Leveraging the Power of High Content Screening Using Iterative Data Mining for Ethnically-diverse, Predictive, Animal-free Hepatotoxicity Platform.

P15-07

Victor S.C. Wong[1], Adeel Shafi[2], Kate Cameron[2], David A. Egan[1]

[1]Core Life Analytics B.V., 's Hertogenbosch, The Netherland;, [2]Cytochroma, Roslin, Scotland

Introduction

Multiparametric high-content screening (HCS) approaches, such as the Cell Painting assay, are growing in popularity for many applications including, but not limited to, *in vitro* pharmacology and toxicology. HCS combines rapid automation and imaging of fluorescent markers to capture unique cellular phenotypes. This is particularly useful for gaining insights into mechanism-of-action by comparing the phenotypic profiles of various compounds with reference drugs or controls. Cytochroma uses HCS approaches to provide *in vitro* toxicology services without the use of animal models. However, HCS approaches precipitated a growing problem where the capacity for data analysis was overwhelmed by the speed at which the data was being generated. Here, we present a workflow in which a web-based platform called StratoMineRTM was used to accelerate the analysis of Cytochroma's RAPIDToxTM platform, a human induced pluripotent stem cell derived hepatocyte toxicity assay. We tested a library of 160+ compounds in quadruplicate at multiple concentrations ranging from 0.75 µM to 200 µM. 48 hours after drug induction, an established Cell Painting protocol was used for direct visualisation and segmentation using fluorescent markers (i.e.,. Hoechst 33342, Syto14 (CellMask), Phalloidin/ AlexaFluor and Mitotracker Deep Red). Images were acquired using the Perkin Elmer High Content Opera Phenix Imaging System, and analysed by Harmony software v4.9, extracting > 350 features based on intensity, texture, size, and morphology. The output of the image quantification was uploaded into StratoMineRTM for rapid data mining and analysis. Moreover, StratoMineRTM also allowed data normalization and transformation, data reduction and unsupervised hit selection. These functionalities allowed Cytochroma to identify a hit rate of 59.84% with 1149 hits that gave significant phenotypes relative to the controls. Using phenotypic distance score ranking from the hit selection, Cytochroma can identify key compounds of interest as well as plotting dose response curves and IC values (e.g., Acetaminophen at 25uM). StratoMineRTM also offers clustering analysis to find similar/dissimilar mechanisms of action. These results demonstrate that the proposed workflow allows combining iPSC technologies, HCS and StratoMineRTM to yield robust and reproducible results for predictive *in vitro* hepatotoxicity.

Approach

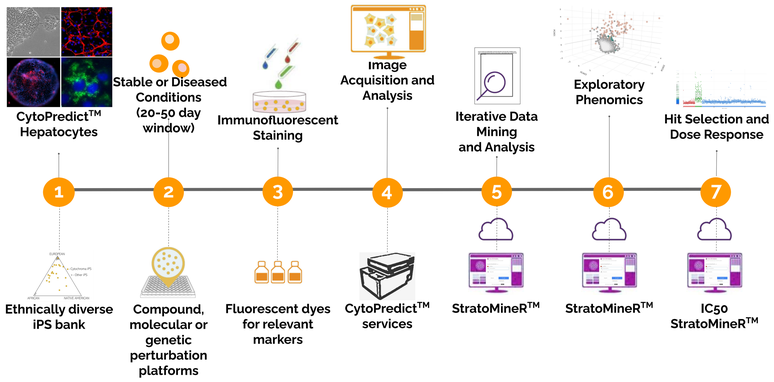

Figure 1. Overview of the predictive toxicology analytical workflow. Cytochroma identifies heterogeneous drug responses in the global population by using its proprietary cell-based ethnically diverse iPSC bank. After automated culturing and maturation of iPSC-derived liver models in multi-well plates (1), they are subjected to compound screening (2). An innovative tagging technology using relevant and specific markers enables direct visualisation (3). High throughput and high resolution automated image acquisition (4) of the liver models is carried out using Perkin Elmer's Opera Phenix system. Quantification of the images are subsequently analysed using Harmony software with advanced segmentation capabilities (5). The high content output from image analysis contains multi-parametric data and is then further analysed using StratoMineR™ (6) developed by Core Life Analytics. This software platform allows for rapid high content data mining and analytics that Cytochroma uses to discover unique phenotypic signatures as well as novel hepatotoxic hits (7). StratoMineR™ also allows for the use of machine learning to handle growing multi-parametric data for upscaling of the high content screening (HCS) platform. This will allow us to actively engage in a data-driven iterative loop to rapidly find novel hepatotoxic compounds, especially as Cytochroma scales up.

Results

Cytochroma's monolayer liver models are differentiated, mature and functional, making them ideal for toxicity testing.

To date, more than 160 compounds have been tested in quadruplicates at multiple concentrations ranging from 0.75 µM to 200 µM. 48 hours after drug induction, an established Cell Paint method is used for direct visualisation using fluorescent tags. Hoechst 33342, Syto14 (CellMask), Phalloidin/ AlexaFluor and Mitotracker Deep Red are used.

Images were acquired using the Perkin Elmer High Content Opera Phenix Imaging System, and analysed by Harmony software v4.9, evaluating > 350 features based on intensity, texture, size, and morphology.

The output of the Image quantification is uploaded into StratoMineR™ for HC analysis (see Workflow)

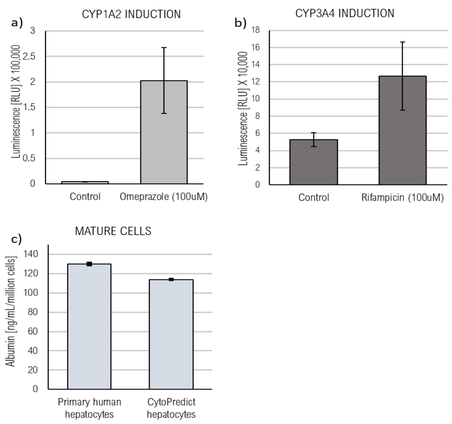

Figure 2. Induction activity of a) CYP1A2 and b) CYP3A4, which are upregulated in response to certain drugs, shows the liver models have the essential function for accurate testing. c) Cytochroma's hepatocytes show maturity and express similar levels to current gold standard primary human hepatocytes.

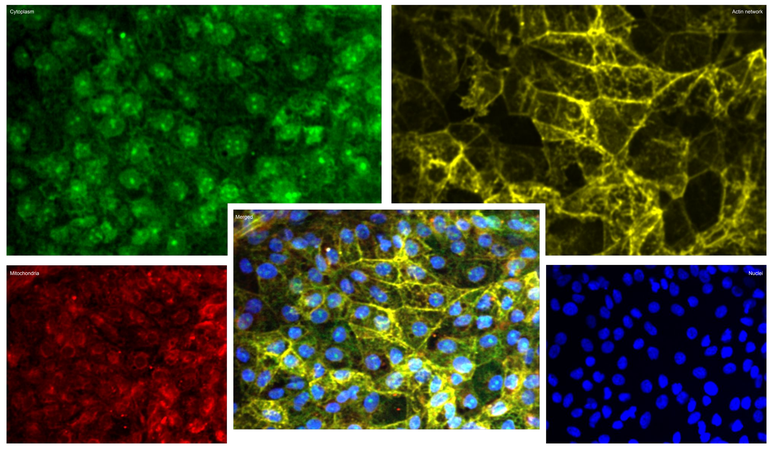

Figure 3. Images on the right are from control wells during Cell Paint toxicity assay. Images were acquired using the Opera Phenix Imaging System. Example of each channel is shown, with a composite image in the middle of nuclei, cytoplasm, actin network and mitochondria.

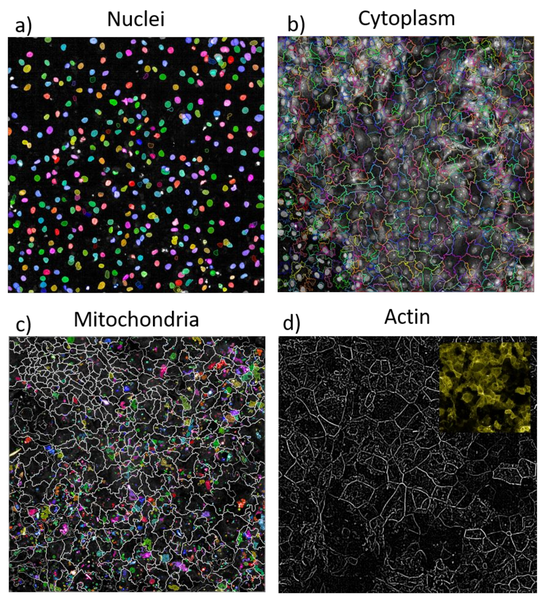

Figure 4. Feature extraction and segmentation using Harmony 4.9 software for calculating STAR morphology. Example, axial and radial distribution of mitochondria were measured, as well as symmetry, profile, and threshold compactness. a) Shows the nuclei segmentation using optimised thresholding parameters. b) Cell outline, calculated using Syto14 and Phalloidin/Alexa Fluor tags. c) Segmentation process for calculating mitochondria within defined cells. d) Using STAR morphology function to calculate key texture properties from actin staining, such as 'ridges'. The insert top right shows the original image.

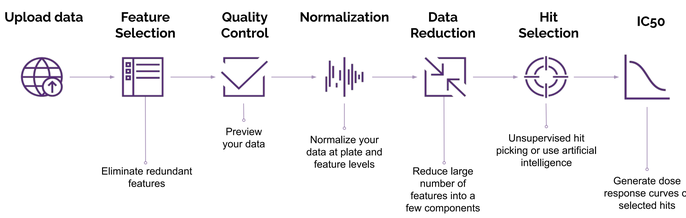

Figure 5. StratoMineR™ is a web-based platform which guides users through a typical workflow in analysis of high content multi-parametric data. Starting with data upload in .csv or .txt formats, subsequent steps in StratoMineR™ allow rapid data mining and analyses. For example, Feature Selection step removes redundant and/or problematic measurements. Quality control allows plate map assignment with control and sample information for the raw data. Normalization takes into account of, and corrects, plate-to-plate variation, heavily skewed measurements, and feature scaling. Data reduction can reduce a large number of measurements into components, and Hit Selection allows calculation of Euclidean distance score relative to a reagent class of interest (e.g., Negative control). Distance scores of hits can then be used to calculate dose response curves in IC50 feature of StratoMineR™.

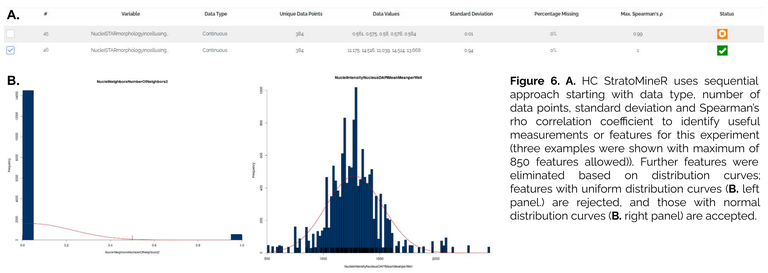

Figure 6. A. HC StratoMineR uses sequential approach starting with data type, number of data points, standard deviation and Spearman's rho correlation coefficient to identify useful measurements or features for this experiment (three examples were shown with maximum of 850 features allowed)). Further features were eliminated based on distribution curves; features with uniform distribution curves (**B.** left panel.) are rejected, and those with normal distribution curves (**B.** right panel) are accepted.

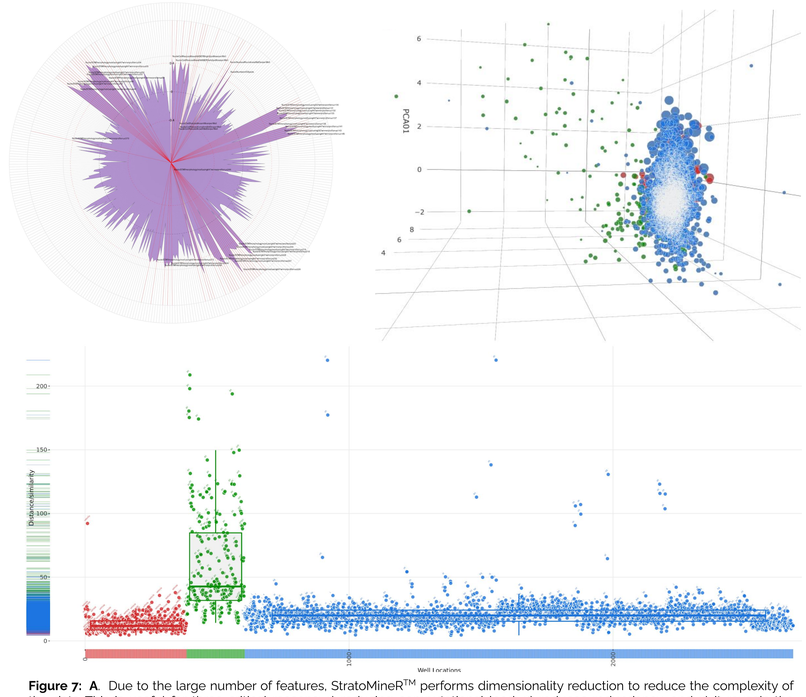

Figure 7: A. Due to the large number of features, StratoMineR™ performs dimensionality reduction to reduce the complexity of the data. This is useful for three critical reasons: 1) reducing computational load, 2) reduces redundancy, and 3) it reveals the biology behind the data by highlighting important features. StratoMineR can also suggest the number of principal components to use, in this case 8 PCAs). A spider plot showing one of the principal components, with features that have significant loadings on the principal component. **B.** Visualizing data points in 3D in relation to 3 components or features is particularly useful as it helps Cytochroma to identify interactions between data points in relation to each other. This is useful to identify samples (blue dots) that clustered with positive controls (green dots) or negative controls (red dots), and may therefore have similar MOA. **C.** Unsupervised hit selection calculates Euclidean distance scores for all wells calculated from the median of the negative controls with $p < 0.05$. Hit rate of 59.84% with 1149 hits.

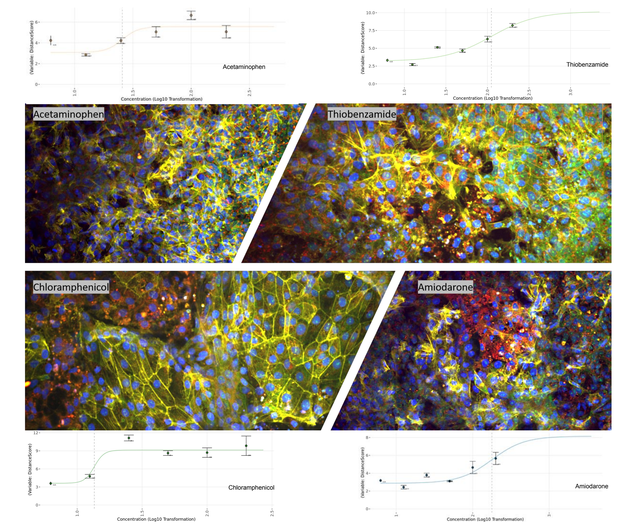

Figure 8: Using phenotypic distance score ranking from the hit selection step, Cytochroma can identify key compounds of interest as well as plot dose response curves for each compound. The Y-axis represents the phenotypic distance taking into account all eight principal components. The x-axis represents the concentration. IC values (uM): Acetaminophen = 25, Thiobenzamide = 111, Chloramphenicol = 13.5, Amiodarone = 178.

Conclusion

We show a streamlined workflow for high-throughput toxicity screening using StratoMineR™ as a tool for phenotypic screening. The unlimited source and scalable production of liver models, coupled with the automated, established, and scalable toxicity assay allows Cytochroma to screen multiple compounds in mature, polarized and stable models that are physiologically representative with enhanced sensitivity. StratoMineR™ is a powerful and intelligent data analytics tool that allows Cytochroma to perform deep analyses on phenotypic features across multiple experiments. The highly representative data outputs can reduce the failure rates that are currently seen going from preclinical to clinical development.

Acknowledgements

Cytochroma thanks the Beatson Institute for using the high-throughput Opera Phenix for acquiring images.

Cytochroma
Roslin Innovation Centre
Roslin, EH25 9RG
Scotland
adeel@cytochroma.org

Core Life Analytics
Sint Janssingel 88
5211 DA, 's-Hertogenbosch
The Netherlands
sales@corelifeanalytics.com