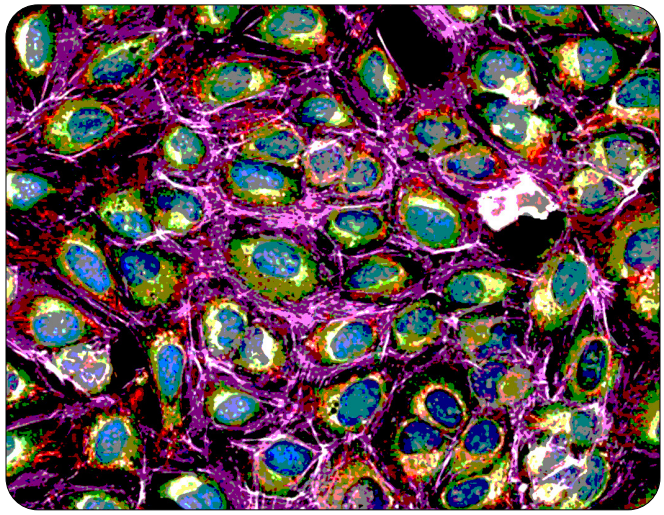
WHITE PAPER

Applying functional genomics strategies to high-content phenotypic screening

The demand for understanding human disease at the level of the gene has prompted the scientific community to adopt novel tools and techniques to address the inordinate challenge of drug discovery. Functional genomic screening is a powerful technique to explore gene-regulated responses and correlate cellular phenotypes to genome profiles in high-throughput screens. In this whitepaper we present key concepts, potential strategies, and best practice guidelines to tackle the challenges of functional genomic screening using high-content imaging (HCI) methods such as cell painting. Cell painting is a phenotypic high-content screening approach which combines cell and computational biology to unravel cells responses when subjected to perturbagens, such as chemical compounds, or gene expression modulators.



What is functional genomic screening?

To identify the cause of human diseases and develop possible treatments, it is critical to understand the underlying genes, proteins, and molecular pathways involved in disease etiology. Numerous approaches are available to rapidly interrogate gene function at the level of the genome, including systematic gene knockdown, knockout, or over-expression. These experimental approaches are collectively known as functional genomics, which has become a key discovery tool in many areas of biological research including target identification, drug resistance, host pathogen interactions, and biological pathway analysis.

Functional genomic screening can be used to determine gene function through the assessment of cellular phenotypes arising from genome-wide perturbations. Screens can be performed in high throughput, with libraries of small interfering RNA (siRNA), short hairpin RNA (shRNA), microRNA, or CRISPR-Cas9 guide RNA (gRNA). Typically, there are two main approaches to functional genomic screening: pooled and arrayed. The decision on what screening format to use is based on a number of considerations, including the assay and cell model being used as well as the read-out and process considerations like automation, high-content data capture, and powerful biostatistical analysis.

This paper will focus on arrayed screens, which are compatible with both binary and multiparametric assays and can be used with a variety of cell types.

Arrayed library screening involves targeting one gene per well across a multiwell plate format. Using this approach, it is possible to explore phenotypes that arise from a vast number of perturbations in parallel. Arrayed screens are

revvity

also amenable to HCI and more sophisticated assays. Using HCI, a wide range of cellular phenotypes can be visualized, including morphological features, protein trafficking, and post-translational protein modifications such as protein phosphorylation.

Arrayed screens are performed using cells grown in 96- or 384-well microwell plates. While experimental conditions vary depending on the biological question being explored, there are three fundamental steps to the workflow:

1. Library delivery
2. Phenotypic data acquisition
3. Data analysis

Briefly, library delivery can be achieved through various transfection methods (e.g., lipid transfection or nucleofection). Once the cells have been perturbed, they are stained either at low or high complexity, and image-based readouts are acquired using automated microscopy. Next, the imaging data is analyzed to explore any novel relationships between phenotypes and the perturbagens that induce them. Due to the size and complexity of the data generated from such screening approaches, high-performance computing and informatics-based systems are often required at this stage. Because there are many types of high-throughput readouts that can be used for screening, it is important to carefully consider the scope of the screen and which readout is most appropriate for the biological question being asked. Orthogonal readouts should also be planned in advance to ensure hits are validated and the most appropriate data captured.

Morphological profiling for unbiased phenotypic analysis

The data acquisition and analysis stages of phenotypic screening are key steps in the pursuit of biological discovery. Conventional high-content screening experiments typically focus on pre-defined, specific features with the aim to identify drugs or drug targets that modulate that phenotype. For such experiments, researchers have an a *priori* assumption of the phenotype of interest. Thus, most imaging experiments extract only one or two features of cells, meaning that vast quantities of quantitative data about cellular state are potentially left unexplored.

In contrast, morphological profiling experiments follow an unbiased approach to profile cells upon perturbations by extracting hundreds or even thousands of phenotypic measurements. Using this technique, various chemical or genetic perturbations can be compared to identify biologically relevant similarities, or clusters, in order to group compounds and/or genes into functional pathways. Phenotypic characteristics of each cell are measured in a process called feature extraction, which ultimately provides the raw data for profiling. The major features include staining intensities, textural patterns, size, shape of the labeled cellular structures, as well as correlations between stains across channels.

One image-based assay that can be used for morphological profiling is the cell painting assay.[1] This assay includes six fluorescent dyes used to label eight cellular components or organelles, which are then imaged in five channels. The aim is to 'paint' the cell as richly as possible with dyes, then use automated image analysis to extract hundreds or thousands of morphological features from each stained and imaged cell to produce profiles. These profiles are then compared and mined to address the biological question at hand.

For therapeutic intervention, functional genomic screening and morphological profiling approaches open up the possibility of detecting subtle phenotypes very powerfully. In contrast to target-based drug discovery, which is effective when a specific disease-causing protein or pathway has been identified, HCI and morphological profiling can be used to predict the biological roles of less-understood genes and pathways. This holds promise for diseases where numerous disease loci might be contributing to morphological effects, or those presenting with multiple phenotypes.

Informatics approaches to analyze big data

The handling and analysis of data is one of the greatest challenges for groups embarking on high-content screening ventures. This not only encompasses the final imaging data, but also the data produced throughout the entire workflow; for example, annotations on perturbagens or cell line concentrations. Downstream analysis then involves interpreting and validating any patterns in the morphological profiles.

As an example, screening experiments using cell painting produce vast amounts of data – typically around 100GB per 384-well plate, equating to approximately 1TB of data per day if 10 plates are screened. Data analysis therefore requires complex automated methods such as machine learning and deep learning to extract meaningful information from the data, particularly when working with more complex cellular systems.

For large datasets, clustering algorithms are commonly used to cluster features based on their similarity. To facilitate this, many laboratories have invested in data storage solutions to enable the long-term storage of their experimental data. This is not only necessary from a regulatory perspective, but also enables researchers to return to old datasets with new biological questions.

Conclusion

Functional genomics screening is a powerful technique to assess cellular phenotypes arising from genetic perturbations. It is useful for providing insight into the genetic underpinnings of biological pathways and the disease state of interest. In this whitepaper, we have discussed how to address some of the challenges of functional genomic screening using high-content imaging (HCI) methods such as cell painting. As technologies continue to advance and more functionally complex solutions become available, so too will our ability to view the data from these screens in higher dimensions to ensure that reproducible and robust results are achieved.

Reference

1. Bray, MA., Singh, S., Han, H. et al. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat Protoc* 11, 1757–1774 (2016). https://doi.org/10.1038/nprot.2016.105

This whitepaper was based on a roundtable discussion featuring Justin Boyd (Director, Translational Sciences, Vaxxinity), Martin Daffertshofer (Global Product Manager - High Content Screening Software, Revvity), Kaylene J Simpson (Associate Professor Head, Victorian Centre for Functional Genomics, Peter MacCallum Cancer Centre), and Steven Haney (Executive Director, Head of Oncology Research, Valo Health).

Revvity, Inc.
940 Winter Street
Waltham, MA 02451 USA

(800) 762-4000
www.revvity.com