CellPress

Review

High-Content Screening for Quantitative Cell Biology

Mojca Mattiazzi Usaj,[1] Erin B. Styles,[1] Adrian J. Verster,[1] Helena Friesen,[1] Charles Boone,[1] and Brenda J. Andrews[1,*]

High-content screening (HCS), which combines automated fluorescence microscopy with quantitative image analysis, allows the acquisition of unbiased multiparametric data at the single cell level. This approach has been used to address diverse biological questions and identify a plethora of quantitative phenotypes of varying complexity in numerous different model systems. Here, we describe some recent applications of HCS, ranging from the identification of genes required for specific biological processes to the characterization of genetic interactions. We review the steps involved in the design of useful biological assays and automated image analysis, and describe major challenges associated with each. Additionally, we highlight emerging technologies and future challenges, and discuss how the field of HCS might be enhanced in the future.

Introduction

Over the past 15 years, high-throughput (HTP) microscopy has been one of the fastest growing fields in cell biology. HTP microscopy is defined by its productivity, which usually means processing thousands to hundreds of thousands of samples, and includes automated sample preparation and data acquisition setup. When combined with automated multiparametric image and data analysis, HTP microscopy is referred to as HCS, a relatively new approach that has been transforming basic cell biology into a large-scale data-driven and exacting science.

Functional genomics strategies for exploring gene expression, genetic interactions, chemogenetic networks, or protein–protein interactions on a genome scale typically involve a single or handful of phenotypic readouts. By contrast, HCS provides the rich spatiotemporal resolution that is needed to fully understand the complexity and dynamics of cell biological processes. The power of HCS resides in its ability to simultaneously extract hundreds of biologically informative measurements from the acquired images, at the level of individual proteins, organelles, whole cells, cell populations, or even an entire organism.

The growth of HCS has been driven by advances in biology and chemistry, as well as mechanization and computation. One of the main drivers has been the development of genetic and molecular tools compatible with HTP microscopy, such as genome-scale mutant collections in budding and fission yeasts [1–4], and reagents for systematically altering gene expression in mammalian cells and metazoan model systems [5,6] (Box 1). In addition, a variety of fluorescent protein tags as well as fluorescent probes that report on numerous aspects of the state of the cell have become available (Box 1). Two other advances that have facilitated the rapid development of HTP microscopy and HCS platforms are hardware improvements that have catalyzed the development of fast, automated microscopes and major innovations in the software for image analysis, which have enabled facile extraction of quantitative measurements. Over the past decade, HTP microscopy has been used extensively to examine the effects of large-scale genetic or chemical perturbations, which impinge on different aspects of cellular organization and dynamics, at the

Trends

HCS combines automated microscopy with quantitative image analysis.

Recent hardware advances and innovations in software for automated image analysis now allow researchers to rapidly screen and analyze hundreds of thousands of images.

In contrast to early analysis of high-throughput imaging data, which often involved testing for deviation of a single parameter, machine learning, both supervised and unsupervised, allows high-dimensional data analysis.

The image analysis pipeline must be designed simultaneously with the development of the biological assay.

HCS has been used to identify genes and activities required for a specific biological process and in various disease models, to identify proteome-wide changes in response to chemical or genetic perturbations, and in chemical and genetic profiling.

[1]The Donnelly Centre, University of Toronto, Toronto, ON M5S3E1, Canada

*Correspondence: brenda.andrews@utoronto.ca (B.J. Andrews).

 http://dx.doi.org/10.1016/j.tcb.2016.03.008

Box 1. Genetic and Fluorescent Reagents for HTP Microscopy

A key factor in the recent applications of HTP microscopy is the availability of genetic and molecular reagents for fluorescence microscopy that are compatible with HTP approaches.

Genetic Reagents

The most versatile toolbox of genetic reagents is available for the model yeast *Saccharomyces cerevisiae* (e.g., [1,3,4]). For fluorescence microscopy, a particularly useful collection is the open reading frame (ORF)-GFP collection [19], in which each ORF has been systematically tagged at its C terminus with GFP. A second invaluable reagent set is the genome-wide gene-deletion collection, available in both *S. cerevisiae* [1,4] and *Schizosaccharomyces pombe* [2]. There are no complete mutant or tagged ORF collections available for multicellular organisms or cell lines, although partial collections exist for several organisms [57–59]. In most cases, these are not easily compatible with HTP applications and screening is conducted on a single gene and/or reporter basis.

Until now, the most widely used approach to screen for effects of reduced gene function in higher eukaryotes has been RNAi technology. There are several commercial RNAi libraries, from genome wide to more pathway focused, available for a range of organisms [5].

Recently, several technologies, such as gene editing, have emerged to ease the construction of genome-wide collections in higher organisms [60,61]. One of the most promising is CRISPR/Cas9 mediated (multi-) gene editing [62], which can be used for the large-scale construction of gene mutations or collections of tagged genes in mammalian and other eukaryotic cells. Genome-scale CRISPRi and CRISPRa libraries, which inhibit and activate gene expression respectively, offer the opportunity to modulate gene expression over an approximately 1000-fold range [6]. Likewise, CRISPR/Cas9-mediated knock-in of a desired reporter cassette [63] (e.g., a gene encoding a fluorescent protein fused to a gene of choice at the endogenous locus) in higher eukaryotes will enable *in vivo* imaging and reduce potential reagent toxicity and the need for preparative procedures associated with immunostaining.

Fluorescent Reagents

There are three main classes of fluorescent reagent that enable researchers to detect specific cellular structures: immunoreagents, fluorescent dyes and/or probes, and genetically encoded fluorescent proteins. Immunoreagents specifically recognize epitopes of the protein of interest and are widely used in screens of mammalian cell lines. The use of antibodies requires additional optimization of the fixation and permeabilization methods and testing for cross-reactivity when multiplexing different unconjugated antibodies. Fluorescent probes include dyes that bind to the cell surface or are taken up by cells and concentrate in specific cellular structures, as well as indicators of the physiological state of the cell, such as calcium indicators, live and/or dead indicators, and probes for pH, redox, and membrane potential. Genetically encoded fluorescent proteins, such as GFP and its derivatives, have revolutionized fluorescence microscopy. There has been a substantial push in recent years in the development of new fluorescent protein variants with different characteristics (various colors, fixable variants, and fluorescent timers) and protein-based fluorescent biosensors that monitor specific changes in the cellular state or the tagged protein and/or peptide (e.g., conformational changes, ligand binding, and post-translational modifications) [64,65].

single cell level in diverse model systems. The method has perhaps seen the greatest use in the pharmaceutical industry, where it has been used in the identification of new lead compounds that can be optimized into drug candidates, drug target validation, assessment of the *in vivo* toxicity of drug candidates, and the identification of the mode of action of orphan compounds [7].

Here, we provide an overview of applications of HTP microscopy and HCS approaches as applied in the basic research environment over the past decade, and highlight some of the key challenges that need to be addressed when developing projects within this experimental realm. The set of examples of HCS experiments outlined below is not meant to offer a comprehensive review, but rather to provide a sense of the breadth of biological questions that can be addressed using the approach.

Biological Questions Addressed with HTP Microscopy

The main goal of HCS experiments is to use images of cells to produce multidimensional profiles that identify aberrant phenotypes, such as those caused by the addition of a compound or genetic mutation, to make functional predictions. Below, we focus on protein imaging and describe five different classes of experiment that have successfully used HCS to answer important biological questions. To provide an overview of the power and potential of HCS approaches and to demonstrate their use in different model systems, we outline a few interesting

examples for each class of experiment and describe some of the limitations and benefits of each approach.

Identification of Genes and/or Activities Required for a Specific Biological Process

Typically the identification of genes and/or activities required for a specific biological process involves perturbing many (or all) genes in a standard genetic background expressing one or a few reporters of the biological process of interest. This type of approach encompasses the majority of screens in the literature, probably because it is relatively straightforward to set up, requiring only a perturbation collection and a robust phenotypic readout that is amenable to HTP and reproducibly reports on a biological compartment or function of interest. In some recent examples of this class of experiment, genes involved in cell division and the proper morphology of the mitotic spindle were identified in screens of both the fission and budding yeast deletion collections using a GFP-tubulin reporter gene [8,9]. The acquisition of multiple time points allows for a more detailed analysis of the biological consequences of the perturbation, an approach that is particularly useful when studying dynamic processes. For example, in another study, automated HTP time-lapse microscopy was used to monitor the effect of an RNAi library in a HeLa cell line expressing a GFP-tagged histone: 109 000 time-lapse movies were collected and the nucleus tracked, leading to the discovery of several novel genes involved in mitosis [10]. In addition, to obtain a more in-depth analysis of an entire process, screens for perturbations using multiple reporters, as opposed to single reporter screens, have been productive. For example, 13 RNAi screens of multiple different endocytic activities and their downstream organelles in HeLa cells identified genes that coregulate specific subsets of endocytic uptake routes and organelle abundances [11]. Other examples of this class are a screen for genes required for secretion in HeLa cells [12] and a screen for regulators of a signal transduction pathway using an immunofluorescence-based readout of activation of the human TORC1 effector RPS6 expressed in microarrays of *Drosophila* cells treated with RNAi [13].

Identification of Genes Required for Differentiation and Proliferation

Experiments to identify genes required for differentiation and proliferation also use perturbation libraries and a few reporters, but they must be done in specific cell types under conditions that induce differentiation, which often makes them technically challenging. Interesting examples of this type of screen include: a screen for stimulators of cardiomyocyte proliferation involving a library of miRNA mimics and markers of cell division [14]; a screen for regulators of human embryonic stem cell identity involving an RNAi library and a GFP reporter for quantitative analysis of the expression of the key pluripotency gene *POU5F1* [15]; and a screen for small molecules that regulate cell differentiation in zebrafish embryonic cells expressing lineage-specific GFP reporters [16]. In addition to the technical challenges, screens in specialized systems are subject to the caveat that their findings may not be relevant *in vivo*. This issue was addressed in the cardiomyocyte study: following myocardial infarction in mice, injection of vectors expressing two of these miRNAs was able to stimulate cardiac regeneration and almost complete recovery of cardiac functional parameters [14].

Models of Disease

HCS screens designed to explore disease mechanisms typically require the use of cells expressing disease-causing mutant alleles in relevant cell types coupled with reporters or assays tailored to query aspects of disease biology. For example, factors that may influence the pathology of Huntington disease were identified by screening *Drosophila* primary neurons expressing an RFP-tagged version of the pathogenic Huntingtin protein for small-molecule and genetic perturbations that revert aberrant neuronal morphology [17]. Other studies have used sensitive reporters activated by pathways known to contribute to disease pathology. For example, compounds that influence the cell biological defects associated with Alzheimer's disease were discovered by treating HEK293 cells expressing a disease-causing presenilin mutant and a genetically encoded FRET-based calcium indicator reporter [18]. Applying an analogous strategy, potential modifiers of

the cystic fibrosis phenotype were discovered by combining RNAi and HCS to assess the activity of the epithelial sodium channel ENaC in human alveolar epithelial cells [19]. As with other cell type-specific models, it is important to confirm any findings *in vivo*; researchers in the cystic fibrosis study were able to validate one of their hits as a possible drug target using cells from the lungs of patients with cystic fibrosis.

Identification of Proteome-Wide Changes in Response to Chemical or Genetic Perturbation

The identification of proteome-wide changes requires a collection of strains where each protein is individually tagged. Presently, an open reading frame (ORF)-GFP collection, in which each strain carries a unique fusion gene that generates an ORF-GFP fusion protein driven by the endogenous promoter, has been constructed only for *Saccharomyces cerevisiae* [20]. By combining automated yeast genetics and HCS of the ORF-GFP collection, changes were quantified in the abundance and localization of the yeast proteome in response to both genetic and chemical perturbation, with measurements recorded over a time series (Figure 1) [21]. The

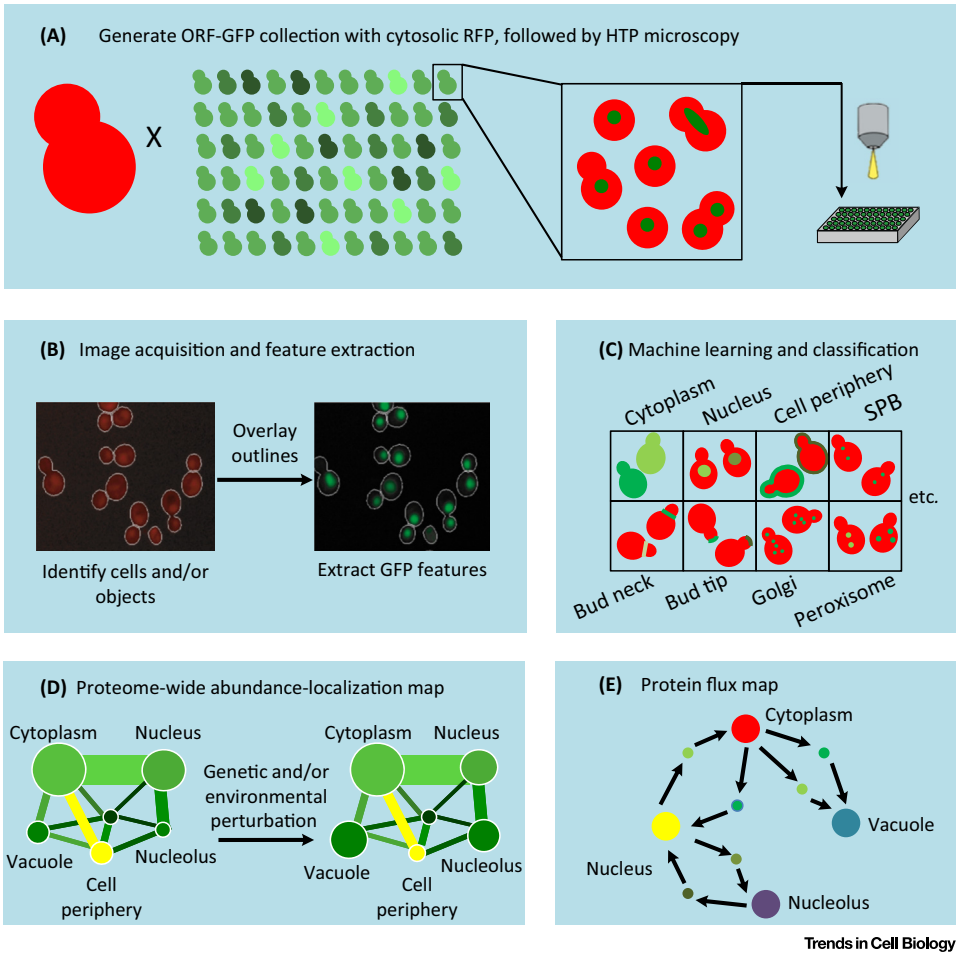


Figure 1. Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. (A) In these experiments a version of the yeast open reading frame (ORF)-GFP collection expressing a cytosolic RFP marker was created and the cells were imaged using automated confocal microscopy [21]. (B) The cytosolic red signal was used to segment the cells, and texture measurements were then extracted from both the red and green channels. (C) A representative subset of proteins that were known to localize to single subcellular compartments was chosen and used to generate training sets for automated classification of the remaining proteins. (D) After classifying all the ORF-GFPs, proteome-wide maps showing abundance and localization were generated. (E) Finally, image analyses of ORF-GFP collections treated with hydroxyurea and rapamycin and in a mutant genetic background were used to derive flux maps of the proteins that changed under these conditions.

challenge with this type of screen lies in identifying multiple different proteins that localize to a given compartment but that may not look identical; consequently, some type of machine learning is required. However, this type of approach offers the benefit of being able to query the entire proteome in an unbiased manner. In the near future, CRISPR-Cas9 technology will enable the creation of similar collections in other systems [22].

Chemical and Genetic Profiling

General phenotypic profiling can also be used to generate diagnostic profiles of the cellular response to various insults, providing a template for interpretation of the effects of unknown compounds or genetic lesions. For example, measurements of cell and nuclear shape and texture were used to generate profiles characteristic of the response of approximately 30 well-characterized drugs in HeLa and rat kidney cells over time and by dose. The comparison of profiles led to the identification of several natural products as tubulin inhibitors [23]. With this type of experiment, the possibility of similarity to an unknown compound is limited by the number of profiles in a set. In a second example, in budding yeast, quantification of 501 cell morphology parameters, based on staining of the cell wall, actin, and DNA, led to the identification of drugs that interfere with cell wall synthesis through comparison with profiles of mutant strains known to be involved in this process [24].

In addition to experiments that query a specific process, cell state, or chemical, imaging screens have also been used to ask more fundamental biological questions. For example, HCS has been used to identify genetic interactions that affect morphology rather than viability. In these experiments, the authors treated human cancer cells [25] and *Drosophila* cells [26] with double RNAi, stained the cells for DNA, actin, and tubulin, and extracted multiple quantitative features to identify genetic interactions. This example shows how HCS may be used to provide deeper insight into previously well-studied questions.

Assay Design

Regardless of the question being addressed, the success of any HTP-imaging experiment relies on thoughtful assay design and appropriate image analysis approaches. The main goal of HTP microscopy assay development is to design a phenotypic readout that is reproducible and scalable but that requires a minimum number of experimental steps. HTP microscopy experiments generally require at least four independent steps: sample preparation, image acquisition, image and/or data handling, and image analysis; each step comes with significant technical challenges that need to be considered beforehand. While each HTP microscopy experiment requires customization to address the biological question of interest, there are some general features associated with optimal assay design, which are summarized in Figure 2. These include the choice of a model system for which an appropriate phenotypic readout has been validated through low-throughput microscopy. The assay must then be adapted to a robust multi-well set-up, and scaled for automated sample handling (Figure 2).

The choice of imaging platform depends largely on the biological application. Although it is possible to carry out an HTP microscopy experiment on conventional motorized microscopes, most researchers opt for commercial automated imagers that are fully adapted to HTP operation, and offer different imaging modalities and flexibility (Table 1). Imaging systems are often coupled to automated liquid-handling devices to increase reproducibility and decrease plate-to-plate variation. Imagers fall into two broad categories: confocal and wide field. In general, confocal imagers offer higher image resolution and are better suited for imaging of small cells, intracellular or complex structures, and samples with strong background fluorescence, but are expensive and require longer imaging times, potentially limiting throughput. Wide-field imagers are faster and provide higher signal intensities when working with dimmer samples, but lack depth of resolution.

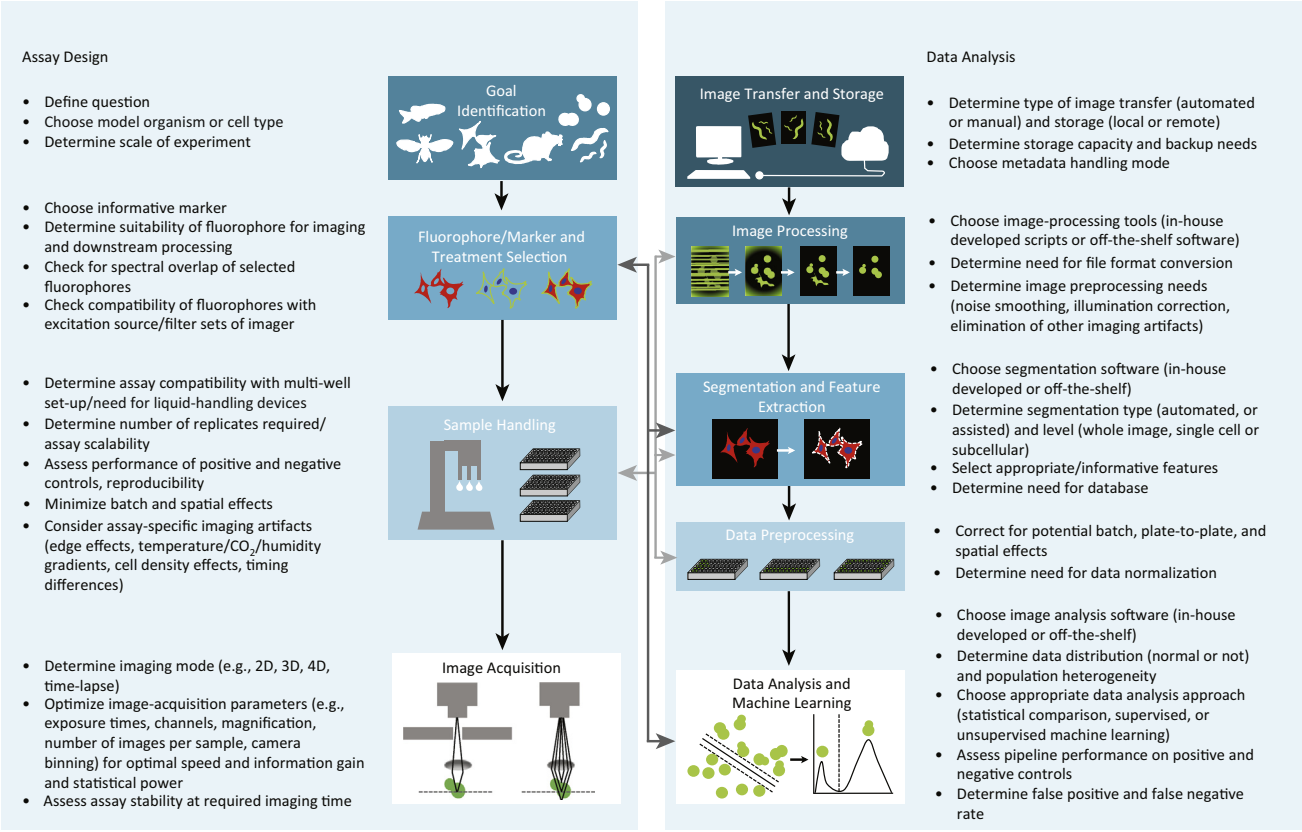


Figure 2. General Steps and Main Considerations of a High Content Screening (HCS) Experiment. Both the assay design and image analysis part of a high-throughput (HTP) microscopy experiment include several steps. Each step comes with specific challenges that need to be considered beforehand and optimized accordingly. It is important to develop the image-analysis pipeline in parallel with the development of the biological assay so that the markers and fluorophores, as well as the number of replicates, reproducibility, and experimental conditions, are appropriate for the types of image and data analyses chosen, and vice versa.

One of the greatest strengths of HTP microscopy is the ability to screen large sample sizes, which allows critical statistical considerations to be addressed, following careful consideration of the number of biological replicates and sites imaged per sample. When dealing with time-sensitive applications, optimal imaging parameters are often a tradeoff between the acquisition speed needed to achieve images of the desired quality and resolution, and the minimum sample numbers required for robust statistical analysis.

Table 1. Automated HTP Microscopy Platforms Frequently Seen in Published Literature[a]

Company	Model	Optics	Light Source	Software
BD Biosciences	BD Pathway 855/435	Confocal	Halide lamp	+
Perkin Elmer	Opera/Operetta	Confocal	Laser/halide lamp	+
GE Healthcare	IN Cell Analyzer 6000/2200	Confocal/wide field	Laser/solid state	+
Molecular Devices	ImageXpress Ultra/Micro	Confocal/wide field	Laser/halide lamp or solid state	+
Olympus	ScanR	Wide field	Halide lamp	+
Thermo Scientific	CellomicsArrayScan	Wide field (confocal optional)	Solid-state LED	+

[a]Most models come with optional kinetic and environmental control units.

Automated Image Analysis

A significant and sometimes underappreciated challenge of HTP microscopy screening lies in the computational analysis of the thousands of acquired images. Manual scoring of images, which dominated the early days of HTP screening, is still predominantly used in low- and medium-throughput studies, likely because most bench scientists lack expertise in image analysis and because significant time may be required to develop an image analysis pipeline. For true HTP experiments, in which thousands of images need to be assessed, manual scoring becomes unfeasible, but, more importantly, manual assignments of morphology or subcellular localization are time consuming and inherently subjective. Automation requires image-processing and data-analysis techniques, such as statistical analysis and machine learning. It leads to quantitative, highly reproducible analysis and, because of its assessment of multiple parameters (sometimes over 100), is highly sensitive, allowing for the identification of more subtle phenotypic differences than would be possible by manual analysis. The main steps of image and data-analysis pipelines include: image processing, object segmentation, feature extraction and selection, and statistical analysis. Combined, these steps ultimately lead to the transformation of bioimaging data into biologically meaningful results. Figure 2 highlights the general workflow and main considerations associated with each step, which we discuss further below. In general, it is crucial to develop the image-analysis pipeline simultaneously with the development of the biological assay and ensure that the analysis steps used lead to reproducibility from batch to batch and corroborate known effects of included positive and negative controls.

Image Processing, Object Segmentation, and Feature Extraction

After automated image acquisition, images first need to be corrected for noise (e.g., background noise or experimental artifacts) and uneven illumination. This is usually achieved by applying different filters, either linear or nonlinear, to smooth away the noise [27]. Next, regions of interest (ROIs) must be defined (e.g., cells or subcellular structures) in each microscopy image through the process of segmentation. This is most commonly done using a binarization or thresholding algorithm, such as Otsu's Method, combined with the watershed algorithm [28,29]. While a good segmentation pipeline can go a long way towards separating ROIs from background pixels and from each other, good experimental design can save both time and stress by minimizing imaging artifacts that can introduce bias. For example, if segmentation is performed using only a cell boundary marker, a thresholding algorithm is likely to accidentally merge adjacent cells (an analysis issue known as 'undersegmentation'), whereas, if a nuclear marker is also present, nuclei can be detected first using thresholding, and a watershed algorithm can then be used to propagate outwards from the nuclei, to properly segment cells.

Following cell segmentation, quantitative measurements that are representative of the cells or subcellular objects can be acquired. A variety of commonly used numeric descriptors of segmented objects have been developed in the computer vision field, and include summaries of the intensity distribution, shape, size, texture, radial distribution, and granularity of the segmented ROIs. Several measurements tailored to the problem of interest can additionally be calculated from the extracted quantitative features, such as the nucleus:cell area ratio. Importantly, not all extracted features are equally informative for a given biological question, and many features can be highly redundant and represent almost identical information. Furthermore, in certain cases, an overabundance of features can inhibit algorithm performance and introduce noise in the data.

To simplify the data without losing information, dimensionality reduction can be used to project the original feature space onto a lower dimensional space. A commonly applied method is principal component analysis (PCA) [30], where the original features are linearly transformed into new sets of features. The transformed features are often difficult to interpret; therefore, feature selection methods can be used to select, by process of iterative elimination, a smaller subset of

features that are most informative for a given question. A review of the different methods can be found in [31].

Data Analysis: Statistical Comparisons

After feature extraction and selection, three general data analysis approaches can be carried out using the quantitative morphometric features: statistical comparisons, supervised machine learning, and unsupervised machine learning.

Statistical comparisons directly operate on feature distributions. Here, the most common action is to determine whether two sets of cell populations, strains, or images (usually comparing a treated to a nontreated or mutant to wild-type sample) are significantly different. HCS can collect hundreds of feature measurements for each ROI, but researchers often use only one or a few image-derived features and calculated statistical metrics for downstream analysis. Despite the utility of direct comparisons of one or a few selected features using statistics such as the Z factor [32], an analysis pipeline that uses only a single measurement severely restricts the detection power of a data set. In addition to the information loss that occurs in single-metric analyses, important cellular information can also be lost when statistically oversimplified data-mining strategies are used, such as arbitrary Z-score cut-offs to identify outliers in a population that might not follow a single distribution curve. Many of the measurements that are ignored are likely to correspond to rare phenotypes of interest (see [33] for detailed review).

HTP microscopy data are inherently information rich, but the utility of this information depends on appropriate analysis of the multiparametric data. Moving away from testing the deviation of a single-parameter from a control population, some groups have explored the full range of available parameters, assessing each feature in relation to a control population to create high-dimensional data analysis methods. For example, in one of the earliest screens, the nonparametric Kolmogorov-Smirnov [34] statistic was used to broadly profile 100 chemicals over a range of concentrations, analyzing 11 different markers and a large set of measurements, to assess multidimensional single cell phenotypes [35]. This led to the successful categorization of blinded drugs and suggested targets for drugs of uncertain mechanism. In this screen, all of the parameters were considered independently; other studies have tried to assess the dependencies between parameters using methods such as factor analysis [36]. Such approaches increase the ability to discover novel phenotypes, but often suffer from what is known as the 'multiple comparisons problem', occurring when a large set of statistical inferences are considered simultaneously, leading to errors in inference [37]. Unfortunately, methods that are often used to adjust for this issue, such as the Bonferroni correction [38,39], will increase the false negative rate [40]. Furthermore, while some measurements are easy to understand (e.g., an increase in nuclear size), others lack a direct biological interpretation (e.g., an increase in the texture entropy in the diagonal direction).

As part of any statistical analysis, it is worthwhile for the researcher to consider whether to minimize false positive or false negative results. For example, if extensive follow-up of the findings is planned, such as might be the case for a disease treatment, the researcher would want to have a low false positive rate. Conversely, if general biological principles are being explored, the researcher may prefer to include more of the data and reduce the false negative rate.

Data Analysis: Machine Learning

An effective way to turn a large number of measurements per cell into something biologically interpretable involves machine learning. In supervised machine learning, algorithms such as logistic regression, random forests, and support vector machines automatically discover novel positive items from training data (reviewed in [41]). In a classification problem, a trained biologist first manually labels examples of ROIs (for example individual cells) or whole images exhibiting a

phenotype of interest and groups them together in distinct classes; additional occurrences of the same phenotype within the full data set are subsequently identified by the trained classifier [42]. This approach has been successfully used by several groups (e.g., [10,21], Figure 1). Classification accuracy can also be implemented as a measure for identifying when a treatment condition is significantly different from the control, an approach that has been used to quantitatively assess drug titrations in human cells [43].

While supervised machine-learning approaches have been used with great success to identify novel instances of known phenotypes, these methods rely on the *a priori* knowledge of aberrant phenotypes or cellular patterns. When phenotypic classes are uncertain or unknown, unsupervised machine learning can be used to identify distinct groups within the data set. One common strategy for unsupervised machine learning is hierarchical data clustering, in which similar data points (either at the single cell or image and/or population level) are grouped together. Clustering approaches have been used to group proteins by their subcellular patterns [44] and drugs by their effects [35].

Supervised and unsupervised machine-learning methods can be combined, to harness the best from both worlds. A supervised neural network trained to identify representative phenotypes was used to analyze the cell morphologies in a *Drosophila* cell line after gene overexpression or double-stranded RNA treatment, and genes were subsequently grouped into informative biological pathways using unsupervised hierarchical clustering [45]. Currently, there exists a large amount of interest in neural network-derived algorithms, both supervised and unsupervised (e.g., autoencoders, which learn feature representations from the image pixels rather than from predefined formulas), although to date few groups have applied such methods to the analysis of HCS data. Instead, many of the existing applications of neural networks have been applied to precomputed image features rather than to the underlying pixels [46].

Population Heterogeneity

Depending on the segmentation parameters, data extracted from microscopy images usually represent the phenotype of a single cell. However, these data can be reported as an averaged summary of the image. When working at the whole-population and/or full-image level, researchers run the risk of neglecting cellular heterogeneity within each treatment population and missing changes that only occur in a subpopulation of cells. Particularly in the case where significant population variability is expected, averaged feature representations will not be representative of the actual biological state within the population. To avoid data misinterpretation, an alternative is to classify or cluster individual cells to identify such cellular subpopulations. For example, by clustering cells based on several morphological measurements, discrete subpopulations of cancer cells could be identified, as well as the response of each subpopulation to different classes of cancer drug [47]. To show the direct effects of cell–cell variability on HCS results, the rate at which different human cell lines were infected by a panel of viruses was assessed [48], revealing that 60–80% of the variability was attributable to the local cell microenvironment, which was represented by variables such as cell size, local cell density, and the mitotic state. The same group later showed that accounting for the local cell microenvironment can significantly improve the reproducibility of cell line RNAi screens [49]. This result has significant implications for phenotypic RNAi screening, because many RNAi screens have poor reproducibility [50].

Concluding Remarks

The steady increase in reports of HCS in the literature in recent years indicates that developing and performing these types of massively multiplexed assays is becoming more accessible to academic laboratories. However, while there have been many recent advances in the development of useful reagent sets, HTP imaging techniques, and multiparametric data mining of high content microscopy data, many aspects of HCS are still in their infancy (see Boxes 1 and 2, and

Outstanding Questions

How do we integrate HCS data with other large-scale data sets to get a complete picture of the cell?

Can CRISPR/Cas9-mediated gene editing be used for large-scale construction of gene mutations and knock-in tags?

What type of coherent strategy for data presentation and accessibility can we generate to unify reporting standards and make data from different sources directly comparable?

How can HCS be combined with other types of imaging technology to derive HTP quantitative data from unexplored sources?

Box 2. Open-Source Tools Available for Bioinformatics for Microscopy

Bioinformatics for microscopy aims to provide tools for data acquisition, storage, management, analysis, and visualization and has had a major impact on HTP cell biology, where advanced image-analysis approaches are needed. Given the variety of applications, it is impossible to create a universal analysis solution. While almost all commercial image-acquisition platforms now come with easy-to-use built-in software that offers at least some basic image visualization, data extraction, and analysis capability, a particular HCS assay often requires a combination of approaches or a completely new methodology to achieve optimal analysis. Several academic laboratories have dedicated research personnel or even core facilities to the development of the tools required for every aspect of image analysis, from image transfer and storage, to data extraction, management, and visualization. An alternative for many labs is the use of several existing user-friendly software solutions, both commercial and open source. Many open-source software platforms were originally developed to solve specific image-analysis problems, but were later expanded to other purposes by the research community. They provide a cost-effective option, can be readily customized to individual needs, and should ultimately encourage more researchers to incorporate automated image analysis in their studies. These open-source tools are available for all the steps involved in carrying out a HTP microscopy screen. Many allow direct linking of data to other open-source software solutions. Examples of open source software available for each of the image analysis steps are listed in Table I. The list is by no means exhaustive.

Table I. Examples of Open-Source Software Available for Bioimage Informatics

Step	Software	Features	Refs	Website
Image acquisition	μmanager	Software for full automation of motorized microscopes	[66]	www.micro-manager.org/
	scanImage	Automation of laser-scanning microscopes	[67]	http://svobodalab.cshl.edu/software_main.html
Image storage and management	OMERO	Enables web-based access, sharing, analysis, and visualization of complex, multidimensional image data and associated metadata	[54]	www.openmicroscopy.org/site
	BisQue	Enables storage, visualization, organization, and analysis of images in the cloud	[68]	http://bioimage.ucsb.edu/bisque
	OpenBIS	Enables web-based querying and visualization of both raw and analyzed data from HCS experiments	[69]	www.cisd.ethz.ch/software/openBIS
Image format conversion	Bio-formats	A Java library that converts >120 proprietary file formats to a common model	[53]	www.openmicroscopy.org/site
Image processing and analysis	BioImageXD	Package for analyzing, processing, and visualizing multidimensional microscopy images	[70]	www.bioimagexd.net/
	ImageJ; Fiji	General purpose image processing and analysis software package	[71,72]	http://imagej.nih.gov/ij/; http://fiji.sc/Fiji
	EBImage	General purpose toolbox for image processing and analysis	[73]	http://bioconductor.org/packages/release/bioc/html/EBImage.html
	CellCognition	Combines object detection and supervised machine learning for classification of morphologies with time-resolved analysis by single cell tracking	[74]	http://www.cellcognition.org/
	CellProfiler	Modular design enables the combining of predefined algorithms into an image analysis pipeline	[75]	www.cellprofiler.org/
Machine learning	CellProfiler Analyst	Exploration and analysis of large, high-dimensional image-derived	[76]	www.cellprofiler.org/

Table I. (continued)

Step	Software	Features	Refs	Website
		data; includes machine-learning tools for identifying complex and subtle phenotypes		
	Ilastik	Modular software framework for (supervised) pixel- and object-level classification, automated and semiautomated object tracking, semiautomated segmentation, and object counting	[77]	http://ilastik.org/index.html
Workflow systems ('visual programming')	KNIME	Enables nonexperts to visually build an image analysis workflow	[78]	www.knime.org/
	Orange	Data mining through visual programming: data visualization and analysis, components for machine learning, add-ons for bioinformatics and text mining, features for data analytics	[79]	http://orange.biolab.si/
	Taverna	Domain-independent workflow management system	[80]	www.taverna.org.uk/
Data mining	WEKA	Collection of machine-learning algorithms for data mining tasks	[81]	www.cs.waikato.ac.nz/ml/weka/

the *Automated Image Analysis* section) and several questions remain (see Outstanding Questions).

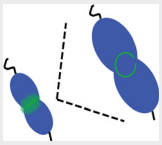
A major shift in assay design that is needed to understand the full complexity and dynamics of cellular processes is increased use of HTP live cell and time-lapse microscopy. While live cell imaging is commonly performed in HCS screens with unicellular organisms, such as yeast, or small multicellular organisms, such as worms, fixed cells are still the norm when working with mammalian cell cultures. With a few exceptions (e.g., [10,11]), producing and analyzing single frame and single time point images is the modus operandi of most reported studies. By collecting high-dimensional images, comprising multiple time points or using 3D or spatially resolved time-lapse microscopy (4D imaging), a large amount of detailed kinetic information can be extracted.

The development of novel gene-editing approaches, in particular CRISPR/Cas9-mediated gene editing, will make HTP genetics and HCS more amenable to many different eukaryotic systems (Box 1). CRISPR/Cas9 is likely to have an impact in two areas: first, in the construction of new types of perturbation reagents and libraries [6] and second, in the building of new mutated or tagged [51] cell lines. HCS is also developing towards screening the effects of genetic or environmental perturbations in more physiologically relevant cellular systems, such as cells in 3D cultures or cell co-cultures, where many aspects of tissue physiology have been preserved [52]. With increasing screening complexity, even more effort needs to be put into the development of suitable analysis methods, such as computer vision methods for the identification of complex cell shapes or temporal tracking of small objects with low signal:noise ratios.

HTP 'omics' technologies generate large amounts of data on the DNA, RNA, protein, lipid, or metabolite content within a cell. HCS with its multiparametric readout and potentially millions of data points generates even larger amounts of data: hundreds of quantitative measurements per image, per cell, or per subcellular structure describing multiple markers and/or reporters, in addition to thousands of images each associated with metadata. This wealth of data poses

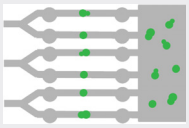
Box 3. Emerging and Hybrid Imaging Technologies

Fluorescence microscopy comprises a vast and expanding selection of image acquisition approaches. Several new and hybrid technologies have emerged in recent years that advance the possibilities of HTP microscopy even further. Here we outline notable examples that have been adapted to HTP or that we predict will be adapted in the near future.

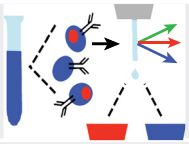


Super-resolution microscopy: This technique allows live-cell imaging with up to 50x better resolution than standard fluorescence microscopy and is just beginning to be adapted to HTP (reviewed in [82]). Recent work used HTP 3D super-resolution photo-activated localization microscopy (PALM) to image synchronized *Caulobacter* bacteria [83], providing a detailed quantitative picture of the FtsZ protein in cell division at nanoscale resolution.

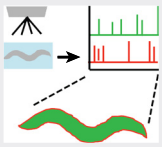
Fluorescence correlation spectroscopy: Using this technique researchers can measure properties of proteins that reflect the fluctuation of fluorescence intensity over short periods of time, allowing them to derive information such as diffusion coefficients and interaction kinetics (reviewed in [84]). Recent experiments described the development of HTP fluorescence correlation spectroscopy to study not just dynamic abundance and localization, but also mobility and protein–protein interactions for 53 nuclear proteins in human cells [85].



Microfluidics: Developments in microfluidics have permitted researchers to perform live-cell imaging in a variety of conditions over extended periods of time (reviewed in [86]). A recent example is the development of an HTP microfluidic single-cell analysis chip, enabling efficient trapping and retention of yeast mother cells, followed by time-lapse microscopy to track single cells throughout their entire replicative lifespan [87].



Imaging flow cytometry: This technique offers the speed and throughput of flow cytometry combined with the resolution of fluorescence microscopy (reviewed in [88]). HTP imaging flow cytometry was recently used to distinguish among B cells at various stages of the cell cycle and ask questions about subcellular protein distribution from their images [89].



Imaging mass spectrometry: In this technique the mass spectrometer collects molecular snapshots of a biological sample and this molecular information is superimposed onto an optical or fluorescence image, giving a picture of the distribution of specific molecules in a tissue or population of cells (reviewed in [90]). While we know of no reports on HTP imaging mass spectrometry projects, an automated imaging mass spectrometer has been developed [91] and this technique is likely to be useful in HTP studies.

specific challenges, some of which have been addressed above. Microscopy is a diverse field, with more imaging techniques emerging and being adapted to HTP (Box 3), as well as improved instrumentation, experimental design, and analysis approaches. These developments bring major challenges, such as the management of terabytes of images and image-associated data, the diversity of data structures, and the often proprietary file formats for data storage that come with commercial imaging systems.

One of the crucial future challenges will be the development of a coherent strategy for data presentation and accessibility. Open Microscopy Environment (OME) offers two possible solutions to these problems, with the development of Bio-Formats [53] and the OMERO server [54] (Box 2). An important consideration with the latter is the trade-off between the flexibility offered by the platform and available resources, both in terms of money and training time required for expert use of the server, which might pose a considerable limitation to the widespread use of OMERO. Since HCS is an evolving field, there are at present no unified reporting standards or requirements for the full availability of data through centralized image and/or data repositories, similar to the standards implemented in the field of microarray technology with the Minimum Information About a Microarray Experiment (MIAME) requirements [55]. To date, there is no systematic reporting standard proposed for HCS through the MIBBI Foundry [56], which hosts a common portal of minimum information checklists for various disciplines. Establishment of such

standards for recording and reporting HCS data would facilitate data exchange, the interpretation and independent verification of the reported data, and the development of much-needed downstream informatics tools, centralized repositories, and image databases.

Microscopy has always been the method of choice for cell biologists. With recent advances in automated microscopy platforms and computational image-analysis techniques, the information derived from microscopy images on a large scale can now be objectively assessed to uncover even subtle phenotypic changes after genetic, chemical, or other perturbations. Despite the challenges associated with HTP microscopy, we can expect this technique to be instrumental in answering diverse biological questions.

References

1. Giaever, G. *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391
2. Kim, D.U. *et al.* (2010) Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe. *Nat. Biotechnol.* 28, 617–623
3. Li, Z. *et al.* (2011) Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat. Biotechnol.* 29, 361–367
4. Winzeler, E.A. *et al.* (1999) Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science* 285, 901–906
5. Falschlehner, C. *et al.* (2010) High-throughput RNAi screening to dissect cellular pathways: a how-to guide. *Biotechnol. J.* 5, 368–376
6. Gilbert, L.A. *et al.* (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 159, 647–661
7. Zanella, F. *et al.* (2010) High content screening: seeing is believing. *Trends Biotechnol.* 28, 237–245
8. Graml, V. *et al.* (2014) A genomic multiprocess survey of machineries that control and link cell shape, microtubule organization, and cell-cycle progression. *Dev. Cell* 31, 227–239
9. Vizeacoumar, F.J. *et al.* (2010) Integrating high-throughput genetic interaction mapping and high-content screening to explore yeast spindle morphogenesis. *J. Cell Biol.* 188, 69–81
10. Neumann, B. *et al.* (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* 464, 721–727
11. Liberali, P. *et al.* (2014) A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell* 157, 1473–1487
12. Simpson, J.C. *et al.* (2012) Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nat. Cell Biol.* 14, 764–774
13. Lindquist, R.A. *et al.* (2011) Genome-scale RNAi on living-cell microarrays identifies novel regulators of Drosophila melanogaster TORC1-S6K pathway signaling. *Genome Res.* 21, 433–446
14. Eulalio, A. *et al.* (2012) Functional screening identifies miRNAs inducing cardiac regeneration. *Nature* 492, 376–381
15. Chia, N.Y. *et al.* (2010) A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* 468, 316–320
16. Huang, H. *et al.* (2012) High-throughput screening for bioactive molecules using primary cell culture of transgenic zebrafish embryos. *Cell Rep.* 2, 695–704
17. Schulte, J. *et al.* (2011) High-content chemical and RNAi screens for suppressors of neurotoxicity in a Huntington's disease model. *PLoS ONE* 6, e23841
18. Honarnejad, K. *et al.* (2013) Development and implementation of a high-throughput compound screening assay for targeting disrupted ER calcium homeostasis in Alzheimer's disease. *PLoS ONE* 8, e80645
19. Almaca, J. *et al.* (2013) High-content siRNA screen reveals global ENaC regulators and potential cystic fibrosis therapy targets. *Cell* 154, 1390–1400
20. Huh, W.K. *et al.* (2003) Global analysis of protein localization in budding yeast. *Nature* 425, 686–691
21. Chong, Y.T. *et al.* (2015) Yeast proteome dynamics from single cell imaging and automated analysis. *Cell* 161, 1413–1424
22. Sternberg, S.H. and Doudna, J.A. (2015) Expanding the biologist's toolkit with CRISPR-Cas9. *Mol. Cell* 58, 568–574
23. Futamura, Y. *et al.* (2012) Morphobase, an encyclopedic cell morphology database, and its use for drug target identification. *Chem. Biol.* 19, 1620–1630
24. Okada, H. *et al.* (2014) Distinct roles of cell wall biogenesis in yeast morphogenesis as revealed by multivariate analysis of high-dimensional morphometric data. *Mol. Biol. Cell* 25, 222–233
25. Laufer, C. *et al.* (2013) Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nat. Methods* 10, 427–431
26. Fischer, B. *et al.* (2015) A map of directional genetic interactions in a metazoan cell. *Elife* 4, e05464
27. Wang, Y.L. (2007) Computational restoration of fluorescence images: noise reduction, deconvolution, and pattern recognition. *Methods Cell Biol.* 81, 435–445
28. Beucher, S. (1991) The watershed transformation applied to image segmentation. *Scan. Microsc. Int.* 6 (Suppl.), 299–314
29. Otsu, N. (1979) Threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man. Cybern. SMC* 9, 62–66
30. Abdi, H. and Williams, L.J. (2010) Principal component analysis. *Wiley Interdiscip. Rev.: Comput. Stat.* 2, 433–459
31. Guyon, I. and Elisseeff, A. (2003) An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3, 1157–1182
32. Zhang, J.H. *et al.* (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen* 4, 67–73
33. Singh, S. *et al.* (2014) Increasing the content of high-content screening: an overview. *J. Biomol. Screen* 19, 640–650
34. Cox, D.R. and Hinkley, D.V. (1979) *Theoretical Statistics,* CRC Press
35. Perlman, Z.E. *et al.* (2004) Multidimensional drug profiling by automated microscopy. *Science* 306, 1194–1198
36. Young, D.W. *et al.* (2008) Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nat. Chem. Biol.* 4, 59–68
37. Miller, R.G. (2012) *Simultaneous statistical inference,* Springer
38. Dunn, O.J. (1959) Estimation of the medians for dependent variables. *Ann. Math. Stat.* 30, 192–197
39. Dunn, O.J. (1961) Multiple comparisons among means. *J. Am. Stat. Assoc.* 56, 52–64
40. Perneger, T.V. (1998) What's wrong with Bonferroni adjustments. *Br. Med. J.* 316, 1236–1238
41. Tarca, A.L. *et al.* (2007) Machine learning and its applications to biology. *PLoS Comput. Biol.* 3, e116
42. Jones, T.R. *et al.* (2009) Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1826–1831
43. Loo, L.H. *et al.* (2007) Image-based multivariate profiling of drug responses from single cells. *Nat. Methods* 4, 445–453

44. Chen, X. and Murphy, R.F. (2005) Objective clustering of proteins based on subcellular location patterns. *J. Biomed. Biotechnol.* 2005, 87–95
45. Bakal, C. *et al.* (2007) Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science* 316, 1753–1756
46. Boland, M.V. and Murphy, R.F. (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
47. Slack, M.D. *et al.* (2008) Characterizing heterogeneous cellular responses to perturbations. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19306–19311
48. Snijder, B. *et al.* (2009) Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* 461, 520–523
49. Snijder, B. *et al.* (2012) Single-cell analysis of population context advances RNAi screening at multiple levels. *Mol. Syst. Biol.* 8, 579
50. Bushman, F.D. *et al.* (2009) Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog.* 5, e1000437
51. Lackner, D.H. *et al.* (2015) A generic strategy for CRISPR-Cas9-mediated gene tagging. *Nat. Commun.* 6, 10237
52. Pampaloni, F. *et al.* (2007) The third dimension bridges the gap between cell culture and live tissue. *Nat. Rev. Mol. Cell Biol.* 8, 839–845
53. Linkert, M. *et al.* (2010) Metadata matters: access to image data in the real world. *J. Cell Biol.* 189, 777–782
54. Allan, C. *et al.* (2012) OMERO: flexible, model-driven data management for experimental biology. *Nat. Methods* 9, 245–253
55. Brazma, A. *et al.* (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 29, 365–371
56. Taylor, C.F. *et al.* (2008) Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat. Biotechnol.* 26, 889–896
57. Kettleborough, R.N. *et al.* (2013) A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* 496, 494–497
58. Mohr, S.E. *et al.* (2014) Resources for functional genomics studies in *Drosophila melanogaster*. *Genetics* 197, 1–18
59. Thompson, O. *et al.* (2013) The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* 23, 1749–1762
60. Durai, S. *et al.* (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* 33, 5978–5990
61. Miller, J.C. *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148
62. Shalem, O. *et al.* (2015) High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* 16, 299–311
63. Yang, H. *et al.* (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379
64. Giepmans, B.N. *et al.* (2006) The fluorescent toolbox for assessing protein location and function. *Science* 312, 217–224
65. Stepanenko, O.V. *et al.* (2011) Modern fluorescent proteins: from chromophore formation to novel intracellular applications. *Biotechniques* 51, 313–327
66. Edelstein, A. *et al.* (2010) Computer control of microscopes using microManager. *Curr. Protoc. Mol. Biol.* Chapter 14, Unit14 20
67. Pologruto, T.A. *et al.* (2003) ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* 2, 13
68. Kvilekval, K. *et al.* (2010) Bisque: a platform for bioimage analysis and management. *Bioinformatics* 26, 544–552
69. Bauch, A. *et al.* (2011) openBIS: a flexible framework for managing and analyzing complex data in biology research. *BMC Bioinform.* 12, 468
70. Kankaanpaa, P. *et al.* (2012) BioImageXD: an open, general-purpose and high-throughput image-processing platform. *Nat. Methods* 9, 683–689
71. Abramoff, M.D. *et al.* (2004) Image processing with image. *J. Biophotonics Int.* 11, 36–42
72. Schindelin, J. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682
73. Pau, G. *et al.* (2010) EBImage–an R package for image processing with applications to cellular phenotypes. *Bioinformatics* 26, 979–981
74. Held, M. *et al.* (2010) CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nat. Methods* 7, 747–754
75. Carpenter, A.E. *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 7, R100
76. Jones, T.R. *et al.* (2008) CellProfiler Analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinform.* 9, 482
77. Sommer, C. *et al.* (2011) Ilastik: interactive learning and segmentation toolkit. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2011*, pp. 230–233, IEEE
78. Berthold, M.R. *et al.* (2008) KNIME: The Konstanz Information Miner. In *In Data Analysis, Machine Learning and Applications* (Preisach, C. *et al.*, eds), pp. 319–326, Springer
79. Demšar, J. *et al.* (2013) Orange: data mining toolbox in python. *J. Mach. Learn. Res.* 14, 2349–2353
80. Wolstencroft, K. *et al.* (2013) The Taverna workflow suite: designing and executing workflows of Web Services on the desktop, web or in the cloud. *Nucleic Acids Res.* 41, W557–W561
81. Frank, E. *et al.* (2004) Data mining in bioinformatics using Weka. *Bioinformatics* 20, 2479–2481
82. Yamanaka, M. *et al.* (2014) Introduction to super-resolution microscopy. *Microscopy* 63, 177–192
83. Holden, S.J. *et al.* (2014) High throughput 3D super-resolution microscopy reveals *Caulobacter crescentus* in vivo Z-ring organization. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4566–4571
84. Liu, Z. *et al.* (2015) Imaging live-cell dynamics and structure at the single-molecule level. *Mol. Cell* 58, 644–659
85. Wachsmuth, M. *et al.* (2015) High-throughput fluorescence correlation spectroscopy enables analysis of proteome dynamics in living cells. *Nat. Biotechnol.* 33, 384–389
86. Streets, A.M. and Huang, Y. (2014) Microfluidics for biological measurements with single-molecule resolution. *Curr. Opin. Biotechnol.* 25, 69–77
87. Jo, M.C. *et al.* (2015) High-throughput analysis of yeast replicative aging using a microfluidic system. *Proc. Natl. Acad. Sci. U. S. A.* 112, 9364–9369
88. Barteneva, N.S. *et al.* (2012) Imaging flow cytometry: coping with heterogeneity in biological systems. *J. Histochem. Cytochem.* 60, 723–733
89. Thaunat, O. *et al.* (2012) Asymmetric segregation of polarized antigen on B cell division shapes presentation capacity. *Science* 335, 475–479
90. Gessel, M.M. *et al.* (2014) MALDI imaging mass spectrometry: spatial molecular analysis to enable a new age of discovery. *J. Proteomics* 107, 71–82
91. McDonnell, L.A. *et al.* (2010) Automated imaging MS: toward high throughput imaging mass spectrometry. *J. Proteomics* 73, 1279–1282