Automated microscopy for high-content RNAi screening

Christian Conrad1 and Daniel W. Gerlich2

1Advanced Light Microscopy Core Facility, European Molecular Biology Laboratory Heidelberg, D-69117 Heidelberg, Germany
2Institute of Biochemistry, Swiss Federal Institute of Technology (ETH) Zurich, CH-8093 Zurich, Switzerland

Fluorescence microscopy is one of the most powerful tools to investigate complex cellular processes such as cell division, cell motility, or intracellular trafficking. The availability of RNA interference (RNAi) technology and automated microscopy has opened the possibility to perform cellular imaging in functional genomics and other large-scale applications. Although imaging often dramatically increases the content of a screening assay, it poses new challenges to achieve accurate quantitative annotation and therefore needs to be carefully adjusted to the specific needs of individual screening applications. In this review, we discuss principles of assay design, large-scale RNAi, microscope automation, and computational data analysis. We highlight strategies for imaging-based RNAi screening adapted to different library and assay designs.

Microscopy for functional genomics and systems biology

High-throughput methodologies such as proteomics, expression profiling, or protein interaction mapping have established an era of systems biology, which aims at understanding biological processes through comprehensive identification of network components and their interplay (Hartwell et al., 1999). Such large-scale approaches can generate valuable parts lists and provide basic insight into their modular organization, but they do not resolve spatial and temporal aspects of protein function and regulation (Meganos and Fraser, 2007). Most biological processes occur spatially confined at distinct subcellular sites and vary between different cells, thus calling for methods capable of sampling spatial and temporal patterns at the single cell level.

Fluorescence microscopy provides an ideal tool to study complex biological processes with high spatiotemporal resolution. Fluorescent proteins allow one to label virtually any cellular structure or signaling component under physiological conditions in live cells (Giepmans et al., 2006). A wide range of fluorescent biosensors and imaging modalities provides the possibility to detect steady-state protein dynamics, post-translational modifications, protein–protein interactions, and small molecules (Lippincott-Schwartz et al., 2003; Giepmans et al., 2006). Microscopy has long been tedious and difficult to perform in a systematic and quantitative way. Therefore, imaging-based assays have in most cases been restricted to manual low-throughput experiments, for example, detailed mechanistic studies of few selected candidate genes. Recent developments in robotics for sample preparation and automation of microscope control now enable one to perform imaging at a large scale (Pepperkok and Ellenberg, 2006). The key challenge often remains the annotation of complex phenotypic patterns in huge image datasets. Many studies still rely on visual scoring and manual annotation, which is slow, error prone, and potentially biased by the user. Significant progress has been made through the implementation of computer vision methods for multidimensional data analysis (Gerlich et al., 2001; Gerlich and Ellenberg, 2003) and supervised machine learning approaches for automated classification of cellular and subcellular phenotypes (Conrad et al., 2004; Neumann et al., 2006; Glory and Murphy, 2007; Jones et al., 2009; Walter et al., 2009).

In this review, we provide an overview of imaging-based screening strategies. We focus on biological assay design, automated image acquisition, and computational analysis. We further discuss advanced imaging options and how throughput and content of screening assays can be balanced. Finally, we present a perspective on how integration of experimental robotics, image analysis tools, and large-scale data resources may be used to further automate the discovery process.

Biological assays: content versus throughput

The most basic readout for imaging-based assays is total cellular fluorescence intensity of immunodetected antigens or overexpressed fluorescent reporters (Fig. 1 A). For example, this can...
be used to score the expression of marker genes (Müller et al., 2005; Loo et al., 2007), DNA content for cell cycle progression (Kittler et al., 2007), lipoprotein uptake (Bartz et al., 2009), mitochondrial Ca\textsuperscript{2+} transport (Jiang et al., 2009), or virus entry into cells (Pelkmans et al., 2005; Brass et al., 2008; Krishnan et al., 2008; Plouffe et al., 2008).

Another class of assays scores cellular morphology features (Fig. 1 B). For example, the pattern of cytoskeletal or chromatin markers can serve to probe cellular morphologies (Bakal et al., 2007; Liu et al., 2009), cell division phenotypes (Gönczy et al., 2000; Echard et al., 2004; Sönnichsen et al., 2005; Neumann et al., 2006; Draviam et al., 2007; Goshima et al., 2007), cell cycle progression (Boutros et al., 2004; Kittler et al., 2007), or DNA double-strand break repair (Doil et al., 2009). Although manual annotation of such assays is possible, this way of analyzing the images is very tedious and may be user biased. Fortunately, computational machine learning methods allow efficient annotation even of subtle morphological features (see Computational image analysis for quantitative phenotyping).

Fluorescent proteins can also be used to assay biochemical events in live cells. GFP-based biosensors have been engineered for visualization of protein–protein interactions (Ciruela, 2008) and posttranslational modifications (Aye-Han et al., 2009) as well as enzyme activity and small molecules (VanEngelenburg and Palmer, 2008). Imaging modalities such as fluorescence correlation spectroscopy (Haustein and Schwille, 2007), photobleaching and photoactivation (Lippincott-Schwartz et al., 2003), and chemical labeling of engineered target proteins (Johnsson, 2009) further enable the study of steady-state protein dynamics in living cells. All of these methods can, in principle, be applied to high-throughput imaging assays, opening new possibilities to screen for factors in very specific aspects of cellular signaling.

Time-resolved live imaging (Fig. 1 C) provides the highest content for assays on complex dynamic processes such as cell division (Sönnichsen et al., 2005; Neumann et al., 2006). Live cell imaging can be automated to a level that enables genome-scale RNAi screens (MitoCheck project; Neumann et al., 2010). Several commercial microscope platforms support automated time-lapse imaging of live cells. The key to live imaging-based screening is a stable incubation and careful optimization of the light dose to avoid photodamage (Schmitz and Gerlich, 2009).

The most severe limitation of live imaging in screening is the complexity of data annotation.

Screening of hypothesis-derived candidate gene sets

Many initial functional genomics screens relied on simple intensity-based readouts, which can be scaled relatively easily to the full genome level. However, this approach provides little information on the underlying phenotype and its variability within the cell population. The application of higher-content imaging assays that incorporate spatial or temporal phenotypic patterns is much more labor intensive and therefore can require preselection of candidate genes to perform low- to medium-throughput screens (Fig. 2 A).
Candidate gene sets can be derived from a low-level functional assay in a genome-wide screen. Alternatively, published information (e.g., from other large-scale experiments or public databases) can serve to compile target gene lists. At this level, a basic hypothesis can be incorporated into the screening strategy. For example, it may be reasonable to screen only proteins localizing to a certain subcellular structure (Skop et al., 2004) or defined classes of enzymes like kinases (Pelkmans et al., 2005). The level of detail needed in a functional genomics assay as well as the number of experimental perturbation conditions will vary considerably depending on the biological question. In some applications, it may be advantageous to sample only very few experimental conditions with ultra high-content assays such as time-lapse imaging, whereas others may not require sophisticated assays and/or lack a rationale to define reasonable sets of candidate genes.

In any screening scenario, assay optimization affects both sample preparation as well as data analysis strategy. The assay optimization often requires more effort and time than that to perform the actual screen (Fig. 2 B). Once an assay has been established, it needs to be scaled and tested for realistic screening conditions. For this, automated microscopy and computational image analysis parameters have to be carefully adjusted, which still remains the most difficult part for cell biologists. Next, negative controls and selected perturbation conditions representative for the subsequent large-scale screens should be used to perform pilot screens. It is advisable to establish standardized quality control procedures already at this stage to ensure subsequent consistent recording of the full dataset.

**RNAi in screening applications**

The discovery of RNAi as a conserved gene-silencing mechanism (Fire et al., 1998; Meister and Tuschl, 2004) has been the starting point for functional genomics studies in organisms previously inaccessible to systematic genetic perturbations. Several types of reagents have been developed to induce RNAi, which have also been used in screening applications. In mammalian cells, the most common are chemically synthesized short double-stranded RNA (dsRNA) oligomers (siRNAs; Elbashir et al., 2001), which can be obtained commercially as libraries targeting the entire genome or subsets targeting specific processes or gene families. As an alternative to chemical
synthesis of siRNAs, long dsRNAs can be enzymatically digested in vitro by Dicer or RNase III into heterogeneous pools of endo-RNase−prepared siRNAs (esiRNAs; Yang et al., 2002; Buchholz et al., 2006). A third method to induce RNAi uses the expression of short hairpin RNAs from plasmid vectors, which are cleaved into siRNAs by the endogenous Dicer enzyme (Root et al., 2006; Snoe and Rossi, 2006; Wiznerowicz et al., 2006). In contrast to the transient knockdown mediated by siRNA or esiRNA, the latter method enables stable RNAi in long-term gene-silencing studies. However, it requires transfection methods that can be less efficient and more toxic than transfection of siRNAs. Furthermore, the levels of vector-expressed short hairpin RNA are often difficult to control, which can be critical for potential concentration-dependent off-target effects (Jackson et al., 2006).

In some cell types like Drosophila melanogaster S2 cells, RNAi can be induced by simply adding dsRNA directly to the culture medium. However, most other cell types require transfection methods to deliver RNAi reagents. siRNAs or esiRNAs can be efficiently transfected by chemical transfection reagents using liquid-handling robotics and multiwell plates. As an alternative method, cell transfection arrays have been developed that contain dried spots of siRNA oligonucleotides mixed with transfection and matrix reagents. A uniform layer of cells seeded onto transfection arrays is then transfected locally at each spot by solid-phase transfection (Ziauddin and Sabatini, 2001; Erle et al., 2004, 2007; Erle and Pepperkok, 2007). Thereby, a large number of replica transfection plates can be generated in a single step. These can be distributed and stored for subsequent use in various screening applications without the need of further robotics (Neumann et al., 2006). Although transfection arrays are cheap and convenient to use, establishing the challenging production workflow and quality control procedures might be solved best in large central facilities. The main limitation of transfection arrays is the risk of cross-contamination from neighboring spots. However, solid-phase transfection can also be applied in multiwell plates, thereby avoiding cross-contamination between different siRNA oligonucleotides (Erle et al., 2008).

**Microscope automation**

Initial imaging-based screening was often implemented on standard motorized epifluorescence microscopes, to which the automation was added by academic researchers. This requires motorized control of stage positioning, fluorescence filters, and camera acquisition, which is now available as preassembled systems from all major microscope providers. One of the most challenging tasks in microscope automation is the focus control (Shen et al., 2006). Initial systems implemented image-based autofocus methods that first record entire z stacks of images to determine the focal plane with maximal information content. This method has the advantage to correctly position the focus even when cellular geometry changes, for example, in morphological phenotypes. However, image-based autofocus is slow and it exposes the specimen to excessive light-causing photodamage. An alternative autofocusing method measures the reflection of a laser at the surface of the tissue culture dish or multiwell plate. This method is appropriate to image adherent cells at a defined z offset, and it is very robust, fast, and minimizes perturbation of the specimen by light exposure. Reflection-based autofocusing requires specific hardware that has recently become available from most major microscope companies.

Another difficulty in automating microscopy for screening applications is the supply of objective immersion media. Therefore, most screens use dry objectives. With short working distance, the numerical aperture available for dry objectives will suit the requirements for most image-based screening assays. However, systems for automated water supply on immersion objectives with higher numerical aperture have been developed by some microscope companies.

With the increased demand for automated microscopes, several companies developed dedicated screening microscopes. These were optimized for throughput and robustness, but they are often less flexible for specific assay requirements. For example, they use proprietary software and image formats, which limit their integration into specialized image analysis pipelines. As a more flexible alternative, powerful open source software has been developed for microscope control, database integration, and image processing (Text box 1).

**Computational image analysis for quantitative phenotyping**

Automated image processing is the key for reliable quantitative measurements in high-throughput microscopy. After image preprocessing by denoising and background correction, the starting point for any cell-based assay is the identification of objects or cells by segmentation algorithms (Fig. 3 A). Segmentation algorithms detect objects based on a priori knowledge of their properties (like brightness, size, homogeneity, or edge information). Which object detection algorithm performs best highly depends on the specific fluorescent markers and cellular morphologies of a particular assay. Image segmentation is still one of the most challenging aspects when implementing an image analysis workflow for screening. To facilitate cell detection, fluorescent DNA or chromatin markers are often used because labeled cell nuclei have well-defined contours and they are well separated from neighboring cells (Megason and Fraser, 2007). Dilation of cell nuclei contours can serve to define cytoplasmic regions.

Because of the high experimental variability of most staining procedures, absolute fluorescence intensities for quantification of marker levels usually do not provide reproducible measures within large-scale experiments. Therefore, intensity-based assays mostly rely on fluorescence ratios related to a counterstained reference marker in individual cells. With this, imaging-based ratio metric measurements can be as sensitive as flow cytometry devices (Gordon et al., 2007).

To annotate distinct morphological phenotypes, pattern recognition methods have been developed, among which supervised machine learning classifiers are most common. These methods first require detailed quantitative description of shape and texture of each segmented cell. Several collections of algorithms
to calculate such statistical features have been published (Haralick et al., 1973; Prokop and Reeves, 1992; Walker and Jackway, 1996), which can be combined to define cellular morphologies by numerical vectors. In the next step of supervised machine learning, a classification algorithm needs to be trained on user-defined examples for cellular morphology classes (Fig. 3 B). A widely used classification algorithm, support vector machine, automatically determines a boundary (termed hyperplane) within the multidimensional feature space, which optimally discriminates the user-annotated cellular morphologies (Fig. 3 C). This classifier can be subsequently applied to annotate cell morphologies in large-scale image data (Fig. 3 D). In practice, supervised machine learning succeeded to accurately classify a variety of phenotypes typically in the range of 70–90%, including localization to subcellular structures, mitotic phenotypes, and virus infection (Boland and Murphy, 2001; Conrad et al., 2004; Neumann et al., 2006; Rämö et al., 2009). In fact, classifiers for subcellular objects have been shown to outperform human annotation for localization patterns of Golgi apparatus, endosomes, or lysosomes (Glory and Murphy, 2007). The orientation of the classification hyperplane and its distance to cellular objects in the feature space can also be directly used for detailed phenotypic profiling (Loo et al., 2007). In any case, a good representation of phenotypical variability in the training data is essential for accurate annotation of the full screening data.

Supervised machine learning can, in principle, also be used to annotate cellular dynamics in live imaging data. This requires a workflow combining morphology classification with the tracking of cells over time. However, because the maximal classification accuracy is rarely higher than 95% per object, cellular trajectories typically contain multiple annotation errors. Some efforts have been made to improve the classification accuracy in time series by suppressing biologically illegitimate transitions of morphological states. Although this has been shown to improve the annotation accuracy in an assay based on mitotic chromosome morphologies (Harder et al., 2009; Zhou et al., 2009), its application to other assays will require implementation of specific biological a priori models. To overcome this limitation, automatically extracted class transition probabilities can serve to correct classification errors without user supervision, providing a generic tool for time-resolved phenotype annotation (will be released as open source software CellCognition in 2010; unpublished data).

The most severe limitation of current machine learning approaches to RNAi screening is the dependence on a supervised training step. The definition of morphology classes is often based on a relatively small training dataset, which may be unrepresentative or lack rare phenotypic aberrations. Iterative supervised machine learning can be performed to raise the chance of discovering rare unknown phenotypes. Thereby, automated preselection of aberrant morphologies enables annotation of a more comprehensive classifier (Jones et al., 2009). Alternatively, iterative clustering of morphology classes can build a model that includes rare and unexpected phenotypes (Yin et al., 2008). Finally, unsupervised machine learning methods are available for clustering cellular morphologies into phenotype classes without any user interaction. Although these methods bear big promise, their performance in real-life biological screening applications still needs to be demonstrated.

**Quality control and validation of hits**

Optimizing an assay for screening requires a quantitative parameter for sensitivity and robustness of signal detection. A widely used measure for assay performance is the Z’ score, which defines the discriminative power of an assay between unperturbed negative controls and cells perturbed with a positive control that induces the respective phenotype (Zhang et al., 1999). Based on this quality score, the best experimental parameters can be determined, for example, the cell seeding density, transfection conditions, or image analysis parameters. Calculating Z’ scores is often based on a relatively small training dataset, which may be unrepresentative or lack rare phenotypic aberrations. Iterative supervised machine learning can be performed to raise the chance of discovering rare unknown phenotypes. Thereby, automated preselection of aberrant morphologies enables annotation of a more comprehensive classifier (Jones et al., 2009). Alternatively, iterative clustering of morphology classes can build a model that includes rare and unexpected phenotypes (Yin et al., 2008). Finally, unsupervised machine learning methods are available for clustering cellular morphologies into phenotype classes without any user interaction. Although these methods bear big promise, their performance in real-life biological screening applications still needs to be demonstrated.

**Open source software projects for image-based screening**

**μManager**

μManager (http://www.micro-manager.org) is software used to control automated image acquisition on motorized microscopes. Supported hardware devices include motorized microscope stands from various companies as well as a large number of illumination sources, shutters, filter wheels, scanning stages, and digital cameras. μManager provides a graphical user interface through integration into ImageJ software (National Institutes of Health).

**CellProfiler**

CellProfiler (http://www.cellprofiler.org) is software for data exploration and image analysis, designed for the needs of high-throughput imaging. It provides a graphical user interface to manage complex image analysis workflows and offers some functionality in feature extraction and machine learning (Sigal et al., 2006; Jones et al., 2008, 2009; Rämö et al., 2009). The software CellClassifier provides support vector machine classification to CellProfiler (Rämö et al., 2009).

**CellHTS**

The software CellHTS is part of a software project Bioconductor (http://www.bioconductor.org), which adds functionality to the statistical software R for the analysis of high-content screens. CellHTS can be used to compute the Z’ scores on the experiment-wide level of screens, offering various normalization graphical visualization options (Boutros et al., 2006). CellHTS also links to Gene Ontology and BioMart data warehouses to serve as a data mining tool.
straightforward, but it has been found to be less robust in more complex screens, for example, when the siRNA transfection efficiency is variable (Birmingham et al., 2009). Improved quality scores for RNAi screening assays can be calculated by taking into account pools of siRNA targeting one gene (König et al., 2007) or multiple experimental replicas of siRNA conditions combined with plate to plate variation (Zhang et al., 2008).

Two recent RNAi screens on cell cycle regulation in mammalian cells demonstrated surprisingly little overlap (~10%) of the hit lists (Mukherji et al., 2006; Kittler et al., 2007). This may in part reflect cell line–specific differences or the use of different RNAi reagents but raises concerns about a generally high rate of false hits in RNAi screens (Echeverri et al., 2006).

A potential source of false positive hits in RNAi screens are off-target effects of the RNAi reagents. The main reason for potential off-target effects in mammalian RNAi screens is the relatively high tolerance of mismatches on the siRNA (Birmingham et al., 2006). Several validation strategies are available. RNAi phenotypes should be reproduced by at least two distinct siRNAs. A correlation of the phenotype penetrance caused by different siRNAs and the depletion level provides support for the specificity of an RNAi phenotype. A powerful validation strategy tests for RNAi phenotype complementation by overexpression of an RNAi-resistant version of the target gene (Echeverri et al., 2006). A resource for such RNAi phenotype rescue experiments is mouse transgenes cloned from bacterial artificial chromosomes for stable expression in cultured
mammalian cells (Poser et al., 2008). Together, these studies highlight that hit lists resulting from primary RNAi screens bear limited value and that without detailed validation and follow up analysis, each “hit” should be considered not more than a candidate.

**Linking screening data to systems biology resources**

The ever-increasing content in image-based screening assays poses new challenges on the statistical analysis and data mining. Multiparametric phenotype descriptors, phenoprofiles, can be used to cluster genes in functional groups of similar RNAi phenotypes, which then can be related to potential common cellular processes, signaling pathways, or protein complex formation (Perlman et al., 2004; Skop et al., 2004; Pelkmans et al., 2005; Sönnichsen et al., 2005; Neumann et al., 2006; Loo et al., 2007). This requires sophisticated normalization and multivariate analysis methods to compensate potential experimental variations on the level of plates, wells, or cellular subpopulations (Text box 1; Boutros et al., 2004).

Comparison with multiple large-scale datasets, for example, subcellular protein localization, protein–protein interaction, and automated mining of published literature (Jensen et al., 2009), can be used to increase the relevance of a hit list obtained by RNAi screening. Furthermore, the comparison of different screens can be used for validation. This has been recently applied in a cell-based screen for antimalaria compounds, which were validated by comparison with >100 published unrelated cellular and enzymatic screens (Plouffe et al., 2008). Public access to original image data and their phenotypic annotations is necessary for comparing multiple large-scale datasets. This has been provided by RNAi screening databases for Caenorhabditis elegans (Gunsalus et al., 2004; Sönnichsen et al., 2005) and Drosophila (Flockhart et al., 2006), and similar resources for mammalian cells will be available in the near future (MitoCheck database; Neumann et al., 2010). However, the lack of standards for cell types, RNAi reagents, assay design, data analysis, and phenotype ontology still severely limits the integrated analysis of multiple large-scale datasets. Future efforts should address such standards, as it has been achieved, for example, in expression array profiling (Brazma et al., 2001).

**Conclusions**

The past few years have established microscopy as a high-content screening method for a broad spectrum of biological questions. Both microscope hardware and analysis software are now available to perform imaging-based screening in standard cell biology laboratories. Yet, advanced imaging modalities like time-lapse imaging require vast resources for screening on a genome-wide scale. The increasing number of large-scale systems biology data resources provides a basis to define meaningful candidate gene sets for hypothesis-based screening, for example, by focusing on genes of a certain functional class or intracellular compartment.

The design of new assays and workflows is still tedious, often limited by rigid software user interfaces and proprietary data formats. Improved modular software and databases will be needed to transform automated microscopy into a tool for hypothesis-driven research projects with daily changing assay needs. The integration of machine learning methods into microscope-controlling software will be an important next step for increasing assay content in screening. Based on classification of an image object or phenotypic event, the microscope could be automatically reconfigured, for example, to alter spatial or temporal resolution. This will also enable the implementation of new imaging modalities into screening applications like fluorescence correlation spectroscopy or photobleaching/photoactivation assays or provide event-driven feedback on robotic drug addition. Once the hurdle of hardware and software integration is overcome, imaging-based assays will open new options for systems biology. In the long run, we may envision completely autonomous “robot-scientists” (King et al., 2009) executing cycles of experimentation and hypotheses generation on automated microscopes.

We thank Dimitry Kofyar, Gabor Cucu, Wolfram H. Gerlich, Thomas Walter, Beate Neumann, and Jean-Karim Heriche for critical comments on the manuscript.

This work was supported by Swiss National Science Foundation research grant 3100A0-114120 and a European Young Investigator award of the European Science Foundation to D.W. Gerlich.

Submitted: 19 October 2009
Accepted: 25 January 2010

**References**

Aye-Han, N.N., Q. Ni, and J. Zhang. 2009. Fluorescent biosensors for real-time tracking of post-translational modification dynamics. *Curr. Opin. Chem. Biol.* 13:392–397. doi:10.1016/j.cpb.2009.07.009

Bakal, C., J. Aach, G. Church, and N. Perrimon. 2007. Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science.* 316:1753–1756. doi:10.1126/science.1140324

Bartz, F., L. Kern, D. Erz, M. Zhu, D. Gilbert, T. Meinhof, U. Wirkner, H. Erle, M. Muckenthaler, R. Pepperkok, and H. Runz. 2009. Identification of cholesterol-regulating genes by targeted RNAi screening. *Cell Metab.* 10:63–75. doi:10.1016/j.cmet.2009.05.009

Birmingham, A., E.M. Anderson, A. Reynolds, D. Islsey-Tyree, D. Leake, Y. Fedorov, S. Baskerville, E. Maksimova, K. Robinson, J. Karpilow, et al. 2006. 3’ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat. Methods.* 3:199–204. doi:10.1038/nmeth834

Birmingham, A., L.M. Selfors, T. Forster, D. Wrobel, C.J. Kennedy, E. Shanks, J. Santoyo-Lopez, D.J. Dunican, A. Long, D. Kelleher, et al. 2009. Statistical methods for analysis of high-throughput RNA interference screens. *Nat. Methods.* 6:569–575. doi:10.1038/nmeth.1351

Boland, M.V., and R.F. Murphy. 2001. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics.* 17:1213–1223. doi:10.1093/bioinformatics/17.12.1213

Boutros, M., A.A. Kiger, S. Armknecht, K. Kerr, M. Hild, B. Koch, S.A. Haas, R. Paro, and N. Perrimon. 2004. Genome-wide RNAi analysis of growth and viability in Drosophila cells. *Science.* 303:832–835. doi:10.1126/science.1091266

Boutros, M., L.P. Brás, and W. Huber. 2006. Analysis of cell-based RNAi screens. *Genome Biol.* 7:R66. doi:10.1186/gb-2006-7-7-r66

Brass, A.L., D.M. Dykshoorn, Y. Benita, N. Yan, A. Engelman, R.J. Xavier, J. Lieberman, and S.J. Elledge. 2008. Identification of host proteins required for HIV infection through a functional genomic screen. *Science.* 319:921–926. doi:10.1126/science.1155725

Braxma, A., P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, J. Aach, W. Ansorge, C.A. Ball, H.C. Causton, et al. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 29:365–371. doi:10.1038/ng1201-365

Buchholz, F., R. Kitterl, M. Slabicki, and M. Theis. 2006. Enzymatically prepared RNAi libraries. *Nat. Methods.* 3:696–700. doi:10.1038/nmeth912

---

*Image-based screening • Conrad and Gerlich* 459
assembly in Drosophila S2 cells. Science. 316:417–421. doi:10.1126/ science.1141314

Gunsalus, K.C., W.C. Yueh, P. MacMenamin, and F. Piano. 2004. RNAiDB and PhenoBlast: web tools for genome-wide phenotypic mapping projects. Nucleic Acids Res. 32:D406–D410. doi:10.1093/nar/gkh110

Haralick, R.M., K. Shanmugam, and I. Dinstein, 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. 3:610–621. doi:10.1109/TSMC.1973.4309314

Harder, N., F. Mora-Bermúdez, W.J. Godinez, A. Wünsche, R. Eils, J. Ellenberg, and K. Rohr. 2009. Automatic analysis of dividing cells in live cell mov-
ties to detect mitotic delays and correlate phenotypes in time. Genome Res. 19:2113–2124. doi:10.1101/gr.092494.109

Hartwell, L.H., J.J. Hopfield, S. Leibler, and A.W. Murray. 1999. From molecular to modular cell biology. Nature. 402:C47–C52. doi:10.1038/3511540

Haustein, E., and P. Schville. 2007. Fluorescence correlation spectroscopy: novel variations of an established technique. Annu. Rev. Biophys. Biomol. Struct. 36:151–169. doi:10.1146/annurev.biophys.36.043006.132612

Jackson, A.L., J. Burchard, J. Schelter, B.N. Chau, M. Cleary, L. Lim, and P.S. Linsley. 2006. Widespread siRNA “off-target” transcription silenced me-
diated by seed region sequence complementarity. RNA. 12:1179–1187. doi:10.1261/rna.25706

Jensen, L.J., M. Kuhn, M. Stark, S. Chaffron, C. Cvejic, J. Muller, T. Doerks, P. Julien, A. Roth, M. Simonovic, et al. 2009. STRING 8—a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 37:D12–D20. doi:10.1093/nar/gkn933

Jiang, D., L. Zhao, and D.E. Clapham. 2009. Genome-wide RNAi screen identi-
fies Letm1 as a mitochondrial Ca2+/H+ antiporter. Science. 326:144–147. doi:10.1126/science.1175145

Johnsson, K. 2009. Visualizing biochemical activities in living cells. Nat. Chem. Biol. 5:63–65. doi:10.1038/nchembio0209-63

Jones, T.R., I.H. Kang, D.B. Wheeler, R.A. Lindquist, A. Papallo, D.M. Sabatini, P. Gos, and A.E. Carpenter. 2008. CellProfiler: Data explorer: analysis and evaluation software for complex image-based screens. BMC Bioinformatics. 9:482. doi:10.1186/1471-2105-9-482

Jones, T.R., A.E. Carpenter, M.R. Lamprecht, J. Moffat, S.J. Silver, J.K. Grenier, A.B. Castooren, U.S. Eggert, D.E. Root, P. Golland, and D.M. Sabatini. 2009. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. Proc. Natl. Acad. Sci. USA. 106:1826–1831. doi:10.1073/pnas.0808843106

Kanda, T., K.F. Sullivan, and G.M. Wahl. 1998. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. Curr. Biol. 8:377–385. doi:10.1016/S0960-9822(98)00715-3

King, R.D., J. Rowland, S.G. Oliver, M. Young, W. Aubrey, E. Byrne, M. Liakata, M. Markham, P. Pir, L.N. Soldatova, et al. 2009. The automation of science. Science. 324:85–89. doi:10.1126/science.1165620

Kittler, R., L. Pelletier, A.K. Heninger, M. Slabicki, M. Simonovic, et al. 2009. STRING 8—a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 37:D12–D20. doi:10.1093/nar/gkn933

König, R., C.Y. Chiang, B.P. Tu, S.F. Yan, P.D. DeJesus, A. Romero, T. Bergauer, A. Orth, U. Krueger, Y. Zhou, and S.K. Chanda. 2007. A probability-based approach for the analysis of large-scale RNAi screens. Nat. Methods. 4:847–849. doi:10.1038/nmeth1089

Krishnan, M.N., A. Ng, B. Sukumaran, F.D. Gilfoy, P.D. Uchil, H. Sultana, A.L. Brass, R. Adametz, M. Tsui, F. Qian, et al. 2008. RNA interference and quantitative image data. Nat. Methods. 5:63–65. doi:10.1038/nmeth0208-63

Lippincott-Schwartz, J., N. Altan-Bonnet, and G.H. Patterson. 2003. Photobleaching and photoactivation: playing protein functions in living cells. Nat. Cell Biol. Suppl:S14–S19.

Liu, T., D. Sims, and B. Baum. 2009. Parallel RNAi screens across different cell lines identify genetic and cell type-specific regula-
tors of actin organization and cell morphology. Genome Biol. 10:R26. doi:10.1186/gb-2009-10-3-r26

Loo, L.H., L.F. Wu, and S.J. Altschuler. 2007. Image-based multivariate profiling of drug responses from single cells. Nat. Methods. 4:484–489. doi:10.1038/nmeth1089

Meister, G., and T. Tuschl. 2004. Mechanisms of gene silencing by double-
stranded RNA. Nature. 431:84–90. doi:10.1038/nature02760

Meister, G., and T. Tuschl. 2004. Mechanisms of gene silencing by double-
stranded RNA. Nature. 431:84–90. doi:10.1038/nature02760

Mukherji, M., R. Bell, L. Suppekova, Y. Wang, A.P. Orth, S. Batalov, L. Miraglia, D. Huesken, J. Lange, C. Martin, et al. 2006. Genome-wide functional
