Available online 8th May 2006

1367-5931/\$ - see front matter © 2006 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.cbpa.2006.04.010

Introduction

Small molecules are valuable tools to study dynamic biological processes with high temporal resolution and can catalyze therapeutic drug discovery [1]. Active small molecules can be discovered by screening chemical libraries for alteration of a specific protein activity in pure protein in vitro assays or for a desired phenotype in cellbased assays [2]. In the past, cell-based assays have usually been established to report on effects on the activity of a single pathway or process, typically using as the readout a luminescence or fluorescence signal measured in a plate reader. In cell imaging assays, fields of cells are visualized using fluorescent tags attached to different macromolecules, and automated microscopy. Such data potentially contain a large amount of information relevant to desired as well as unexpected phenotypic changes, and the approach is sometimes called 'high content screening' [3]. The potential of this method in

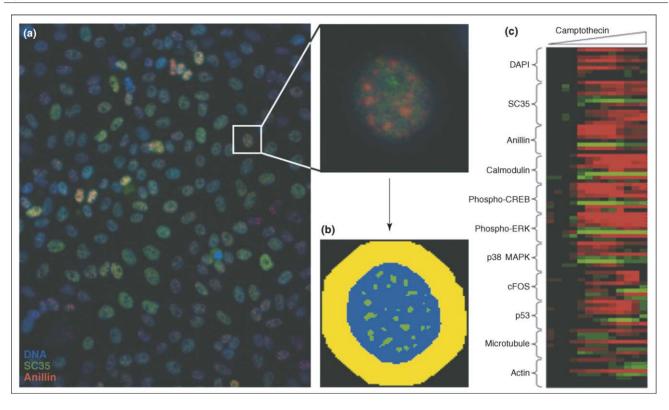
academia and industry for the discovery and characterization of useful compounds is high, but many challenges remain, particularly at the level of target identification and data analysis. In this review, we discuss recent progress, considering both small molecule screens and RNAi screens, which use similar methods and face related challenges.

Small molecule imaging screens typically involve the following steps [4,5]. First, cells are plated into opticalbottom multiwell plates, treated with small molecules and incubated for an appropriate time. Proteins of interest are then rendered fluorescent by some combination of small molecule fluorescent probes, immunofluorescence with antibodies, and expression of green fluorescence protein (GFP)-tagged proteins in cells. The last method can be applied to both live and fixed cells. Next, fluorescent images are captured by automated microscopy. The images are then analyzed to provide measures of phenotypic change, identifying wells that contain desired, undesired or unexpected phenotypes. Examples of phenotypic images are shown in Figures 1a and 2. In small molecule screening, it is eventually necessary to identify the biochemical target causing the phenotypic change. This can be challenging, especially if phenotypic change results from perturbation of more than one macromolecule.

Recent screens have provided the community with useful small molecule tools to study cellular processes of varied complexity, such as cell migration [6°], cytokinesis [7°°], mitotic spindle length determination [8°], melanogenesis [9°], aggresome formation [10°], membrane transport [11°] and nuclear export [12°]. Many of these screens result in a variety of small molecules that target different facets of the process. This is well illustrated by the discovery of small molecules that affect different steps of the PI3K/PTEN/Akt signaling cascade that controls nuclear export of the Forkhead family of transcription factors [12°], a process mutated in many cancers.

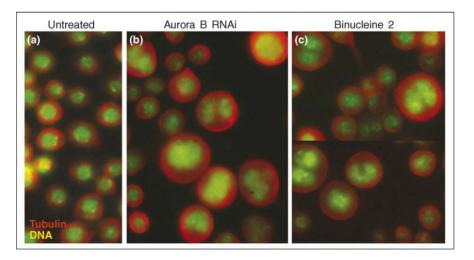
When discussing small molecule phenotypic screens, it is important to acknowledge the synergy from the parallel development of genome-wide RNAi screening. Although small molecules and RNAi function through different mechanisms [13], they can have similar effects on cells and can employ the same screening methods. Because genome-wide RNAi screens are now possible in mammalian and *Drosophila* cells, and in *Caenorhabditis elegans* animals, there has been a surge of interest in the technologies required for successful screening [14]. For example, the CellProfiler image analysis software was recently

Figure 1



Use of image analysis to generate a multidimensional phenotypic profile [26**]. (a) Primary image of human cells stained with probes for DNA and two nuclear proteins, Sc35 and Anillin, acquired from a 384 well plate. (b) The image was segmented to identify individual nuclei. An annulus around each nucleus was drawn to represent cytoplasm. Intensity, shape and number parameters were collected for each probe on a cell-by-cell basis. These data were used to compare treated wells to controls by statistical testing. This procedure was repeated using probes for 10 proteins. (c) Multidimensional phenotypic profile describing the action of the drug camptothecin from image analysis. Each column represents a different dose of drug, increasing from left to right. Each row represents measurements using a different parameter from image analysis, grouped by fluorescent probe. Each pixel shows the Z-score from a statistical test, black is no difference from controls, red and green are significantly different. Such profiles were used to compare different drugs.

Figure 2



Comparison between small molecule and RNAi phenotypes [7**]. Tubulin is shown in red, DNA in green. (a) untreated Drosophila Kc167 cells, (b) cells exposed to dsRNA targeting Aurora B kinase for 4 days, (c) cells treated with 100 µM of the small molecule binucleine 2 for 24 h. Note that the phenotypes in (b) and (c) are similar and differ from (a). They contain a combination of binucleate cells and cells with large diffuse DNA. released to analyze RNAi data in single cells (http:// www.cellprofiler.org), but it should be equally applicable to small molecule data. The community studying C. elegans, where RNAi screens first became available, has systematically defined and characterized phenotypes allowing standardized phenotypic profiling [15°,16]. As more screens are carried out and phenotypic profiling becomes routine in all organisms, comparisons of small molecule and RNAi phenotypes will be an important tool in small molecule target identification. We have made a first attempt at this strategy, which is discussed below. In addition to technology sharing, small molecules and RNAi can also be combined in modifier screens to explore mechanistic connectivity.

To cause a specific phenotypic change in an imaging screen, a small molecule must be active on living cells (either by targeting cell surface molecules or by permeating into the cytoplasm), and it must also display some specificity in its activity. These are major advantages over pure protein screens. However, phenotypic screens are not as commonly used because they still face some experimental limitations, especially in the areas of image analysis and target identification, which are discussed below in more detail.

Screening in model organisms

A significant limitation of cell-based assays is the challenge of finding single cell systems that recapitulate the complex biology of tissues and diseases. This can be addressed to some extent by using primary cells, but a more general approach is phenotypic screening in whole organisms [17,18]. Small model organisms, including embryos of zebrafish and C. elegans nematodes, and perhaps also flies, can be cultured in multiwell plates and treated with large numbers of diverse compounds [19–21]. Imaging is the logical readout for such screens, using either classic morphology, or GFP expressed in specific cells or organs. In two recent zebrafish examples, small molecules were identified that suppress the effect of a mutation causing a cardiovascular defect [22**], and that suppress a mutation causing mitotic arrest by an Sphase delay [23**]. In both cases, the screens were performed on mutant organisms, looking for compounds that rescued a molecularly defined defect. This approach has the advantage of targeting the screen to a specific area of biology, and also of requiring an absence of toxic side effects in the hits. C. elegans embryos also appear ideal for screening in that they are small, robust, well characterized genetically, and much used for RNAi screening.

Image analysis

Perhaps the single largest challenge in image-based screening is data analysis. This problem starts with the sheer quantity of data; several terabytes is not untypical for a single screen. Storing and retrieving such large

amounts of data can be challenging [24]. Increasingly, relational databases are used to facilitate image analysis, although it may be most efficient to store only metainformation about images (information on what the image is, and the results of image analysis) in the database, and store the actual images as ordinary files. One approach to databases for image handling is provided by the Open Microscopy Initiative (OME), and open-source academic consortium [25].

Once the images are collected and stored, the next problem is to turn them into quantitative phenotypic information. Although we know that images are rich in biological information, there is currently no single best method to quantitate and extract this information. This problem must be solved on a screen-to-screen basis, and optimal solutions typically require interaction of screeners with image analysis experts.

Image analysis typically starts by identification of individual cells and also of regions of interest within each cell (nuclei, organelles, etc.). Next, measurements of intensity, size, shape and texture are made in the identified regions. The more probes that are multiplexed in a given experiment, the more measurements can be made. Ten or more measurements for a single problem are possible [26°,27°,28°]. In some cases, the measurement is obviously biologically relevant, for example the fraction of a signaling protein in the cytoplasm vs. nucleus. In other cases, the interpretation is non-obvious, such as a measure of the texture and size of a nucleus. Biologists naturally focus on interpretable measurements, and many screens have been based on these. However, several studies have demonstrated that measurements on cells that are not obviously interpretable nevertheless contain useful phenotypic information (Figure 1) [26°,27°,28°]. As a practical issue, images can be analyzed with homegrown software, or commercial packages. Commercial software tends to be more reliable and 'user friendly', but it can be expensive and may not provide the flexibility needed to ask specific questions. An interesting development is the release of open-source image analysis packages, for example, CellProfiler and ImageJ (http:// rsb.info.nih.gov/ij/). As these attract interactive user communities, they may ultimately offer the best of both worlds.

Most scientists approaching a screen tend to view phenotype in a relatively one-dimensional way, asking questions about specific pathways or processes. However, cells contain many different features, and cell images can almost always provide information on more than one of them. For example, DNA morphology reports on cell cycle state, but also on protein synthesis, apoptosis and other pathways [29]. Thus, phenotype is inevitably multidimensional, and is best described not by a single number, but by a vector in some multi-dimensional

'phenotype space'. A familiar example of multidimensional analysis is transcript profiling, where the levels of many thousands of RNAs are used as a measure of phenotype [30]. Powerful methods have been developed for comparing and quantifying phenotypes from RNA profiling. Progress has been made on the problem of turning images of fixed cells in into quantitative biological information [26°,27°,28°], but this field is still in its infancy.

Time lapse movies of cells expressing GFP-tagged proteins contain large amounts of information that are potentially complementary to data from fixed cells. Currently, the challenge of storing and analyzing movies is sufficiently daunting that there have been few examples of live imaging screens. The large information content inherent in movies was illustrated in a genome-wide RNAi screen in C. elegans, where human interpretations of phenotypes from movies were integrated into a multidimensional phenotypic profile capable of distinguishing many different cellular defects in early embryos [15**]. Using advanced computational methods, it may be possible to fully automate aspects of phenotype scoring in C. elegans [31]. Important information can also be extracted from movies using fairly simple analysis methods. In one recent example, automated analysis of periodic motion was used to score compounds from a small library for effects on heartbeat in zebrafish [32]. This high-throughput method was as effective as, and much cheaper than, testing on dogs for detecting adverse effects, and may allow such testing much earlier in the drug development process. As the informatics challenges are progressively solved, we expect to see more use of live cell imaging in small molecule screening.

Target identification

Target identification has always been a rate-limiting step in cell-based small molecule screens. Advances in image analysis and RNAi screening have added to the toolkit for solving this problem. Traditionally, small molecule targets have been identified by affinity methods or candidate-based approaches. Both of these approaches can work well in some circumstances and can be more challenging in others. For example, a candidate-based approach will not work if the target protein was previously unknown. Sometimes the protein target is already known and the small molecule can be used to uncover a new role of the protein (e.g. a screen for inhibitors of aggresome formation discovered an unexpected role for histone deacetylase in this process [10°]). Usually, a candidatebased approach works best if the target is determined through a series of secondary assays. In a screen for small molecules that inhibited cell migration, a very broad problem that involves many cellular machines, Yarrow et al. used secondary assays to narrow down potential targets [6°]. They discovered that the small molecule Rockout showed a phenotype consistent with inhibition of the Rho pathway and inhibited Rho kinase in vitro. In a similar strategy, the small molecule secramine, which inhibits membrane export from the Golgi to the plasma membrane, was shown to inhibit the small GTPase Cdc42 in a RhoGDI-dependent manner [11°]. In some assays, the readout can be designed with specific candidate targets in mind. For example, Vogt and colleagues screened for small molecules that caused accumulation of phosphorylated Erk, a member of the MAPK cascade [33]. Some of the compounds inhibited MKP-3, a phosphatase targeting Erk, *in vitro* and overcame the effects of overexpressed MKP-3, suggesting that they target this enzyme. To facilitate target ID by candidate testing, it is useful to include in any screen a collection of compounds with known mechanisms of action, including toxins and medicines. The valuable information that can result from screening such collections has been discussed [34]. Collections of several hundred biologically active molecules are commercially available (e.g. ICCB Known Bioactives from Biomol; http://www.biomol.com).

In affinity methods, a small molecule is tethered to a solid support and cell extract is passed over this support to see which proteins will bind. For this approach to be successful, the small molecule has to be attached to a linker without changing its activity. This will be easier if a chemical library is designed to include a linkage site. A tagged triazine library has been used in several screens [35], most recently to identify small molecules that increased pigmentation [9°]. One of these small molecules, melanogenin, inhibited prohibitin, a mitochondrial protein not previously known to be involved in melanin generation. Affinity chromatography may also be easier if the library is screened for activity in cell-free systems, so the same system can be used directly as the input for biochemical purification. In one example, a screen for alterations in mitotic spindle morphology in Xenopus extract identified a compound, diminutol, that targets NAPD-dependent oxidoreductase [8°]. In an interesting combination of affinity chromatography and automated image analysis, Tanaka et al. compared the morphological profiles of kinase inhibitors with various structural scaffolds and showed that a compound with an unusual profile unexpectedly targeted carbonyl reductase 1 [27**].

We developed a target identification strategy complementary to affinity and candidate-based approaches, which involves comparisons between small molecule and RNAi phenotypes. If the phenotype is the same, the protein targeted by RNAi and the small molecule is likely to be the same. In our proof-of-principle study, we expanded this simple concept to a genome-wide level. We carried out parallel genome-wide RNAi and small molecule screens to identify proteins involved in cytokinesis, the final step of cell division, and small molecules that target these proteins [7**]. We grouped the RNAi phenotypes into different categories and compared these to small molecule phenotypes. Using this strategy, we were able to identify a small molecule inhibitor of the Aurora B kinase pathway (Figure 2). Phenotypic comparisons also have limits because the phenotype of small molecule inhibition might be different from an RNAi phenotype, where the protein has been depleted from the cell [13]. A small molecule could activate a protein or inhibit one domain of a multi-domain complex, which might result in a different cellular effect. The identification of small molecule targets remains a challenge, but some encouraging progress has been made and a combination of all existing strategies should make it less of an obstacle to phenotypic screening by imaging in the future.

Conclusions

Image-based screening has huge potential in small molecule discovery, but until recently has been limited by technical challenges. Phenotypic screens are now possible in model organisms such as zebrafish, making it an ideal platform to discover non-toxic small molecules that affect multicellular processes such as heart function. Parallel developments in small molecule and genome-wide RNAi screening have given a boost to screening by imaging especially in the area of automated image analysis and phenotypic profiling. A combination of these new methods and traditional approaches has been useful in identifying small molecule targets. Challenges still exist, but recent developments have been very positive and should allow screening by imaging to become more widely used in the future.

Acknowledgements

We thank Christine Field for helpful comments on the manuscript. USE was supported by a Merck-sponsored fellowship from the Helen Hay Whitney Foundation. Research on small molecule screening in the Mitchison lab was supported by grants from the National Institutes of Health and the National Cancer Institute (PO1 CA078048 and GM023928-25), Merck and Merck KGA.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest
- Smukste I. Stockwell BR: Advances in chemical genetics. Annu Rev Genomics Hum Genet 2005. 6:261-286
- Spring DR: Chemical genetics to chemical genomics: small molecules offer big insights. Chem Soc Rev 2005, 34:472-482.
- Abraham VC, Taylor DL, Haskins JR: High content screening applied to large-scale cell biology. Trends Biotechnol 2004,
- Yarrow JC, Feng Y, Perlman ZE, Kirchhausen T, Mitchison TJ: Phenotypic screening of small molecule libraries by high throughput cell imaging. Comb Chem High Throughput Screen 2003. 6:279-286
- Mitchison TJ: Small-molecule screening and profiling by using automated microscopy. ChemBioChem 2005, 6:33-39
- Yarrow JC, Totsukawa G, Charras GT, Mitchison TJ: Screening for cell migration inhibitors via automated microscopy reveals a Rho-kinase inhibitor. Chem Biol 2005, 12:385-395.

This study sought small molecule inhibitors of cell migration during scratch wound healing. A series of secondary assays led to the discovery of a Rho kinase inhibitor.

- Eggert US, Kiger AA, Richter C, Perlman ZE, Perrimon N, Mitchison TJ, Field CM: **Parallel chemical genetic and genome**wide RNAi screens identify cytokinesis inhibitors and targets. PLoS Biol 2004. 2:e379.

Parallel screens identified proteins involved in cytokinesis and small molecule inhibitors. This is the first study to use systematic comparisons between small molecule and RNAi phenotypes as a target identification

- Wignall SM, Gray NS, Chang YT, Juarez L, Jacob R, Burlingame A,
- Schultz PG, Heald R: Identification of a novel protein regulating microtubule stability through a chemical approach. Chem Biol 2004, 11:135-146.

This paper describes an imaging screen in a cell-free system. Small molecules that affect the size of the mitotic spindle were discovered and the target of one compound identified by affinity chromatography.

- Snyder JR, Hall A, Ni-Komatsu L, Khersonsky SM, Chang YT,
- Orlow SJ: Dissection of melanogenesis with small molecules identifies prohibitin as a regulator. Chem Biol 2005, 12:477-484. This screen for small molecules that affect pigmentation used a library that was designed to include a tagging site. This facilitated target ID by affinity chromatography.
- Corcoran LJ, Mitchison TJ, Liu Q: A novel action of histone
- deacetylase inhibitors in a protein aggresome disease model. Curr Biol 2004, 14:488-492.

A small molecule that inhibited protein aggregation in a cellular model system for amyotrophic lateral sclerosis (ALS) targeted histone deacetylase, revealing a new role for this protein.

- 11. Pelish HE, Peterson JR, Salvarezza SB, Rodriguez-Boulan E,
- Chen JL, Stamnes M, Macia E, Feng Y, Shair MD, Kirchhausen T: Secramine inhibits Cdc42-dependent functions in cells and Cdc42 activation in vitro. Nat Chem Biol 2006, 2:39-46.

Secramine, an inhibitor of membrane traffic from Golgi to the plasma membrane, was shown to affect the small GTPase Cdc42. This is one of the first examples of a small molecule that targets a protein in pathways connected to Rho-family small GTPases.

- 12. Kau TR, Schroeder F, Ramaswamy S, Wojciechowski CL, Zhao JJ,
- Roberts TM, Clardy J, Sellers WR, Silver PA: A chemical genetic screen identifies inhibitors of regulated nuclear export of a Forkhead transcription factor in PTEN-deficient tumor cells. Cancer Cell 2003, 4:463-476.

This screen discovered small molecules that affect different steps of the nuclear export machinery and identified possible targets for cancer chemotherapy.

- 13. Eggert US, Field CM, Mitchison TJ: Small molecules in an RNAi world. Mol BioSyst 2006, 2:93-96.
- 14. Friedman A, Perrimon N: Genome-wide high-throughput screens in functional genomics. Curr Opin Genet Dev 2004, **14**:470-476.
- 15. Sonnichsen B, Koski LB, Walsh A, Marschall P, Neumann B,
- Brehm M, Alleaume AM, Artelt J, Bettencourt P, Cassin E et al.: Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature 2005, 434:462-469.

This genome-wide RNAi screen in C. elegans embryos gave the first example of comprehensive phenotypic profiling from the analysis of

- Gunsalus KC, Ge H, Schetter AJ, Goldberg DS, Han JD, Hao T, Berriz GF, Bertin N, Huang J, Chuang LS et al. Predictive models of molecular machines involved in Caenorhabditis elegans early embryogenesis. Nature 2005, 436:861-865
- 17. Carroll PM, Dougherty B, Ross-Macdonald P, Browman K, FitzGerald K: Model systems in drug discovery: chemical genetics meets genomics. Pharmacol Ther 2003, 99:183-220.
- 18. Zon LI, Peterson RT: In vivo drug discovery in the zebrafish. Nat Rev Drug Discov 2005. 4:35-44
- 19. Peterson RT, Link BA, Dowling JE, Schreiber SL: Small molecule developmental screens reveal the logic and timing of vertebrate development. Proc Natl Acad Sci USA 2000, 97:12965-12969.

- 20. Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, Ahringer J: Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature 2000, 408:325-330
- 21. Gonczy P, Echeverri C, Oegema K, Coulson A, Jones SJ, Copley RR, Duperon J, Oegema J, Brehm M, Cassin E et al.: Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature 2000, 408:331-336
- Peterson RT, Shaw SY, Peterson TA, Milan DJ, Zhong TP, Schreiber SL, MacRae CA, Fishman MC: Chemical suppression of a genetic mutation in a zebrafish model of aortic coarctation. Nat Biotechnol 2004, 22:595-599. See annotation to [23**].
- Stern HM, Murphey RD, Shepard JL, Amatruda JF, Straub CT, Pfaff KL, Weber G, Tallarico JA, King RW, Zon LI: **Small molecules** that delay S phase suppress a zebrafish bmyb mutant. Nat Chem Biol 2005, 1:366-370.

These studies are examples of screens in whole animals. The screens identify small molecules that can overcome mutations in zebrafish.

- Swedlow JR, Goldberg I, Brauner E, Sorger PK: Informatics and quantitative analysis in biological imaging. Science 2003, 300:100-102
- 25. Goldberg IG, Allan C, Burel JM, Creager D, Falconi A, Hochheiser H, Johnston J, Mellen J, Sorger PK, Swedlow JR: The open microscopy environment (OME) data model and XML file: open tools for informatics and quantitative analysis in biological imaging. Genome Biol 2005, 6:R47.
- 26. Perlman ZE, Slack MD, Feng Y, Mitchison TJ, Wu LF, Altschuler SJ: Multidimensional drug profiling by automated microscopy. *Science* 2004, **306**:1194-1198.

This study used automated image analysis to generate cytological profiles for cells that were treated with different drugs. Comparisons between the profiles of compounds with known and unknown mechanisms should be a useful target identification strategy in the future.

27. Tanaka M. Bateman R. Rauh D. Vaisberg E. Ramachandani S. Zhang C, Hansen KC, Burlingame AL, Trautman JK,

Shokat KM et al.: An unbiased cell morphology-based screen for new, biologically active small molecules. PLoS Biol 2005, 3:e128.

A screen followed by unbiased automated image analysis was used to discover a small molecule with an unusual profile. The target of the small molecule was identified and characterized

28. Huang K, Murphy RF: From quantitative microscopy to automated image understanding. J Biomed Opt 2004, 9:893-912.

Recent review of image analysis methods.

- 29. Perlman ZE, Mitchison TJ, Mayer TU: High-content screening and profiling of drug activity in an automated centrosome-duplication assay. ChemBioChem 2005, **6**:145-151.
- 30. Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998, 95:14863-14868.
- 31. Ning F, Delhomme D, LeCun Y, Piano F, Bottou L, Barbano PE: Toward automatic phenotyping of developing embryos from videos. IEEE Trans Image Process 2005, 14:1360-1371.
- 32. Burns CG, Milan DJ, Grande EJ, Rottbauer W, MacRae CA Fishman MC: High-throughput assay for small molecules that modulate zebrafish embryonic heart rate. Nat Chem Biol 2005, 1.263-264
- 33. Vogt A, Cooley KA, Brisson M, Tarpley MG, Wipf P, Lazo JS: Cell-active dual specificity phosphatase inhibitors identified by high-content screening. Chem Biol 2003, 10:733-742.
- 34. Root DE, Flaherty SP, Kelley BP, Stockwell BR: Biological mechanism profiling using an annotated compound library. Chem Biol 2003, 10:881-892.
- Moon HS, Jacobson EM, Khersonsky SM, Luzung MR, Walsh DP, Xiong W, Lee JW, Parikh PB, Lam JC, Kang TW et al.: A novel microtubule destabilizing entity from orthogonal synthesis of triazine library and zebrafish embryo screening. J Am Chem Soc 2002. 124:11608-11609.