SECTION I

ESSENTIALS OF HIGH CONTENT SCREENING

CHAPTER 1

Approaching High Content Screening and Analysis: Practical Advice for Users

SCOTT KEEFER and JOSEPH ZOCK

1.1 INTRODUCTION

The topic of this book is the study of cells. What is in them, on them, around them, and between them. How they eat, sleep, grow, react to stimuli, and die. How they complete tasks and work as a team by signaling, influencing, stimulating, inhibiting, and sometimes destroying each other. High content screening (HCS) (1) is an imaging approach to cell-based assays that has had an impact in the fields of neurobiology (2–5), oncology (6–9), cell signaling (10–13), target identification and validation (14–17) and *in vitro* toxicology (18–21). If you have opened this book, you have most likely heard of high throughput screening (HTS), understand the premise, and have probably seen it utilized somewhere in the workflow of your organization. You have probably also heard of HCS and, hopefully, want to learn more about how it works and where it should be implemented. You might be a drug discovery scientist trying to transition targets from biochemical to cell-based assays. Or you could be an academic cell biologist who wants to generate a larger amount of statistically relevant data in a shorter time frame. Either way, our guess is that you are not viewing HCS as an all-encompassing career move, but rather as a new set of tools to get your job done. This chapter attempts to provide a frame of reference to fit HCS into your mindset by comparing the similarities and differences with several current assay methods. Some of the advantages of cellular imaging will then be discussed as we cover key process steps. Finally, we will leave you with some advice in the form of six points to remember to get the most out of your HCS data. The goal is to open your eyes to the possibilities this new tool has for rapidly expanding the breadth of cell biology that can be quantified, leading to new discoveries in both basic and applied scientific research.

High Content Screening. Edited by Steven Haney

1.2 WHAT IS HCS AND WHY SHOULD I CARE?

High content screening can be defined as an automated imaging approach to understanding compound activities in cellular assays where, in each well of a microplate, you can measure spatial distribution of targets in cells, individual cell and organelle morphology, and complex phenotypes. It provides the flexibility to measure cell subpopulations and to combine multiple measurements per cell, while simultaneously eliminating unwanted cells and artifacts. Recently, the term *high content analysis* (HCA) has also emerged to describe the broader view of multiparametric interrogation of cellular processes in any format.

The "content" is a set of output feature numbers, derived from an algorithmic extraction of fluorescence intensities per pixel within the digitized image of a cell. The number of measurements made for each cell can climb into the hundreds, depending on the number of fluorescent probes used. The raw data that are generated can then be combined to define a staggering number of biological states and phenotypes. These measurements, when applied to the screening of potentially bioactive entities, can describe a compound's cellular bioavailability, potency, specificity, and toxicity. Remarkably, this can often be achieved with one HCS assay by multiplexing assays with probes spread across the visible spectra.

The "C" in HCS also stands for "context." All HCS assays are performed with intact, living cells and, therefore, preserve the state of cell physiology created by the assay environment. In the early years of the pharmaceutical industry, before the development of the mainstream tools of modern molecular biology and biochemistry, context was one of the only ways scientists had of understanding a potential drug candidate's pharmacology. The process was essentially to make a test animal sick, "treat" it with a compound or extract, and observe it for indications that the animal was getting better or getting worse. Often, odd behaviors were noted that, although attributed to the treatment, could not be readily explained. Hence the moniker "black box" science. Then, over time, the application of advances in protein chemistry combined with genetic engineering allowed the isolation or creation of active proteins outside of the cell and the biochemical assays were born. This format could be completed in small volumes, and technologies to take advantage of this drove the number of assays carried out to over 10,000 a day (HTS) and eventually over 100,000 a day (uHTS). The problem was that the context of the "box" was lost in the process.

Why is context so important? We, as a scientific community, collect an enormous amount of biochemical assay data from HTS and try to use it to understand both general cell biology and compound effects, and yet some of the most important questions remain unanswered due to a lack of context. An analogy might help here. You are a brain surgeon trying to remove a tumor without destroying function. Your patient is on the table and you are stimulating different parts of the brain around the tumor to see the response. Stimulation in one spot causes the right index finger to move. Stimulation in another spot causes the right wrist to move. In an effort to not hit the wrong spot you ask the patient to watch a screen and recite aloud either the text or a description of a picture flashed before him. These pieces of behavioral information need to be collected and pieced together to get an idea of what the

tumor might be doing. Additionally, this process needs to have an intact patient to do it. It is not really about the index finger, or the wrist, and even if you have very specific and sensitive ways to identify and measure them, without the context of the whole patient you will not have the right information to be successful.

So it is with cell biology. HCS effectively shines a light into the black box, allowing for context of cell physiology and behavior while collecting multiple pieces of information simultaneously. Context allows for the determination of function. From quantifying the activation of multiple transcription factors in a cell signaling model, through identifying differentiated cell states in a stem cell assay, to assessing true target function in a genome wide RNAi knockdown study, HCS is the detection method of choice.

Finally, the "C" in HCS also stands for "correlation." Trying to interpret correlated results from multiple biochemical assays is often difficult because of compounding variability (lot to lot, pipetting, environmental, and so on). Additionally, each cell in the well has the potential to be in a different physiological state (i.e., cell cycle), often causing a blunting of activity readouts after population averaging. Systemic noise can be great enough to mask the interesting revelations you are trying to uncover. The best way to overcome these issues is to be able to make multiple measurements in each cell (biological variability) in the same well (environmental variability). High content screening not only collects data in this way, but it allows the results to be analyzed collectively from each cell to create highly correlated insights into how various targets react as a network.

1.3 HOW DOES HCS COMPARE WITH CURRENT ASSAY METHODS?

Useful assays that can be validated for screening have a common set of important characteristics, including selectivity, sensitivity, scalability, and robustness to automation. In this way HCS is no different than other current screening methods. The requirements for accurate pipetting, incubation, reagent control, plate washing, and proper assay development are very much the same. As the throughput requirements increase, automation of the assay process steps becomes necessary and is straightforward with commercially available instrumentation and robotics. So what are the advantages of HCS compared to current assay methods?

An enzyme-linked immunosorbent assay (ELISA) is designed to capture and quantify the amount of specific proteins or peptides by their epitopes using high affinity antibodies to create a target "sandwich." Recently, bead-based ELISA formats like Luminex® (Luminex Corporation, Austin, Texas, USA) have expanded the number of targets that can be simultaneously measured from a single sample. Typically, the target proteins are either already purified or come from an extract of cells in a particular biological state, resulting in the loss of spatial context. Therefore, it is impossible to readily identify which cells had a protein and where it was inside them. Many HCS assays also use antibodies as immunocytochemical affinity tags to label various cellular proteins, but retain the advantage of individual cell measures and subcellular location.

In one example, Gasparri et al. developed a multiparameter high content assay for proliferation of human dermal fibroblasts with fluorescent indicators for brdU incorporation, histone H3 phosphorylation, pRb phosphorylation, and KI-67 expression (22). Cross-validation by ELISA and flow cytometry uncovered comparatively fewer false-positive (fluorescent artifacts) and false-negative (cell loss) rates with the HCS assay, leading to the assertion that HCS data were inherently of higher quality. In summary, the authors cited higher accuracy of data, both single-cell and population readouts, and the ability to report morphological features as important advantages of the HCS approach.

Secondary signal assays like luciferase measure transcription activation indirectly and also require the cells to be disrupted into an extract before the luciferase reaction creates the chemiluminescent signal, thus losing the resolution of individual cell responses. This type of assay requires genetic engineering of a target promoter/luciferase gene chimera into the cells that competes with endogenous transcription factors (not measured). High content screening can directly measure endogenous protein levels and their positions over time. The individual cell responses are maintained, allowing the identification of subpopulations of cells with similar responses in each well. The following HCS example would be impossible with a standard second signal assay approach.

Vogt et al. performed a high content screen of a small compound library for inhibitors of ERK dephosphorylation (23). They confirmed the hits by visually inspecting cell images and with standard western blotting techniques. Analysis of the data showed that this group of compounds was enriched for known cdc25 inhibitors. *In vitro* enzyme assays showed that the ERK inhibitors identified in the high content screen inhibited at least one of the DSPases (MKP-3, cdc25B, cdc25A) *in vitro*. The authors then performed a multi-parameter high content assay for MKP-3 inhibition by transiently transfecting a c-myc-tagged version of MKP-3 into cells, then assaying for ERK phosphorylation via an intensity increase in the nuclear compartment in the two subpopulations. They reported a significant measurable difference in phospho-ERK accumulation between the MPK-3 overexpressing cells and the untransfected cells in the same wells. Additionally, the group determined that the compound having the best cellular activity was not one identified as potent in the biochemical screen, suggesting that performing this type of cell-based assay earlier in the drug discovery process is useful.

Flow cytometry or automated cell sorting, which has been the gold standard in cell biology for the last 30 years, has likely the most critical advantage when approaching a cell-based assay. The context of the cell is retained. By keeping the cell intact, flow cytometry permits measurements such as intensity, size, and count to be made. In addition, the multiple spectra capability of flow cytometry permits the multiplexing of targets. This is advantageous, because multiplexing tends to scale well and will often provide more insight as to a sequence of events rather than a single target screen.

There are, however, a few limitations to flow cytometry that one must consider. Cell sorting in general does not lend itself to adherent cell lines and there are a very limited number of morphologies that can be measured. Structure-related

TABLE 1.1 Capabilities of Various Cell-Based Assays.

Assay Type	Detection Mode	Detection Method	Intensity-Based	Cell-Based	Intact Cells	Multiplex Capability	Cell by Cell	Location	Subcelluar Structure	Multicellular Structure	Assay Development to Screening
Substrate Conversion	Colormetric/Fluorometric	Fluorometer	●	◐	X	X	X	X	X	X	●
RIA	Scintillation	Scintillation counter	●	◐	X	X	X	X	X	X	●
Elisa	Chemiluminescense	Spectrophotometer/ Luminometer	●	◐	X	◐	X	X	X	X	●
SPA	Scintillation	Scintillation counter	●	◐	X	◐	X	X	X	X	●
Luciferace	Chemiluminescense	Luminometer	●	●	X	◐	X	X	X	X	●
Geneblazer	FRET	Ratiometric fluorescence	●	●	●	◐	◐	X	X	X	●
FLIPR	Fluorescent	Plate reader	●	●	●	◐	X	X	X	X	●
Flow Cytometry	Fluorescent	Multilaser PMT flow cytometer	●	●	●	●	●	◐	○	◐	○
HCS/HCA	Fluorescent	Multispectral fluorescence imager with analysis	●	●	●	●	●	●	●	●	●

● Good Capability ◐ Limited Capability ○ Poor Capability **X** No Capability

measurements on a cell are difficult, if not impossible to make due to the flow of the sample. The process is to flow a stream of single cells, passing them through a laser beam to be detected, so it requires the use of large sample volumes (lots of cells) and results in high quantities of potentially hazardous wastes. Other considerations for flow cytometry include its high cost, large size, high maintenance, and extensive training requirements for the instrumentation, and have put it far beyond the reach of many laboratories. Most recently, these concerns have been addressed by vendors, who have built high quality, small application focused, benchtop systems, permitting this type of technology to be delivered even to the most modest of laboratories.

Microscopy, compared to HCS/HCA, has essentially the same technology and biological requirements, but workflow requirements for automation and reproducibility are quite different. Few microscopes have the walk up and run capability to scan multiple plates with multiple fluorochromes. Once scalability is necessary, attempts at in-house solutions can create a whole new set of issues. One might be able to reduce cost by building a system using a microscope and then integrating parts and pieces from various vendors instead of having a tested integrated solution. This takes considerable time and effort, resulting in a system that needs your expertise to maintain. Turnover of resources in this situation will be problematic, and the return on investment will erode as the number of plates and assay types that need to go through the system increases. The conclusion is that HCS platforms are much more than just a "microscope in a box" and provide technology transfer capability within the organization.

The greatest advantage of image-based platforms is the ability to see and record the biology by means of a picture. Truly, a picture is worth not only a thousand words, but with HCS, a thousand data points as well. Table 1.1 shows a variety of cell-based methods and how these map to features that are often significant when looking to implementing a cell-based assay.

Instruments for HCS are the best of many worlds, and as you proceed through this chapter and book, you will undoubtedly see the broad scope of its applications, technologies, and functionality. The unique superset that HCS provides combines the best features from imaging and fluorescence microscopy, microtiter plate readers, and the single-cell analysis of the flow cytometer. Combining these tried and true capabilities provides researchers of all kinds with a powerful and relevant new tool set to investigate and scale cell biology.

1.4 THE BASIC REQUIREMENTS TO IMPLEMENT HCS

All assays can be represented by the simple equation:

$$\text{Defined biology} + \text{change agent} + \text{detection} = \text{measured biological change.}$$

HCS is no different in this respect, so we will use these parts of the equation as the topics for discussion.

Defined biology in HCS assay development always starts with cells. Choosing a cell type or types depends on a number of criteria that need to be balanced. Primaries cell cultures or established cell lines? Human or other animal model? Direct or downstream targets? Biological relevance, availability (commercial or in-house), assay tractability, and reproducibility are usually the main considerations. Unfortunately, these features are often diametrically opposed, resulting in compromises about the choice of cell types for the screen. Literature searches can often lead to the cell types that are currently being used for a particular research focus area, like PC12 for neurite outgrowth, 3T3L1 for adipogenesis, and U20S for G protein coupled receptors (GPCR) activation. As the popularity of HCS and other cell-based assays has grown, so has the interest in manipulating cells to create commercially available biological models. Examples include engineered "redistribution" cell lines containing translocating green fluorescent protein (GFP) chimeras (24, 25), "division-arrested cells" where treatment retards the cell cycle, leaving the cells in a transcriptionally active, yet nonproliferating state (26), and screen-ready neurons that can be thawed, plated, and assayed. There is also a significant interest in using HCS with nonmammalian cells (27). We could spend several chapters covering cell types, but instead we want to focus on two common misconceptions that can have dramatic, and sometimes tragic, consequences for your HCS screens. The first is that a stable clone is actually stable. The second is that passaging cells over time does not change their physiology.

By the very nature of the process they go through to become "stable clones," cells producing heterologous proteins will attempt to limit nonessential protein production. Often, dramatic variability will be seen in physiological response between high, medium, and low expressors and an increase in the nonexpressing cell subpopulation over time. The consequences of these events while screening result in decreasing Z' scores (28), increasing false-positive or false-negative rates, and the inevitable decision to either halt the screen or redo parts of it.

To illustrate the second misconception, let us look at your relatives. You are genotypically 99.99% the same as your father, grandfather, and great grandfather (and only a few percent more away from a chimpanzee) and yet you look only similar and behave much differently, having been exposed to a different set of environmental conditions. Your phenotype has changed over a very short number of life cycles. It is the same for the cell lines used in HCS experiments, where modifying the environment can lead to changes in a cell's ability to respond to stimuli in a predictable way. Therefore, it is absolutely critical to (1) standardize cell passage procedures and (2) limit the number of passages (and therefore cell doublings) prior to assay plating by (3) bringing up a new set of cells from a frozen cell bank large enough to cover the screening campaign (and retests).

1.4.1 Cell Banking

Whether you are using primary, immortalized, or engineered cell lines in your assay(s), it is of the utmost importance to have enough cells with the same physiology to complete a screening campaign. This is not a calculation to take lightly and should

include contingencies for several "catastrophic failures" plus retests. Fortunately, most HCS assays require significantly less overall cell mass than other cell-based assays, and moving to a higher density (i.e., 384-well) microplate format can lower that requirement even further. If creating cell banks is not a core skill in an organization, several commercial options exist, many of which will even store a cell bank until it is needed. For primaries, where expansion for banking is often impossible, cells from multiple animal donors can often be mixed to give a large lot of pooled cells for screening. There is the chance, however, that pooling may just increase the variability of the response, where an approach using normalized results on non-mixed populations may be ultimately more reproducible.

1.4.2 Plating, Cell Density, and the Assay Environment

The process of transferring cells from a flask to a compartmentalized screening environment, typically a multiwell microplate, is generally called cell plating. There are a wide variety of Society for Biomolecular Sciences (SBS) standard microplates that can be used with a range of well densities (6, 12, 24, 48, 96, 384) in both plastic and glass. Using the "physiology is key" mindset, there are several important rules to follow, assuming that cell passage and banking are already being controlled as above. The first is to plan for the state of the cell environment during the assay window. This window is the actual time the cells spend in contact with the sample and the time required to react to the sample in some meaningful way. If planning for the cells to remain as individuals during the assay window, then plating density should allow for enough space between cells to cover the incubation time before the assay window (which may be several days if the cells have been trypsinized during plating). Examples of assays requiring this kind of plating include cell motility, transcription factor activation, morphology, and colony formation.

At the other end of the spectrum, an assay may require an intact monolayer of cells to achieve the correct biological state. To achieve this you must plan accordingly, calculating the degree of cell loss and doubling time to ensure a monolayer during the assay window. Examples of this kind of assay include receptor internalization, gap junction assays, viral plaque assays, wound healing assays, and tube formation assays. When considering using cells that need to be differentiated to achieve the right biological context, then the process has added complexity, as the pre-assay time will effect decisions regarding plating densities.

Another important consideration in the design phase is the assay environment itself. Cells may require a substrate or matrix to attach to in order to achieve the desired biological state. There are numerous examples of plate coatings used in HCS assays (e.g., Collagen IV plates for neurite outgrowth, BD Matrigel™ (BD Biosciences, San Jose, California) for endothelial tube formation). The assay may be built around two or more cell types interacting with each other, and methods need to be devised to mix cell populations appropriately. An excellent example of this is the "gap junction" assay by Li et al. where labeled cells are deposited on top of a nonlabeled confluent cell layer (29). As the labeled cells create active gap junctions, the dye moves through the portal between the cells and is measured as a

growing population of label positive cells in the immediate vicinity of the donor cells. This is a very distinct pattern compared to cytotoxic events, so internal controls for false positives are a benefit of this approach. The assay may require more complex structures inside the wells to simulate the appropriate conditions for a biological response like those found in chemotaxis/migration assays (30, 31) or in cell motility assays (32).

1.4.3 Compound Addition and Incubation

Compound addition for HCS assays has a special set pitfalls to avoid, due primarily to the properties of DMSO, the most common solubilizing agent for most compound libraries. First, most cell types will not tolerate concentrations of DMSO above 1% final concentration, and we strongly recommend staying well below that threshold to minimize uncontrolled fluctuations in physiology and reaction to treatment. The second consideration is a phenomenon coined "the liquid plummer effect." Essentially, like the drain cleaner that pours through standing water to get to the clog, concentrated DMSO with solubilized compounds is initially relatively immiscible with, and heavier than, the media cells. Without immediate mixing, the added aliquot will sink to the bottom of the well (where the cells are) before it gets mixed completely, causing exposure to much higher (toxic) levels of both compound and DMSO. The resulting temporary gradient forms from the center of the well outward, causing a high degree of response variability across the well, ranging from immediate cell death (middle) to no response (edges). There are several ways to solve this problem, from mixing in the tip to compound predilution in warm media and transfer of larger amounts of fluids to increase mixing. There are also newer "touchless" pipetting technologies (solenoid, piezo, and acoustic) that deliver the compounds in very small (picoliter) aliquots, facilitating distribution in the wells.

Once compounds have been delivered to the cells, incubation is required to allow the targeted biologies to develop appropriately. Although the timeframes for this incubation can vary from a few minutes to a few days, the general rule is to minimize and control the time out of the incubator as much as possible. As attempts are made to increase the throughput of the HCS, variability will necessitate the use of automated incubators designed for handling microplates.

1.4.4 Post-Assay Processing

When the biological processes to be measured have run their course, it is time to label the cell components needed for proper image analysis. Often this will include marking regions (structures) of interest and targets of interest with spectrally distinct fluorescent tags. These fluorescent reagents loosely fall into three classes: (1) auto-fluorescing proteins (AFPs) that are engineered into the cells as chimeras, (2) fluorescing dyes that are taken up by the cells and concentrate in various organelles based on charge or molecular affinity, and (3) antibodies that have affinity for target epitopes and can be tagged directly or indirectly with fluorescent molecules.

These classes are often used in combination to paint a multifaceted picture of multiple components inside each cell. The choices made about reagents during assay development often relate back to whether the biologies being imaged and measured can be done with fixed cells or must be performed as a live cell single endpoint or extended multiple endpoint type imaging. Certain reagents, such as antibodies, can only be used after fixation and membrane permeabilization to allow the large antibodies to reach their targets. Typically there is a primary antibody to the target that is allowed to bind. After the unbound primary is washed away, a fluorescently labeled secondary antibody with affinity for the first antibody is bound, creating an accumulation of fluorescence at the target site. Other reagents, such as JC-1 for mitochondrial integrity or Fluo-4 (Invitrogen, Carlsbad, California) for calcium flux can only be used in live (unfixed) cells. There are numerous commercial sources for fluorescent dyes, antibodies, and even AFP engineered cell lines designed to measure specific signaling events. New dye types such as "quantum dots" are coming into the arena and hold the promise to increase the useful number of distinct spectral signatures available. Additionally, HCS reagent kits from several vendors are becoming increasingly available and provide biologically validated components and protocols that can dramatically shorten assay development and project target deadlines.

Live endpoint can also be followed up with fixed endpoint for attempts to correlate target biologies that happen at different timepoints. An example of this would be measuring a calcium flux (live) in a cell line with an engineered GFP translocation of a transcription factor (live), followed by cell fixation and staining of the cytoskeleton to measure a downstream morphology endpoint (fixed).

No matter what are the ultimate design and endpoints of the HCS assays, the key considerations to keep in mind are reagent stability, specificity, and availability. Success, as with all biological assays, depends on being able to reproduce the same set of circumstances, with minimal perturbation, over extended periods.

1.4.5 HCS Imaging Hardware

Choosing an HCS platform is a more complicated affair now that there are numerous vendors delivering HCS instrumentation to the market. In-depth discussions of different imaging options and their advantages are provided in the chapters that follow. At some point you will be evaluating which HCS platform matches your needs, and we have two thoughts to keep in the back of your mind during this process: (1) "Fight for the right light" and (2) "Don't try to use a sledge hammer to pound in a tack." Because the goal of any assay is sensitivity, and the output of an HCS assay is light, the "fight" is to optimize the proper light capture by choosing the right materials. Anything that generates, collects, focuses, or blocks light on the way to the camera is in play. Components like light source, optical train, autofocus modes, filter/dye selection, and even plate type must be evaluated for their ability to consistently provide you with the right light. The second piece of advice is just as important. Many new users of HCS try the highest magnification their systems

have because they think this will give them the best chance of success. The reality is that you want to use the lowest magnification possible while still resolving your objects, because this allows you to collect a larger (and statistically more relevant) number of objects in the shortest timeframe.

1.4.6 HCS Analysis Software

As expected, there is new software with which to become familiar when trying to implement HCS into a discovery workflow. This software falls into two main categories: the acquisition or control software (how you operate the machine) and the analysis software (how you get the data out of the images). Philosophies differ among vendors and users of HCS equipment as to the flexibility and ease of use of products. However, we recommend a few general rules when shopping or implementing this technology.

1. *Look for a complete platform that reflects the workflow.* Instruments/software that adapt to multiple workflow environments are typically the best. Is the instrument being used as a workstation for both imaging and data review, or will the data review be carried out somewhere else? The more workstations that can be set up, the more flexible can the workflow be.
2. *Consider the ease of use of the control software.* The harder the control software is, the less likely someone is to use it.
3. HCS is a very visual technology that allows the user to visualize the biological phenomena before committing to a scan. Because of this, one must be able to interact easily with a sample, interactively changing focal planes, object types, exposure times, filter sets, and enabling/disabling confocal or optical sectioning capability.

Analysis software, which is covered in detail later in this book, falls into two basic camps. The "out-of-the-box" camp typically consists of preconfigured flexible analysis routines that may or may not have been validated (or proved to work with supporting data), while the "free-form" scripting routines or drag-and-drop type algorithms allow the user to create from scratch or build upon an existing algorithm foundation. It should be noted that both types provide detailed analysis capabilities and are very flexible in the breadth of biologies that can be measured.

1.4.7 Informatics

All the advantages of HCS, with its flexibility and in-depth interrogation of each cell, come with a price. The price is not the dollar amount of the instrument (although that is important too); it is the attempt to manage anywhere from 500 GB to 6 TB a year of this multiparametric data. Arguably, the systems are not worth much if it takes too much time or it is too difficult to extract the data . . . after all isn't that why you're implementing this sort of technology?

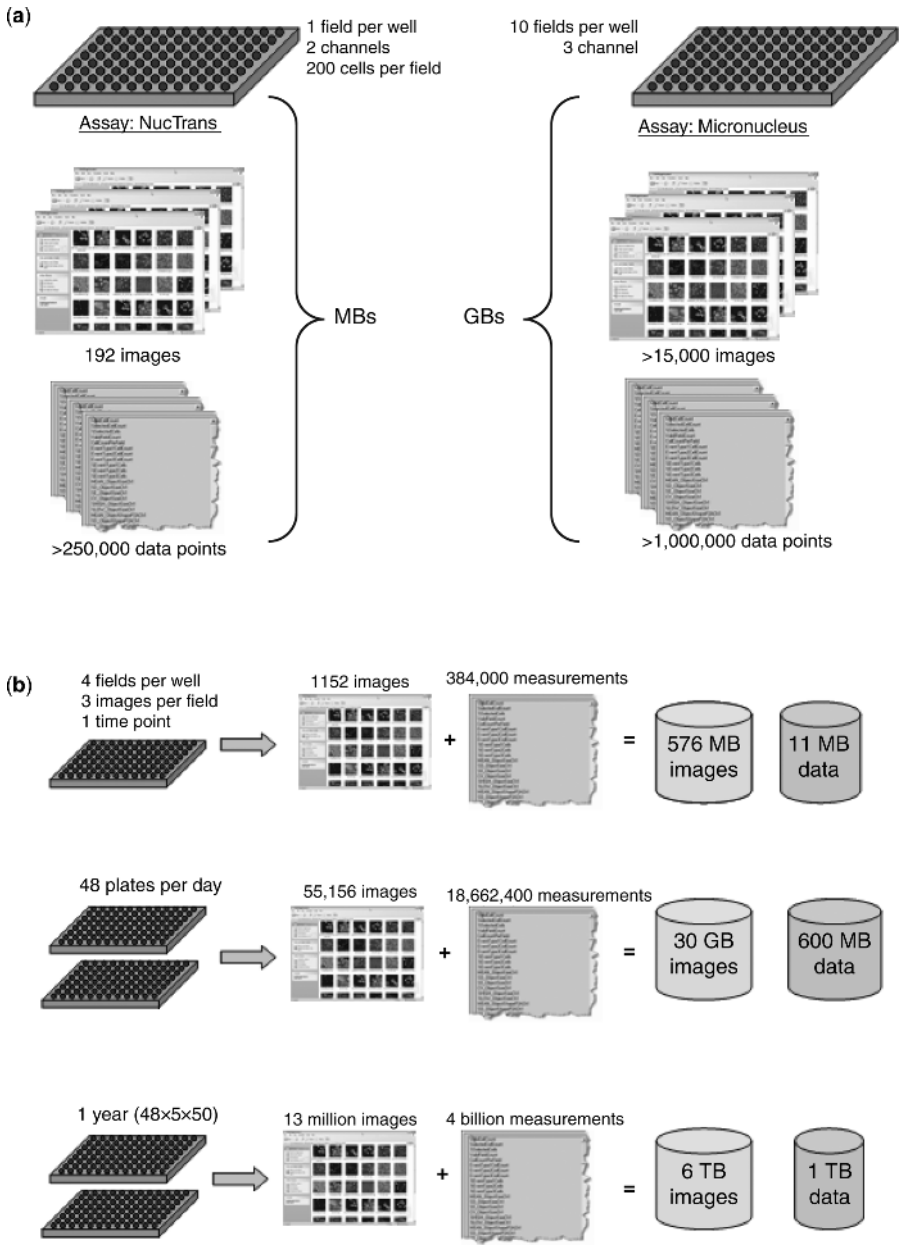

Figure 1.1 (**a**) Estimate of the amount of data generated from a simple two-color translocation (left) assay and a three-color Micronucleus assay (right). (**b**) Data generated from a per plate basis (top), per day basis based on 48 plates per day (middle) and over a year (bottom) screening at this level of throughput.

There are two items from this section you must take away:

- An IT department must be involved up front in the decision-making process in implementing HCS.
- Do not underestimate the necessity to collect, organize, visualize, archive, or report this massive amount of information.

Think forward enough so that the infrastructure that warehouses this data is suitable for the next 18 months and be sure that the informatics solution you choose is scaleable to meet future needs. The more seamless an informatics solution is, the more time that will be available to focus on the science and the data rather than on the mechanics of how it got to a server.

In the following is a quick example of the amount of data generated from a few real world examples: Fig. 1.1a demonstrates the typical amount of information that is generated from a simple assay comprising two fluorescent channels scanned on one plate. One imaged field per well generates megabytes worth of images and data, and just 10 imaged fields per well generates gigabytes worth of images and data. Taking that scenario, Fig. 1.1b shows what happens when we scale up and begin to scan multiple plates in a multiplexed environment. With just four imaged fields and three fluorescent channels, we generate nearly 11 MB of images and data. Using a reasonable throughput of 48 plates a day (instruments will vary as to capacity) systems will generate 30+ GB of images and 600 MB of data per day. Over a year at this pace, it is not unreasonable to get 6+ TB of image files and 1 TB worth of data. Few systems are operating at this capacity, but the point is that these systems generate a lot of data and you (and your IT department) need to be aware of it.

1.5 THE PROCESS

So you have assembled all the basics and are becoming familiar with the cell biology and the tools. You have been thinking about how to create the biological scenario: picking the cells, picking the environment, looking for control compounds that cause the changes in the cells that are to be measured. Literature searches can be very helpful here, as the majority of biological phenomena have been studied to some level before. Additionally, there are some great compilations of bioactive molecules that can be used as positive controls from companies like EMD and Tocris Cookson. You have gone through a mental (or empirical) evaluation of the types of fluorescence reagents you are going to stain/label the cells with and have produced the first images of both nonstimulated and stimulated states by manipulating magnification and exposure. The general rule is that if you cannot see the difference in your two biological states, the algorithm will not either. This does not mean you necessarily need to know the best way to quantitate the difference; the algorithm should give you various outputs of intensity, shape, and texture from which to choose.

At this point, the newly initiated often find themselves on the slippery slope of how to begin to algorithmically extract the right data from the images. This is

especially acute when multiplexing several targets and trying to understand what information the images are providing. As more targets are added and fluorescent targets are combined, it is critical to ensure that artifacts are not created in the images. Sources of such artifacts include (1) emission "bleed through" when excitation spectra overlap too much, (2) quenching of fluorescence by unwanted fluorescence resonance energy transfer (FRET)-type reactions, and (3) visual changes (high background, morphological, and so on) coming from interactions between reagents and/or biological states.

One approach for moving from the image to some defined outputs is called the "IVS Method" or "Identify, Validate, and Select." Make sure you have correctly identified the objects you want to measure first (usually through thresholding), then make sure you are choosing to measure *only* valid objects meeting your criteria, and finally you can decide to select (gate) objects, which have been identified in downstream channels as positives. Apply the IVS to each "virtual assay" in your multiplexed set independently, mentally treating them as separate assay developments. When you are satisfied that you have a basic grasp of each and are collecting the right kind of outputs then go back and look at the assays as a whole, looking for dependencies that could create artifacts. Remember, every virtual assay will have its own natural biological variability and preparation variability, so multiplexed assays must be thought of as a composite of individual assays when assessing assay quality. There are numerous ways to assess whether an assay is good enough to use (33), but the current gold standard of assay validation is the Z′-score (28), which will be described in detail in a later chapter.

The last piece of advice concerning the HCS development process is iteration. Establishing that your parameters work with larger data sets is critical for producing robust assays. No one should ever develop an assay from a single set of images because, invariably, the initial image set you took missed something visually critical that you need to consider. This might be as simple as noticing staining variability, which needs to be addressed as part of the development, or as complex as an unwanted cellular phenotype that routinely shows up to confound your data and can be removed algorithmically. Alternatively, you may realize that this aberration is providing yet another piece of information (e.g., apoptotic cells showing up in a transcription factor translocation assay) and set your parameters to include measuring this subpopulation as a part of the assay.

1.6 AN EXAMPLE APPROACH

The following is an example from a poster that Vivek Abraham, Brent Sampson, Oleg Lapets, and Jeffery Haskins from Cellomics® (now Thermo Fisher Scientific) presented at the SBS in 2004 that demonstrates a multiplexed HCS approach to assess the progression of toxicity and the difference between apoptotic and necrotic cell death. The group used common forms of cell death that included mitochondrial transmembrane potential, phosphatidylserine extrusion, nuclear morphology, and plasma membrane integrity as sensitive measures of cell death.

Each target was amenable to the HCS approach, because fluorescent labeling and image analysis routinely perform these types of analysis as separate and distinct assays. The uniqueness of this approach was to use these markers to create a story of what was happening as a cell progressed towards death rather than a single decision point of alive or dead.

Metrics that exhibit profound cell death due to compound treatment were reflected in increased detection of fluorescence in the plasma membrane integrity channel, while other indicators of poor cell health were measured by more subtle nuclear morphology changes or a decrease in mitochondrial potential. What was demonstrated by the group is the power to collectively use the different targets in conjunction with one another to detect the progression of toxicity and the difference between apoptotic cell death and necrotic cell death.

The approach used to demonstrate this progression used HCS and the associated image analysis software to set up a series of specific events that would describe whether a cell falls into an early-stage reversible toxic event (cells that have decreased mitochondrial potential, *but not* any nuclear condensation or increased plasma membrane permeability) or a late-stage cellular injury (condensed nuclei or increased cell membrane permeability). Figure 1.2 shows the logic that was used to determine if a treatment was early- or late-stage irreversible toxicity.

In the same assay, using the phosphatidylserine serine (PS) extrusion via Annexin V staining, the group could additionally discriminate, using the same approach, the difference between an apoptotic event (PS extrusion but not increased permeability) and necrotic cells (PS extrusion and nuclear fragmentation but not increased permeability) (Fig. 1.3).

The approach taken by the group was not only to obtain the desired decision points of the tested compounds quickly, but to describe more subtle phenotypes that their

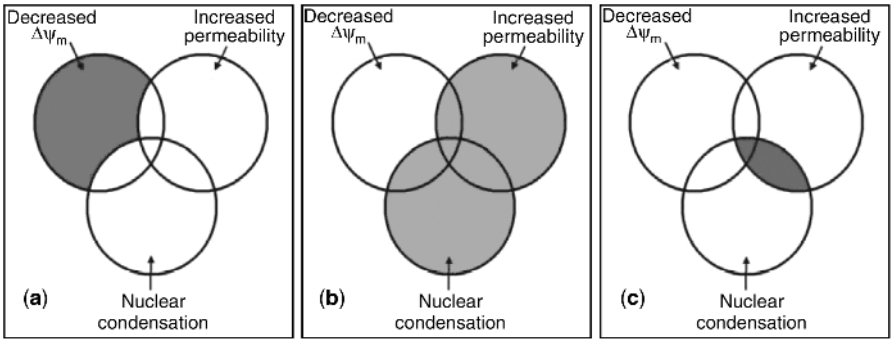

Figure 1.2 Definitions used to assess progression of toxicity. (**a**) Decrease in mitochondrial potential with neither nuclear condensation nor increased membrane permeability (early toxicity). (**b**) Nuclear condensation or increased membrane permeability (late-stage, irreversible toxicity). (**c**) Both nuclear condensation and increased membrane permeability (also late-stage, irreversible toxicity).

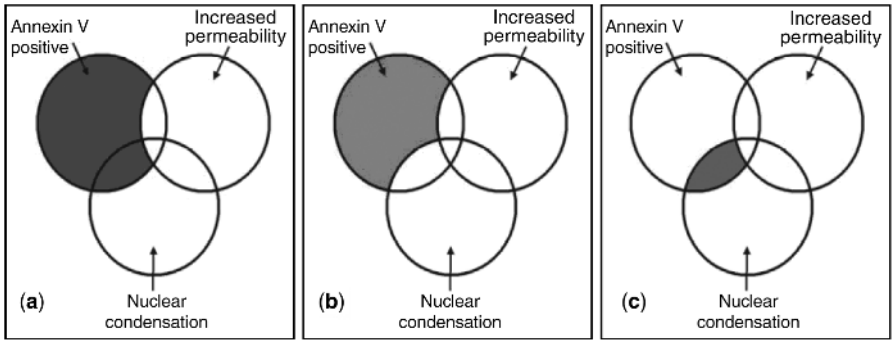

Figure 1.3 Definitions to distinguish apoptosis from necrosis. (**a**) Annexin V labeling in the absence of membrane permeability (occurrence of apoptosis in the absence of necrosis). (**b**) Annexin V labeling in cells that show no nuclear condensation (early apoptosis). (**c**) Annexin V labeling in the presences of cells with nuclear condensation (late apoptosis).

compounds effect via early- versus late-stage toxicity and apoptosis versus necrosis. Each target on its own tells a specific story about that target, but collectively may be used together to make a decision, such that one can discern the difference in the profile of a compound with respect to its cell health with exquisite sensitivity.

The approach of this assay requires good and thorough homework. The right reagents, cell types, HCS platforms, and a planned approach should not be taken for granted. This section, along with the "Process" section, hopefully presents itself as a practical aspect when considering implementing an HCS assay and will spark more innovation to create robust cell-based assays.

1.7 SIX CONSIDERATIONS FOR HCS ASSAYS

Although not a comprehensive list of considerations one must make when approaching or using an HCS assay, the following, in our opinion, are some the most critical factors that must be thought about when embarking on HCS implementation.

1.7.1 Garbage In, Garbage Out (GIGO)

The GIGO philosophy applies directly in the world of high content analysis. The "garbage" in HCS may have many sources, but most are due to processes that happen in the preparation of plates rather than instrument problems. As HCS begins with cells, the most common mistakes have to do with cell plating and the labeling of cells for detection. Again, it cannot be over-emphasized, one must be extremely careful when seeding a plate, regardless of the preferred density, because voids on plates can lead to imaging areas that can generate erroneous data and extended scan times.

When treating cells with a sample compound or condition, methods must involve the gentlest technique possible. Preferably, automation should be employed whenever possible, but the vast majority of cell preparation is done manually on the bench. Essential tools such as a vacuum aspirator and variable-speed pipettor are necessary for quick, accurate, and reproducible work. The two most common errors in plate preparation are the washing off of cells because the sample condition caused cell death, and the lack of signal that comes from the cells as a result of fluorescent labeling anomalies. As you would suspect, both problems are typically catastrophic when attempting to image the plate.

A note on plate types and their relationship to HCS assays. Currently, there are many different plate types available for imaging. By their nature, the thickness of the plastic or glass may restrict the types and magnifications of objective one can use. Use plates designed for imaging; the manufacturer should specify this.

Due to the nature of HCS and the camera technology used in most platforms to derive data from the image, it is safe to say that image quality is of critical importance. If one cannot create suitable images for analysis, whether through cell plating and culture problems, fluorescent labeling, or plate type, data will be extracted from the plate improperly and will lead to erroneous results. As a rule, bad images will result in bad data 100% of the time. Unfortunately, the reverse is not true. A good image may not necessarily lead to good data. As the process section of this chapter describes, good data is the outcome of detectable biological changes that can be quantified and repeated.

1.7.2 This Is Not a Plate Reader

Instruments for HCS deliver an incredible depth of information about what is going on in each well and in each cell beyond what a plate reader can provide. The imaging approach also allows the rejection of fluorescent objects that do not meet the criteria (stuff I want to measure versus stuff I do not want to measure), which plate readers cannot do. To the novice this can be an overwhelming amount of data if they are used to the relatively low content of information coming from standard plate reader formats. The tendency is to initially ignore the high content and focus on the output feature you have decided on as the number describing your activity. Do not fall into this trap. The other data are there for a reason and can often convey the important subtleties that are going on in the well. Certainly, looking at a number of cells or fields required per well can tell you a lot about acute toxicity. Changes in nuclear size, shape, and texture can also be incredibly predictive of toxicity. Even looking at the size of the standard deviations can suggest whether you might have a nonuniform distribution of activities in the well or a rogue population undergoing a significantly different pathway (e.g., apoptosis). Incorporating Boolean-based classification and reference well capabilities found in several of the platforms provides substantial advantages over plate readers and will lead to a higher understanding of the complexities of biologies measured.

1.7.3 Understand Your Biology

Understanding your biology is key to developing robust high content assays. If you cannot visually distinguish the positive from the negative biological state, it will be difficult for any imaging algorithm to do it. If, somehow, you manage an algorithmic miracle and distinguish a difference, you still have to defend the results to your peers, who may be very wary of a biology that cannot be explained visually. Therefore, successful HCS assays usually measure a series of visually simple and distinct phenomena. This is not to suggest that the underlying biology being measured is simplistic. On the contrary, cell motility, neurite outgrowth, angiogenesis, and colony formation have multiple biological pathways involved, yet the phenotypic endpoints measured are combinations of object areas, object intensities, object processes, or object connectivity. From an imaging standpoint, understanding the biology also means realizing that assay artifacts (not just dust bunnies) can also cause either increased variability or misinterpretation of the data generated by an algorithm that has not been optimized to avoid artifacts. A classic example includes apoptotic cells (and sometimes mitotic cells), which look positive for nuclear translocation of a transcription factor, not because it has clearly moved from the cytoplasm to the nucleus, but because the loss of cytoskeletal structure causes the cells to round up, making the cell look positive for the translocation event. It is up to you to understand this and use object rejection criteria to remove these cells from the analysis if you so choose. A second classic example involves cell loss due to cytotoxicity. The cells that remain in the well are identified and measured for the primary endpoints, even though they do not really represent the original population you are seeking to quantify. Typically, these cells are abnormal in some way, making them resistant to the toxic insult. Therefore, the primary endpoints are not valid and, if you only look at these endpoints, you will end up in trouble. Again, you need to understand what is really happening in the well by looking at an alternative feature like "cell count per field."

1.7.4 Subtle Changes Can Be Measured and Are Significant

Imagine yourself at a lab meeting ready to present work on compounds you have tested, which cause predictable changes in nuclear size/area, but the results you have from your HCS data do not indicate even a twofold change. All is not lost, however; your data, albeit a small window, are significant, and there are ways to describe the data, or characterize the population of relevance providing for a larger window.

The process called "population characterization" is a powerful way to both calculate and present data, especially when windows between positive and negative controls are small. This approach bins each cell into user-defined (biologically relevant) categories of above the mean or below the mean of the referenced population. You control the stringency of how high or low below the mean to bin the cells via manually set thresholds based on raw settings or standard deviations of the mean. The result is that even subtle shifts become more significant, because the reported outputs describe the percent of cells in the high and low bins versus the total or average fluorescence seen from all the cells. Besides a more sensitive

way to measure a significant biological change, plate-to-plate variability is mitigated, because each cell is referenced to the control population within that plate.

1.7.5 HCS Workflow — Flexibility is the Key

It is critical now to think about your system as the time it takes from when you put your biology on the system to make a decision about the biology. It is not about how fast the system images, but how quickly, accurately, and reproducibly you can make a decision about your experiment . . . then how quickly you can do it again and again.

Depending on your day-to-day workflow, systems that generate results on the fly are much faster at getting the user to a decision, as opposed to ones that require off-line image processing. Having the ability to change modes between on-the-fly or post-processing and the flexibility to develop assays disconnected from the instrument are usually most desirable. The more flexible and proven a platform is in its ability to do these things, the more likely it will adapt to your and other people's workflow. HCS platforms provide a wide array of workflow needs, and understanding your needs will allow you to choose the right platform.

1.7.6 HCS is Hard — How Do I Learn It and Become Proficient at It?

Either you have found out by now, or will find out shortly, that designing, carrying out, and interpreting HCS assays is not easy. Keeping in mind the GIGO concept, sorting out what went wrong before the plate even gets to the instrument is typically one of the largest hurdles to obtaining good data. Once an assay has been validated, progressing towards a standardization process will allow for greater repeatability and confidence in the results. Ultimately, the most difficult part of HCS is not in the detection or in the preparation, it is in the extraction of the data and interpreting what that data means. Thousands of wells and millions of calculated data points will be overwhelming without the right tools.

It is our experience that there are three areas in which you must immerse yourself in HCS in order to become a champion. First, just like in life, it is about who you know and the connections you make. When HCS came on the market in the late 1990s, customers used vendors as the primary resource of information and contact for expertise. Vendors typically had two things in mind. Make it work reasonably well and sell more. Today HCS is approaching the mainstream, with hundreds of instruments and thousands of users worldwide who can share knowledge of their experiences rather than the fortunate few. The ability to connect with these people through user group meetings and conferences is essential and will remain one of the most powerful tools in learning what HCS can do for you.

The second area is the use of training opportunities, whether in the field or on site. Every vendor has some level of initial training that shows new users how to run the equipment. Although essential, the instruments are becoming easier to use, and the challenge then shifts to one of biological interpetation of the data.

A professionally established training program that involves a variety of colleagues from different types of institutions is desirable because of the interaction that takes place between the groups. If experience really equals productivity, then using the collective experience from both internal and external resources must be better than a closed approach to learning.

Finally, when implementing HCS, it is 15% platform and 85% biology. It is our recommendation that you look at the support network you plan to use, and ask the question, "Can these people support the 85% of questions I have about the biology I want to use on this type of platform?" If the answer is no, then consider how you are going to build your team in order to accomplish your projects before you read any further. If the answer is yes, then you are ready to read the rest of this book and use the experience of these experts in the field of high content screening. Good luck!

REFERENCES

1. Giuliano KA, et al. High content screening: a new approach to easing key bottlenecks in the drug discovery process. *Journal of Biomolecular Screening* 1997;2:249–259.
2. Richards GR, et al. A morphology- and kinetics-based cascade for human neural cell high content screening. *Assay and Drug Development Technologies* 2006;4:143–152.
3. McIlvain HB, et al. Pituitary adenylate cyclase-activating peptide (PACAP) induces differentiation in the neuronal F11 cell line through a PKA-dependent pathway. *Brain Research* 2006;1077:16–23.
4. Simpson PB, et al. Retinoic acid-evoked differentiation of neuroblastoma cells predominates over growth factor stimulation: an automated image capture and quantitation approach to neuritogenesis. *Analytical Biochemistry* 2001;298:163–169.
5. Fennell M, Chan H, Wood A. Multiparameter measurement of caspase 3 activation and apoptotic cell death in NT2 neuronal precursor cells using high-content analysis. *Journal of Biomolecular Screening* 2006;11:296–302.
6. Soncini C, et al. PHA-680632, a novel Aurora kinase inhibitor with potent antitumoral activity. *Clinical Cancer Research* 2006;12:4080–4089.
7. Inglefield JR, et al. Apoptotic responses in squamous carcinoma and epithelial cells to small-molecule toll-like receptor agonists evaluated with automated cytometry. *Journal of Biomolecular Screening* 2006;11:575–585.
8. Lovborg H, Gullbo J, Larsson R. Screening for apoptosis – classical and emerging techniques. *Anticancer Drugs* 2005;16:593–599.
9. Namn J-S, et al. 5-aza-2′-deoxycytidine restores the E-cadherin system in E-cadherin-silenced cancer cells and reduces cancer metastasis. *Clinical and Experimental Metastasis* 2004;21:49–56.
10. Bertelsen M, Sanfridson A. Inflammatory pathway analysis using a high content screening platform. *Assay and Drug Development Technologies* 2005;3:261–271.
11. Borchert KM, et al. High-content screening assay for activators of the wnt/fzd pathway in primary human cells. *Assay and Drug Development Technologies* 2005;3:133–141.

12. Vakkila J, DeMarco RA, Lotze MT. Imaging analysis of STAT1 and NF-kB translocation in dendritic cells at the single cell level. *Journal of Immunological Methods* 2004;294:123–134.
13. Schlag BD, et al. Ligand dependency of 5-HT2C receptor internalization. *Journal of Pharmacological Experimental Therapeutics* 2004;310:865–870.
14. Moffat J, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 2006;24:1283–1298.
15. Chang KH, Zandstra PW. Quantitative screening of embryonic stem cell differentiation: endoderm formation as a model. *Biotechnology and Bioengineering* 2004;88:287–298.
16. Guiliano K, Chan Y-T, Taylor DL. High-content screening with siRNA optimizes a cell biological approach to drug discovery: defining the role of p53 activation in the cellular response to anticancer drugs. *Journal of Biomolecular Screening* 2004;9:557–568.
17. Björklund M, et al. Identification of pathways regulating cell size and cell-cycle progression by RNAi. *Nature* 2006;439:1009–1013.
18. O'Brien PJ, et al. High concordance of drug-induced human hepatotoxicity with *in vitro* cytotoxicity measured in a novel cell-based model using high content screening. *Archives of Toxicology* 2006;80:580–604.
19. Morelli JK, et al. Validation of an *in vitro* screen for phospholipidosis using a high-content biology platform. *Cellular Biology and Toxicology* 2006;22:15–27.
20. Diaz D, et al. Evaluation of an automated *in vitro* micronucleus assay in CHO-K1 cells. *Mutation Research* 2007;630:1–13.
21. Haskins JR, et al. Thiazolidinedione toxicity to isolated hepatocytes revealed by coherent multiprobe fluorescence microscopy and correlated with multiparameter flow cytometry of peripheral leukocytes. *Archives of Toxicology* 2001;75:425–438.
22. Gasparri F, et al. Quantification of the proliferation index of human dermal fibroblast cultures with the ArrayScan high content screening reader. *Journal of Biomolecular Screening* 2004;9:232–243.
23. Vogt A, et al. Cell-active dual specificity phosphatase inhibitors identified by high-content screening. *Chemical Biology* 2003;10:733–742.
24. Lundholt BK, et al. Identification of AKT pathway inhibitors using Redistribution® screening on the FLIPR and the IN Cell 3000 Analyzer. *Journal of Biomolecular Screening* 2005;10:20–29.
25. Almholt K, et al. Changes in intracellular cAMP reported by a Redistribution® assay using cAMP-dependant protein kinase-green fluorescent protein chimera. *Cellular Signaling* 2004;16:907–920.
26. Vasudevan C, et al. Improving high-content-screening assay performance by using division-arrested cells. *Assay and Drug Development Technologies* 2005;3:515–523.
27. Wheeler RT, Fink GR. A drug-sensitive genetic network masks fungi from the immune system. *PLoS Pathogology* 2006;2:e35.
28. Zhang JH, Chung TD, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of Biomolecular Screening* 1999;4:67–73.
29. Li Z, et al. Identification of gap junction blockers using automated fluorescence microscopy imaging. *Journal of Biomolecular Screening* 2003;8:489–499.

30. Mastyugin V, et al. A quantitative high-throughput endothelial cell migration assay. *Journal of Biomolecular Screening* 2004;9:712–718.
31. Richards GR, et al. Quantitative assays of chemotaxis and chemokinesis for human neural cells. *Assay and Drug Development Technologies* 2004;2:465–472.
32. Bhawe KM, et al. An automated image capture and quantitation approach to identify proteins affecting tumor cell proliferation. *Journal of Biomolecular Screening* 2004;9:216–222.
33. Sittampalam GS, et al. Design of signal windows in high throughput screening assays for drug discovery. *Journal of Biomolecular Screening* 1997;2:159–169.