When High Content Screening meets High Throughput

The terms 'High Content Imaging' (HCI) and 'High Throughput Screening' (HTS) were introduced more than a decade ago[1] and are defining the use of automated microscopy and automated image analysis in the context of drug discovery. Considered historically as two very separate disciplines with very few crossovers, this paper discusses whether you can ever do high-content imaging assays in high throughput.

By Dr Vincent Unterreiner and Dr Daniela Gabriel

Ever since the technology has evolved significantly to enable not only medium throughput assays for target identification or secondary screens, but also higher throughput assays compatible with primary hit identification using large compound libraries[2,3]. With the clear benefits of performing phenotypic cellular assays generating biologically relevant and multi-parametric data sets, the technology has established itself as a powerful tool for drug discovery. This evolution was partly made possible by the hardware improvements of the automated microscopes (eg auto-focus, plate/sample positioning enabling the use of high density formats) as well as the enhancement of the image analysis software enabling fast data extraction saving both time and costs in the screening process. In addition, innovations in the automation of plate preparation enabling the performance of non-homogenous assays in 1536-well plates (eg high density plate washers) were of great benefit for the establishment of the HCS technology for large scale screening campaigns. This article describes how we perform HCS with high throughput at the Lead Finding Platform of the Novartis Institutes of Biomedical Research (NIBR) explaining the benefits and challenges we are facing in primary hit finding.

Until recently, HCI and HTS have been considered as two separate worlds sharing only some borders. HCI is enabling multiplexed assays providing cellular or sub-cellular resolution and generating multivariate data sets. These assays can deliver different insights of the compounds' mode of action as well as their putative unspecific effects (eg toxicity). Due to their complexity, the HCI assays were often limited in throughput and were generally used to screen focused libraries or to perform secondary or counter screens. Homog-enous fully automated HTS assays are enabling fast data acquisition compatible with the screening of large compound collections. These assays generally deliver uni-variate data (eg cAMP accumulation, Ca^{++} release, protein production) and their resolution is limited to the well level[4]. High-throughput HCI assays are resulting from a combination of both worlds enabling fully automated primary screening of large compound collections with high resolution in multiplexed mode. To validate the use of imaging technology versus conventional assays, the variability and sensitivity of an imaging assay has been compared to a reporter gene assay (RGA) for the screening of inhibitors for PI3K – Akt – Foxo3A pathway[5]. Both assay formats were equally reproducible, with the HCS assay having a better statistical quality. In addition, the HCS assay was more sensitive than the RGA although no additional chemical scaf-

Figure 1
The automated plate preparation platform dedicated to immuno-staining protocols. A Catalyst 5 robot (Thermo Scientific) is integrated with 3 BNX1536 (Bionex), 2 Cytomat 5 (Thermo Scientific) for incubation at 37°C or 4°C as well as with a Teleshake (Thermo Scientific). The system is operated by the Momentum software

folds were identified as hits. While this study represents only one HCS and RGA assay format the outcome might change when comparing other pathways or assay setups.

Why enable HCS in a high-throughput format?

There are mainly four reasons for high-throughput high-content screens. The first and most obvious reason is to fill a gap. Many targets are not suited to be screened with biochemical or conventional cellular assays. HCS is expanding the field by enabling screens that used to be impossible with a decent throughput. Examples of these assays are the quantitative analysis of protein aggregation and granularity, as well as relocation events or morphological changes. The second reason is that in contrast to classical cellular assays (eg a reporter gene which provides an indirect readout that can be far downstream of the target), HCS assays are enabling a more focused readout on the target of interest (eg protein phosphorylation) in addition to the monitoring of the compounds' effect on the whole cell physiology (eg toxicity or morphological changes). Third, the use of multiplexing and sophisticated image analysis in HCS offers more information from the hits than traditional screens. For instance, multiple nodes of a cellular pathway can be measured already at the stage of primary screening (eg protein translocation triggered by a phosphorylation event), enhancing the content and quality of the derived hitlist. Fourth, the use of multi-parametric image and data analysis can reduce the rate of false positive hits lowering the need to perform counter screens to sort out compounds having unspecific effects. Image visualisation tools can support quality control to reduce the false positives as well as to identify assay artifacts that could lead to false negatives (eg absence of staining).

How are high-throughput HCI assays performed?

We started implementing HCS at the Lead Finding Platform of NIBR in Basel in May 2005. The first high-throughput HCS was carried out only recently in 2010. With our high-content instrumentation we support assays for multiple disease areas of NIBR, eg oncology, respiratory diseases, immunology, cardiovascular and infectious diseases. Depending on the readout type, the sensitivity and the statistical quality of the assay and taking into account the time and cost constraints, the team decides if the imaging technology can be used or if a conventional cell-based assay is better suited. The HCS assay formats and readouts are variable including nuclear translocation, protein phosphorylation, receptor internalisation, intracellular trafficking of proteins and virus infection. Not all of these assays are amenable to being tested with a million compounds due to technical or biological constraints such as plate format, incubation times, or cell line stability; therefore these parameters need to be assessed project specifically.

HTS process

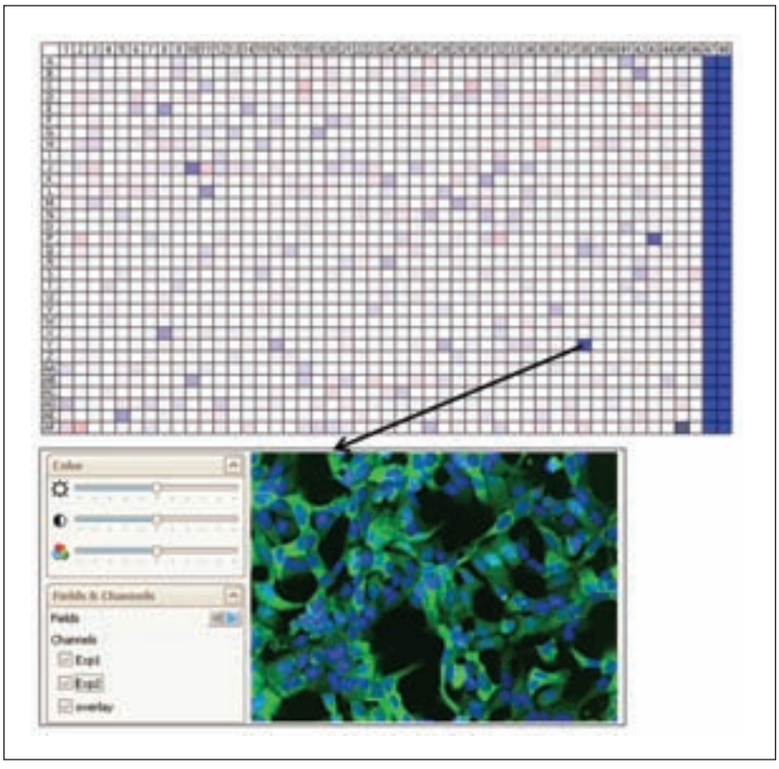
Generic processes for plate preparation, image acquisition and data handling need to be implemented to efficiently perform HCI in high throughput. The imaging time is generally the bottleneck compared to the plate preparation time which is faster in most cases. In order to reach the maximum throughput, flexibility in the use of the

imagers and automated plate preparation platforms can be introduced. To do so, HTS are conducted mainly with end point assays using fixed cells allowing the decoupling of the sample preparation from image acquisition, as well as the sharing of the plate preparation systems between various projects.

A typical process for a high-throughput HCI screen in 1536-well plates could be the following:

First, cells are cultivated and plated into 1536-well assay plates using the automated cell culture platform SelecT (TAP Biosystems). Then these assay plates are transferred on to a plate preparation platform (Agilent) to perform compound transfer with an Echo 550 (Labcyte) as well as reagent dispensing, incubation, fixation and washing. Once the cells are fixed, the complex and non-homogenous immuno-staining protocols are performed on a dedicated automated platform using the Catalyst 5 robot (Thermo Scientific) and designed especially for this purpose. The system is equipped with high density washer/dispenser BNX1536 (Bionex) as well as Cytomat incubators (Thermo Scientific) necessary for incubation at various temperatures and illumination conditions (**Figure 1**). The immuno-staining platform in NIBR was implemented to gain additional flexibility by decoupling the compound addition and fixation steps from the antibody staining process. Once plate preparation is completed, the plates are stored at 4°C until they can be measured on one of the imagers available in the screening unit.

Figure 2
Example of a 1536-well plate heat map displayed in the data analysis software. By clicking a well, the respective image is displayed



High-content imagers and image analysis

For an imager, the following features are important to perform high-throughput assays: ability to handle 1536-well plates and high-speed image acquisition (20-100 mins per 1536-well plate). Furthermore it is advantageous to perform image analysis in parallel to image acquisition (ie 'on the fly' analysis). Finally, the storage capacity of the instrument should be high enough to cope with terabytes of data, or alternatively be set up for automated data transfer to a dedicated database. The choice of the imager is dependent on two assay requirements: resolution and throughput. For assays requiring sub-cellular resolution and high throughput, our preferred imager is the Opera QEHS (PerkinElmer), a confocal imager equipped with four lasers and four CCD cameras allowing on-the-fly image analysis. For assays requiring high throughput based on fluorescence intensity measurements with no need for sub-cellular resolution, the Acumen eX3 (TTP LabTech), a plate scanning device equipped with three lasers allowing on-the-fly analysis of the fluorescence intensity distribution, is optimal. For medium-throughput HCI we generally utilise the IN Cell Analyzer 2000 (GE) which is a wide-field imager equipped with a large-chip CCD camera. It can handle 96, 384 and 1536-well plates; however with image analysis being decoupled from the image acquisition process. Performing the follow-up assays for primary hits identified with a laser scanning device such as Acumen by high resolution images from Opera or InCell200 can improve throughput tremendously. These three instruments complement each other well and the choice of the optimal imager as well as the image analysis software are key criteria to exploit at best the full potential of the technology for HTS.

A high-content imager usually creates a specific image format linked to metadata (eg channel, objective lens, pixel resolution etc) which makes the use of third party software more difficult since it might require an adaptation/conversion of the image or file format. This can slow down the screening process and increase the image storage space making the use of third party analysis software cumbersome. For HTS campaigns, the simplest solution is to use the image analysis software provided with the imager since no transfer of

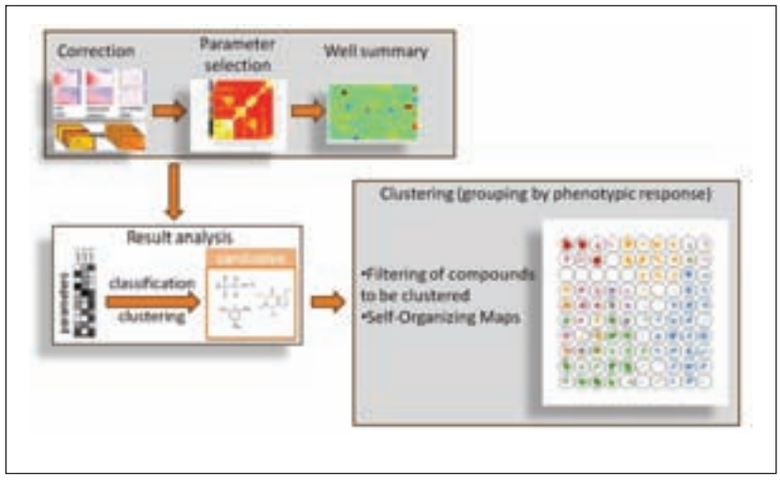


Figure 3
Analysis of multi-parametric HCS data. Data pre-processing consists of correction for plate pattern and plate effects, parameter selection by dimension reduction and well summary. The parameters are being further analysed by either classification or clustering. Visualisation of clustering results in self-organising maps helps to identify similar phenotypes

images or image conversion step is required. However, the most important is to use robust analysis scripts to account for plate-to-plate variation in intensity of staining. In addition, data analysis software for assay quality control in order to recognise quality issues as early as possible is needed. A special requirement for HCS is the link from data to images, allowing a prompt visualisation of the images to quickly identify staining issues or assay artifacts (**Figure 2**).

Multi-parametric data analysis

In HTS assays, a rather small number of project-specific readout parameters (<10) are collected. However, HCS is able to provide much more information with data sets containing readouts on multiple cellular parameters. To exploit the high content of images through multi-parametric data analysis, more sophisticated software tools are needed. Using our recently developed in-house software tool we are able to classify samples into hits or inactive based on a multitude of readouts[6]. Another analysis type can cluster sample responses into groups similar to control compounds or samples having similar phenotypes (**Figure 3**). Performing image analysis with algorithms generating many different parameters and the subsequent analysis of the data generated require a strong computational power. Generally this will not be performed on the computers delivered with the imagers, requiring images and data to be transferred to dedicated databases before analysis. These processes are time- and resource-consuming, which can restrict multi-parametric analysis to selected subsets (eg primary hits based on uni-variate readout). Compared to hit identification with uni-variate readout, we have recently observed a clear reduction of the number of false positives when applying multi-parametric image analysis (eg calculating the Mahalanobis distance to positive controls based on more than 100 parameters).

What are the challenges?

High-content screening in high throughput as described above is now well established in the lead finding department of NIBR, however some technical or process-related challenges still exist and need to be addressed. First, depending on the assay complexity and the imager chosen to perform the screening, the throughput figures can vary substantially (**Table 1**). One prominent factor impacting the throughput is the imaging time. Optimising the number of exposures, exposure time, magnification and number of images acquired per well can clearly influence the plate processing time and therefore the duration of the screening campaign.

Second, the throughput discrepancy between the plate preparation process and imaging can result in delayed quality control to detect errors in cell plating, antibody or compound distribution. Furthermore, in case the delay between plate preparation and imaging extends to several days, 1536-well plates with low volume bear the risk of evaporation. The use of specifically-designed

Assay readout	Imager	Plate format	Compounds screened	Colors	Images per well	Imaging time per plate (min)
Protein aggregation	Opera	1536	> 1 Million	3	2	100
Fluorescence intensity	Opera	1536	550 K	2	2	55
Nuclear translocation	Opera	1536	> 1 Million	2	2	25
Fluorescence intensity	Acumen	1536	> 1 Million	3	2	30
Granule detection	IN Cell 2000	384	20 K	3	3	60

Table 1: Examples of high-throughput HCI assays with different imaging requirements and their effect on the plate reading time

1536-well assay plates (Greiner) comprising a tightly closing lid can mitigate this risk[7]. Third, when using third party software, images have to be transferred to a specific database and converted to a specific format before the analysis. Considering the vast amount of data generated, this process needs to be run in parallel to imaging. Data management and data mining can be problematic when using different imagers or image analysis software packages. Based on the variety of image and metadata formats it can be a challenge to enable a link from image to result files. Software tools enabling the search of plates, compounds, or treatments combined with image and data visualisation for comparison of images and data generated by different imagers are still a major need in the field of HCS.

Summary and outlook

Despite the challenges mentioned, HCI is rendered possible for high throughput screening in primary hit finding campaigns using dedicated automated platforms able to deal with complex immuno-staining protocols in 1536-well formats. In order to do so, high-speed imagers allowing fast 'on the fly' image analysis combined with data analysis software and a dedicated IT infrastructure to manage and mine the wealth of data produced are required. The technology has proven to be mature for primary screening in drug discovery projects adding value by enabling novel assay formats that used to be impossible with conventional technologies. From now on, when the question is raised: Can you do high-content imaging assays in high throughput? The answer will be: "Yes, we can!" However, there is still room for improvement. The implementation of complex assays exploiting, for example, primary cells, live cell imaging, cell migration or 3D imaging in real high-throughput format, are still a challenge. The screening of a large number of plates for these assays will require substantial adaptations in cell culture, plate preparation, imaging devices and screening processes. Nevertheless, the benefit of physiologically relevant assays early in the hit identification process is increasingly acknowledged which augments the demand to implement phenotypic and disease relevant assays in higher throughput.

Acknowledgement

We would like to acknowledge G. Hofmann, Y. Ibig-Rehm, D. Siebert, M. Pfeifer, M. Goette, E. Schmidt and X. Zhang for fruitful discussions and contributions to HCS projects. E. Althof and F. Grandjean are thanked for automation support, A. Kümmel and P. Selzer for generation of the multi-parametric data analysis tool, J. Lin for image visualisation, H.P. Gubler, M. Schröder and I. Hossain for IT and P. Fürst for managerial support. DDW

Dr Vincent Unterreiner *is a Scientist II at the Novartis Institute for BioMedical Research where he is developing and performing high content imaging assays for drug discovery. He studied cell biology and pharmacology at the University Louis Pasteur in Strasbourg where he obtained his Master's degree in 2000.*

Dr Daniela Gabriel *has been heading up the Medium Throughput Screening group at the Novartis Institute for BioMedical Research since 2005. In 1998 she received her PhD in Biochemistry from the Max Planck Institute of Biochemistry in Martinsried, Germany.*

References

1 Taylor, DL, Woo, ES, Guiliano, KA (2001). Real-time molecular and cellular analysis: the new frontier of drug discovery. Curr. Opin. Biotech. 12, 75-81.

2 Xu, G, Mawji, I, Macrae, C, Koch, C, Datti, A, Wrana, J, Dennis, J, Schimmer, A (2008). A high-content chemical screen identifies ellipticine as a modulator of p53 nuclear localization. Apoptosis 13:413-422.

3 Ibig-Rehm, Y, Götte, M, Gabriel, D, Woodhal, D, Shea, A, Brown, NE, Compton, T and Feire, AL (2011). High-content screening to distinguish between attachment and post-attachment steps of human cytomegalovirus entry into fibroblasts and epithelial cells. Antiviral Res., Vol. 89, No. 3, pp.246-256.

4 Thomsen, W, Frazer, WJ, Unett, D (2005). Functional assays for screening GPCR targets. Curr. Opin. Biotech. 16, 655-665.

5 Unterreiner, V, Ibig-Rehm, Y, Simonen, M, Gubler, H and Gabriel, D (2009). Comparison of variability and sensitivity between nuclear translocation and luciferase reporter gene assays. J. Biomol. Screen., Vol.14, No. 1, pp. 59-65.

6 Kümmel, A, Selzer, P, Beibel, M, Gubler, H, Parker, CN and Gabriel, D (2011). Comparison of multivariate data analysis strategies for high-content screening. J. Biomol. Screen., Vol. 16, No. 3, pp. 338-347.

7 Pfeifer, MJ and Scheel, G (2009). Long-term storage of compound solutions for high-throughput screening by using a novel 1536 well microplate. J. Biomol. Screen., Vol. 14, No. 5, pp. 492-498.