

Learning machine	1152
Raising flags	1153
Cell bio curve ball	1155

Digging deep and wide into single cells

Vivien Marx

Instruments maximize the yield in high-content single-cell imaging, which calls for carefully planned experiments and data analysis.

The myriad physical changes that can be observed in an individual cell under different environmental conditions represent a wealth of information for scientists to mine. This information can be captured by imaging cells under many experimental perturbations at a resolution sufficient to observe subcellular details. For such high-content screening, scientists must design their experiments to maximize their instruments' functionality. Researchers also need to plan data analysis well before a data mountain stacks up in front of them.

Reacting to the desire of scientists to harvest as much detail as possible from single cells, companies are providing new capabilities by evolving their instruments and adding software tools, special reagents, cell lines or new types of cell culture media. High-content analysis has in the past been mainly at home in pharmaceutical research, but this approach is increasingly spreading to academic institutions, where scientists can make wide ranging cellular perturbations and look at thousands of cellular changes. At the same time, pharma researchers are widening their view of cellular parameters—for example, with the types of assays they use.

Experimental perturbations can include gene silencing with siRNA or chemicals. The readout may take the form of fluorescent reporter expression or changes in cell components or overall cell shape.

Chris Bakal at the Institute of Cancer Research (ICR), which is part of the University of London, has two high-content imaging microscopes in his lab. When the instruments run at full capacity, they can generate 200,000 images a day. In his research aimed at linking genotype and phenotype, he uses gene silencing to knock down genes one a time and study

cellular changes in different cell lines. Each well in a 384-well plate delivers a fountain of results, with cells in each well reacting to a specific perturbation, allowing hundreds of conditions to be analyzed within minutes.

He calls his line of research “high content in high throughput,” adding that he and his team leverage speed and automation to be able to focus more on the content side of the science. “We develop methods to describe single-cell phenotypes and population heterogeneity from images acquired using high-throughput microscopes,” Bakal says. Given the complexity of the methods involved, throughput hits analytical speed limits. “We can easily acquire hundreds of thousands of images in a day—but it takes months to analyze them.”

In his view, imaging technology itself is now “extremely robust,” enabling fast querying of cells in multiple-well plates “with high accuracy.” During his postdoctoral fellowship in Norbert Perrimon's lab at Harvard Medical School, he explored and tested different imaging systems. “Of course you have to trade quality and throughput all the time,” he says. During his fellowship he created quantitative ways to profile cell shapes; such profiling allows scientists to connect each facet back to the genetic factors regulating these cell traits¹.

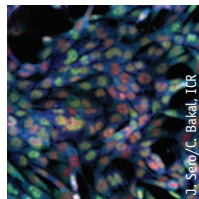
As he tried different imaging technolo-



“We are extracting hundreds of different features from single cells,” says Chris Bakal.

gies, Bakal saw that many systems and software packages have been developed with drug discovery rather than cell biology in mind. In screens, drug developers are looking for as little as a single compound showing an effect that merits further study, much like finding a needle in a haystack. “But we're much more beyond that now, and what we're trying to do is find the relationship between hay and hay,” he says, alluding to the subtle cellular changes he studies and his search for complex relationships between genes and descriptions of cellular networks. “That is much more challenging.”

Calculating many of these cellular features is time consuming and not usually part of a classic type of drug screen. “If you're doing millions of compounds, you just might not have the time to calculate the features: the protrusions, the ruffles, of every single cell in your library,” he says. “But if you are really interested in cell morphology like we are, then that is exactly what you want to calculate.” To satisfy his research needs, he adds to the instruments' software with his own algorithms to dig deep into the cellular features he is studying. When he set up his lab, he chose a confocal automated imaging system



Discerning differences among cells takes microscopy, reagents, software and a fine eye. Shown here are triple-labeled genetically identical mouse embryonic fibroblasts.

from PerkinElmer called Opera, which is equipped with its own image analysis software, because he felt it offers “the best balance” between high quality and high throughput.

“We are extracting hundreds of different features from single cells,” he says. He and his team compare the morphology of a cell in which a gene has been knocked down to a cell with the gene intact. They might then look at how the cells with knockdowns react to a drug. “If we don’t have high quality, it’s not going to help us,” he says, emphasizing that his work requires the “highest resolution possible.” The more accurately the parameters describe the shape of a cell or an organelle, the better the functional insight about the genomic processes that lead to the phenotype.

Peering through a microscope, the educated human eye automatically focuses on interesting phenotypes. “I kind of intuitively ignore the normal cells and can see very subtle differences,” Bakal says. “But we’re training the computers to do that and pick out the phenotypes.” The system applies automated ways to computationally screen for the most frequent phenotype. That view eliminates the unchanged cells from the analysis and highlights, for example, the remaining 30% that do show an atypical phenotype. Those changed cells can then be studied in greater detail to understand which genes are creating these changes.

Learning machine

Knowing that scientists such as Bakal and others do deep dives into cells, PerkinElmer is expanding its software



High-content analysis has been pharmaceutical-driven technology, a situation that is changing, says Jacob Tesdorpf.

tools, explains Jacob Tesdorpf, who directs the company’s high-content instruments and applications division. Inside the new version of the company’s image analysis software is an updated version of a tool called PhenoLOGIC, a machine-learning

functionality that lets users teach the instrument about imaged cells.

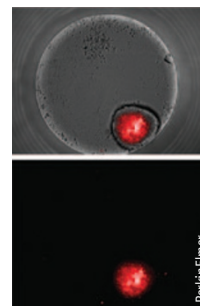
After a plate is imaged and later dis-

played on screen for analysis, a user can click on all the cells with a certain phenotype and, separately, click on cells that show specific changes. From the clicks, the software learns to categorize these two groups and can ascribe descriptors, such as fluorescence distribution or shapes, to these two cell classes.

Once this learning phase is completed, the tool creates an abstraction of how to tell these cells, referred to as groups A and B, apart. “The software can then automatically look at thousands of images and basically say, ‘These cells are A and these cells are B,’” Tesdorpf says, which gives scientists an assay. One type of cell might have a high probability of becoming metastatic. Being able to discern its phenotype helps researchers hunt for a way to undermine this ability.

PerkinElmer also adds to its instrumentation with a range of fluorescent agents for use in cell culture: for example, to measure activity of proteases involved in cancer. But, Tesdorpf explains, it is not enough to just offer these probes—the reagents have to be adapted for a high-content environment in new ways. One such probe, HypoxiSense, binds to biomarkers expressed during hypoxia, when cells inside a tumor are oxygen-starved. The hypoxic condition changes gene expression and might be partly responsible for the tumor’s ability to grow and spread.

Although imaging can be easier in two-dimensional cell culture, subtle changes in tumor cells can remain hidden; the changes are more apparent when the cells are grown in three-dimensional cultures, Tesdorpf explains. His company struck up an assay-development partnership with InSphero, which specializes in growing cells in three-dimensional spheres. Company investment in setting up assays with combinations of hardware and reagents is useful because it shows current and potential users new screening approaches, he says.



PerkinElmer adds to instrumentation with fluorescent agents. HypoxiSense is an indicator of oxygen-starved regions (red) in tumors.

High-content analysis has been a pharma-driven technology, but that situation has been changing. A “significant” amount, “if not the majority,” of the revenue comes from diverse academic sites, he says. In these labs, genome-wide screens are helping scientists to increase throughput and content yield as they screen for subtle effects in cells of interest, he says.

Several efforts are allowing scientists to perform analyses that yield more insight about single cells and that also add throughput, he says. For example, genome-wide siRNA libraries are available, as are probe-development programs such as the Molecular Libraries and Imaging Program (<http://mli.nih.gov/mli/>) from the US National Institutes of Health. The program’s organizers note that industry has built a capacity for high-throughput screening that has been lacking in the public sector. This government initiative offers academic researchers access to such large-scale screening instruments as they hunt for chemical probes to study gene function or signaling pathways. Other programs that bring more users to high-content screening are the ones devoted to rare diseases and conditions. Some labs in these programs have set up “fully fledged screening facilities,” Tesdorpf says.

Although PerkinElmer has traditionally aimed much of its technology at the pharma-based screening market, the boundaries between pharma and academia are blurring in certain ways, he says. Commercial drug developers are expanding their use of single cell analysis screening in ways that had been more common in academia. Academic labs, in turn, are actively translating work more toward drug discovery, pushing potential compounds of interest further along than before, and are therefore in more need of screening facilities to do so.

Other companies, too, such as GE Healthcare Life Sciences or Thermo Scientific, have instruments and software packages for image analysis and reagents for high-content analyses of single cells. Thermo Scientific has an ArrayScan instrument family, equipped with high-content analysis software and accompanying antibodies, dyes and reagents.

In addition to its high-content screening instrument ImageXpress, Molecular Devices sells software for image analysis. Some firms focus on just one aspect of high-content analysis, such as software or

services. Definiens offers image analysis software to analyze multiple parameters and statistically validate the readouts, for example. Imaging on a contract basis for academic and commercial customers, ImageIQ—a company spun out of the Cleveland Clinic—delivers large-scale quantitative data sets to scientists.

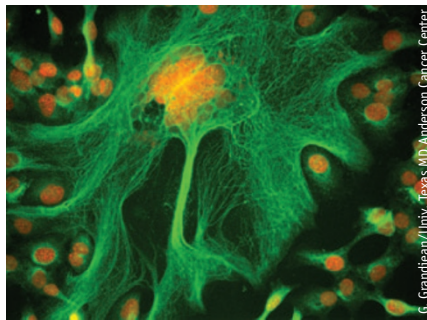
Raising flags

Nick Thomas, principal scientist in cell technologies at GE Healthcare Life Sciences, also sees a demand shift between academia, pharma and contract research organizations. The company’s high-content screening instrument, the IN Cell Analyzer, has traditionally been at home in pharma companies, but “the balance has shifted a little” over the last year in terms of customer base, he says. Interest is on the rise from contract research organizations. In addition, he says, “quite a reasonable proportion of these instruments” and their components are sold into academia. Buyers might be core facilities, or multiple investigators may pool grant money to purchase an instrument they will jointly use.

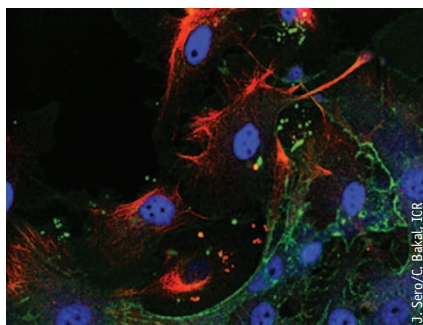


“Studying what’s going on in mitochondria is an ideal thing to tackle with high-content imaging,” says Nick Thomas.

As this shift happens, his company, too, is reacting by developing equipment for scientists and the assays to go along with them, including reagents, and software for analysis. GE Healthcare Life Sciences’ research and development group has previously engineered cells with GFP-fusion proteins



In human ovarian cancer cells, a cancer drug renders microtubules (green) rigid, hindering cell division.



Cells can take on many different shapes. This diversity makes it even harder to image proteins. Here, live epithelial cells are labeled with antibodies against the proteins E-cadherin (green) and vimentin (red). DNA is stained blue.

to be used as probes. “Some of those GFP sensors are in their own right extremely sensitive” and can be used to look at cell specifics, such as changes in cell cycle. “Those tools were developed to complement the power of the instrumentation,” he says.

This approach to add tools essentially integrates wetware and hardware. The process, involving staff biologists at the company who work with these tools and design them for use by biologists, is what sets his company apart from other vendors, he says. “It’s not just something that’s designed, specified and put together by engineers in isolation.”

Another area his group addresses is enabling drug developers and academics to look at changes in single cells over extended periods of time. Many drugs act on mitochondria, which can be particularly damaging to either heart cells or liver cells. Both organs have high energy requirements, needing cells with many mitochondria. Thomas says mitochondria are an ideal target for high-content imaging that allows an investigator to look at mitochondrial morphology or integrity over a large concentration range of a drug or an extended time frame of treatment. Experiments run on these instruments can take on a different breadth, he says. A drug that has an impact on mitochondria can be flagged early and removed from a development pipeline.

GE has two IN Cell models. One has a charge-coupled device (CCD) camera with a five-megapixel imaging chip, which allows users to survey a wide area with low-magnification objectives, maximizing the number of cells per image. “Because you are using a high-megapixel

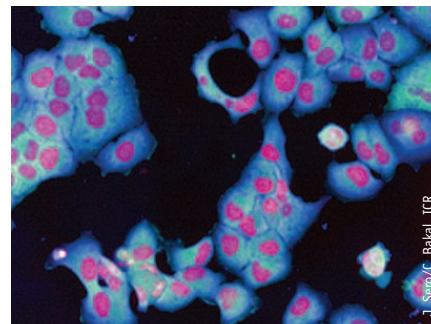
camera, you still have sufficient resolution for image analysis,” Thomas says.

The other model uses an optical engine built around a complementary metal-oxide semiconductor (CMOS) camera. One important change was added in reaction to requests from scientists seeking confocality without a dip in imaging speed. “Conventionally, what they were asking for was impossible,” he says. But with a CMOS detector, it is possible to read out individual pixels in coordination with laser-based line scanning, thus enabling confocality on demand. Alternatively, the camera can be run in wide-field mode for high-speed imaging, as with a conventional CCD system.

Thomas believes the arrangement now allows scientists to conduct their experiments more efficiently. With the way it rejects out-of-focus light, confocality delivers high resolution and improved contrast. Small and filamentous organelles, such as mitochondria, and the differences between one well and another, become sharper in a confocal image. Yet confocality takes time.

If a scientist using their CMOS-based instrument is running a four-color assay and needs confocality in only one channel—for example, to look at mitochondria in high detail—confocality can be limited to the red channel and switched off in the blue, green and orange channels. “The bottom line is, you save a hell of a lot of time,” he says. With this instrument, users have confocality only when needed, so “you don’t take a hit on imaging speed.” He believes this flexible feature differentiates the instrument in the marketplace.

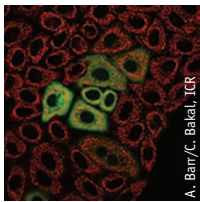
One academic site that has set up the IN Cell, first the previous model and now the newer one, is the Department of Experimental Therapeutics at the



Describing cell shapes requires accuracy, particularly for cancer cells. Here, live breast cancer cells are labeled with fluorescent dyes for DNA (magenta) and the cytosol (blue).

University of Texas MD Anderson Cancer Center. Geoffrey Grandjean landed a staff position there and helped set up a core service that performs high-throughput, high-content siRNA screening using the IN Cell. Seeing the data sets on a regular basis sparked his interest in a research career, and he is now a graduate student in experimental therapeutics.

For his own work, Grandjean's plan is to use the instrument to detect tumor cells on the basis of a metabolic trait. It has long been known that cancer cells use glycolysis to obtain energy, which is insensitive to the amount of oxygen around, rather than standard mitochondrial respiration.



A. Barr/C. Bakal, ICR

Each well in a 384-well plate delivers a fountain of results. Here, HeLa cells express a green probe in the endoplasmic reticulum and have their mitochondria labeled in red.

high-throughput capability and the wide viewing angle of the wells.

The confidence that his experiments will work stems from previous projects in which he was involved that used the GE imaging system. These studies looked at how microtubule stability was affected in cells after treatment with the cancer drug paclitaxel in a human ovarian cancer cell line².

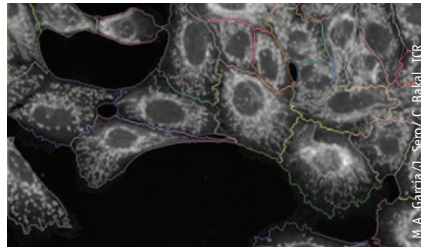


Image analysis software can automatically detect boundaries between cells, helping to extract features such as mitochondrial shape.

Image resolution was an important factor in this work, as was the software's ability to deliver quantified data on many parameters from each image, he says. The instrument is equipped with a small incubator allowing scientists to, for example, track the same cell as it is imaged once every 12 hours over the course of a week.

Experiments that knock down genes one at a time, set up in triplicate, allowing comparison of treated versus untreated cells, translate to thousands of wells on hundreds of plates. "You can't visually inspect everything manually well by well with hundreds of thousands of wells," Grandjean says. The data from automated experiments can give important clues about where to focus research next.

Cell bio curve ball

Although high-content screening of single cells offers a detailed view of many wells on a plate, cell biology presents mighty challenges. As Bakal explains, gene silencing never works across 100% of a cell population, thus complicating the interpretation of experimental results.

In four different cells, for example, silencing may work, not work or work only

partially in different ways. "You have these four bits of information which you have to consolidate and get a gene-specific signature for," he says. It is time consuming to find out what the knockdown of a gene is doing to cells as he and his colleagues pursue the question, "What is the real phenotype that happens when you knock down this gene?"

One important partner in this quest is computation. Without computational tools to analyze these large and rich data sets, the results are "never going to make sense to us," says Bakal. His institute has a strong computational infrastructure, but moving high-content data sets around for analysis is challenging, even for his team. "We don't have the right algorithms, and the computational power still is not right where we need it to be," he says. "The bottleneck is the analysis; it's not the imaging."

For junior scientists setting up a lab, automated high-content imaging experiments must include logistics as part of experimental design, Bakal says. Buying a high-content imaging system without thinking of the analysis that must follow can mean "you're all of a sudden going to end up with all this data, which you can't move anywhere and then can't analyze." With the logistics worked out, though, a high-content imaging system can help probe an increasing number of questions in cell biology.

1. Bakal, C., Aach, J., Church, G. & Perrimon, N. *Science* **22**, 1753–1756 (2007).
2. Ahmed, A.A. *et al. Cancer Res.* **71**, 5806–5817 (2011).

Vivien Marx is technology editor for *Nature* and *Nature Methods* (v.marx@us.nature.com).