Introduction to ChromaLive

Martin Cottet1,*, Ibrahim Bilem1, Karine Audette2, Simon Mathien2, Victor Wong3, Félix Pérusse-Lavoie1, David W. Andrews4

1: Saguaro Technologies, Quebec, Canada 2: High-throughput screening platform, IRIC, Université de Montréal, Montreal, Canada 3: Core Life Analytics, 's-Hertogenbosch, Netherlands 4: Biological Sciences, Sunnybrook Research Institute, Toronto, Canada *: For correspondence: Martin Cottet, mcottet@saguarobio.com

Phenotypic profiling approaches using **high-content screening** and **cell painting techniques** are becoming more central to successful drug discovery efforts. Here, we introduce **ChromaLive**, a novel non-toxic dye that labels live cells and provides unique phenotypic fingerprints for accurate compound profiling. By conducting a variety of in vitro assays with ChromaLive, we demonstrate its ability to effectively quantify disease-relevant phenotypes with ease and flexibility, and its potential for improving efficiency in compound profiling efforts. Specifically, we demonstrate the absence of cellular toxicity from ChromaLive, along with its simple mix-and read staining protocol, and a biologist-friendly analysis workflow that uses both supervised and unsupervised analytical approaches. These features of ChromaLive enable kinetic measurements of phenotypic changes in live cells at high-throughput, therefore offering a valuable tool for researchers to study dynamic cellular processes and the potential of new drug compounds

Introduction

High Content Screening in Drug Discovery

In the field of drug discovery, the traditional approach has involved identifying specific targets of interest and designing small molecules to modify their function (Hughes et al., 2011). Although, this target-based approach has led to successful clinical translation of numerous drugs and compounds, it has also encountered a growing attrition rate (Waring et al., 2015). In contrast, phenotypic drug discovery focuses on probing phenotypic changes in biological systems to identify effective compounds (Lee and Berg, 2013). In fact, phenotypic profiling approaches have emerged as a key component of the drug discovery process due to its capacity to capture the complexity of biological systems. Indeed, it can reveal biological scenarios other than those originally investigated with target-based approaches (Swinney et al., 2011; Vincent et al., 2022).

One central approach in phenotypic drug discovery is high-content screening (HCS), which involves the analysis of cellular images to identify and characterise phenotypes of interest. Unlike standard high-throughput screens that provide only averaged information over a cell population, HCS captures more detailed and comprehensive data. This method has been made possible by the development of automated imaging systems and image analysis tools driven by artificial intelligence (AI) (Bickle, 2010; Danuser, 2011). By performing HCS with perturbing agents like small molecules or gene-editing approaches, researchers have already demonstrated its potential in identifying new therapeutic treatments (Oppermann et al., 2016; Lin et al., 2020)

In addition to automated imaging systems and AI-driven analysis, the development of the popular Cell Painting technique, which involves labeling specific components of the cell (Bray et al., 2016), has also been crucial to the success of HCS. However, Cell Painting requires fixation of cells and multiple staining and washing steps. This process can introduce various challenges, including increased experimental design complexity and the risk of disrupting samples. Moreover, as an end-point assay, it may overlook valuable kinetic information and overlook high-potential compounds.

We believe that the development of a "mix-and-read" stain for live cells can address these limitations, simplify screening procedures, and provide new insights into cellular physiology by enabling kinetic measurements throughout the assay.

ChromaLive for simpler, live-cell HCS

In this article, we introduce ChromaLive dye, a revolutionary multi-chromatic dye that stains cellular components of live cells in a non-specific manner (Figure 1b). Our goal with this product is twofold: to ensure minimal disruption of cellular physiology and to enhance the visualization of diverse cellular features for precise phenotypic profiling. ChromaLive represents a groundbreaking "mix-and-read" fluorescent stain. The use of ChromaLive has several advantages in live cell research. Firstly, its non-toxic nature and minimal impact on live cells make it an ideal choice for cell labeling and tracking cellular 2 phenotypes over prolonged cell cultures. Moreover, the unique staining and spectral properties of this dye allow for simultaneous highlighting of distinct cell states, making it exceptionally suitable for comprehensive phenotypic profiling in live cells.

The data presented in this article highlights the non-toxic nature of ChromaLive and its compatibility with prolonged phenotypic studies on living cells. Additionally, we provide insights into information extraction through automated imaging and data analysis, enabling robust phenotypic profiling. We compare compound testing and phenotypic profiling using different analytical approaches to showcase the advantages of ChromaLive and to leverage its unique spectral

properties. Lastly, we highlight the ability of ChromaLive to facilitate kinetic measurements of phenotypic changes in live cells.

ChromaLive, a first in class non-toxic dye

Non-toxic in cell cultures

To evaluate the impact of ChromaLive on cell cultures, multiple approaches were employed to compare cells grown in the presence of ChromaLive with those grown in standard medium. Initially, a cell counting experiment was conducted using various cell lines cultured for six days (Figure 2a). Remarkably, the presence of ChromaLive in the medium did not exhibit any discernible influence on cell proliferation. Furthermore, cell viability was assessed over a 48-hour period using RealTime-Glo™ (Promega) with different concentrations of ChromaLive (Figure 2b), and results were consistent with cell counting data, showing no significant differences between conditions

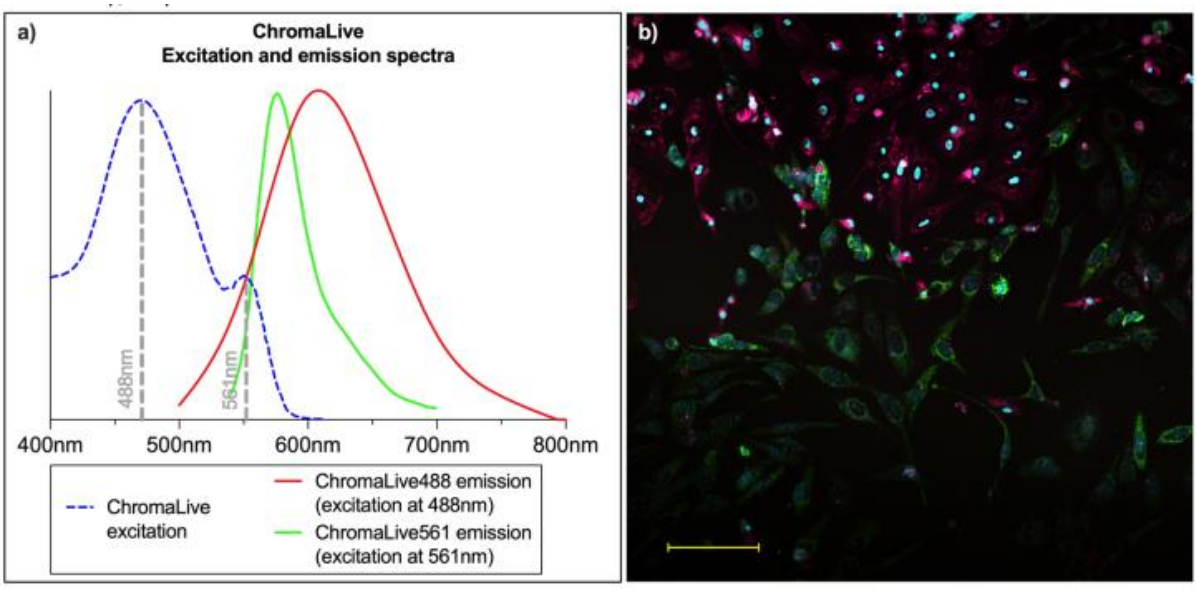

Figure 1. ChromaLive multi-chromatic imaging. a) Excitation and emission spectra of ChromaLive. Of note: ChromaLive is excited at 488nm and 561nm, with different resulting emission spectra (in green, ChromaLive561: emission spectrum when excited around 561nm, in red, ChromaLive488: emission spectrum when excited around 488nm). b) Differentiated HepaRG cells, cultured for 72h in presence of ChromaLive and stained with Hoechst 33342.
Live cells imaged on a Zeiss LSM800 confocal microscope (magnification 20x), in a "mix-and read" format (cells in culture medium, with ChromaLive and Hoechst present). (Red: ChromaLive488-Yellow, Magenta: ChromaLive488_Red, Green: ChromaLive561, Cyan: Hoechst. Scale bar: 100nm

Minimal impact on gene expression patterns

To gain further insights into the potential physiological impact of ChromaLive, an extensive RNA sequencing assay was performed. In this assay, we compared complete transcriptomes of MCF-7 breast cancer cells grown in medium containing a control amount of DMSO (0.1%), ChromaLive, or commercial dyes such as DRAQ5 and Green CMFDA (CellTracker™ Green CMFDA). Principal component analysis (PCA) (Figure 2c) and volcano plots (Figure 2d) were utilised to analyse and visualise the data. Notably, the results demonstrate that ChromaLive has a very limited impact on gene expression over 24 hours, compared to DRAQ5 (known to exhibit a certain level of cytotoxicity as a DNA intercalator (Mari et al., 2010; Richard et al., 2011), and even CellTrackerTM Green CMFDA, commonly regarded as a non-toxic dye (Zhang et al., 1992). For teams interested in more detailed information, the complete sequencing data are available upon request (see contact information).

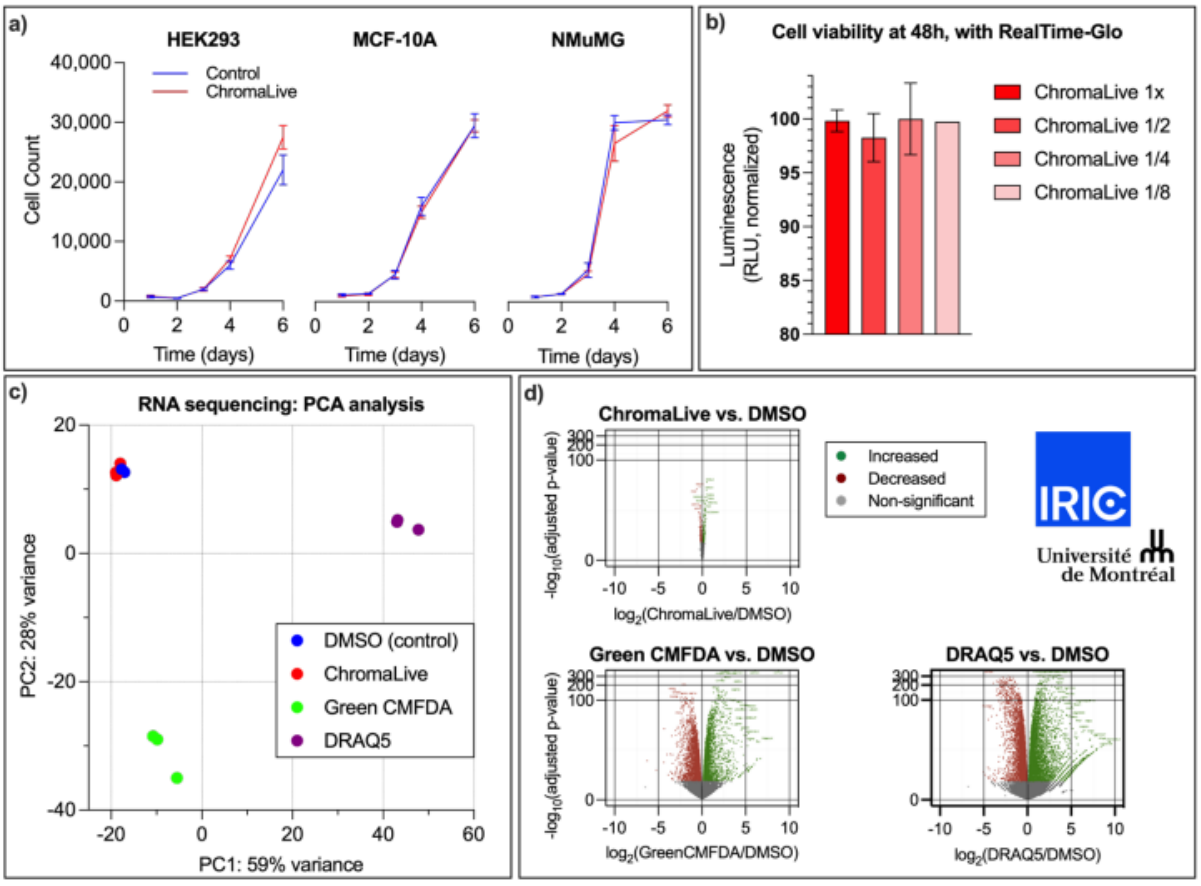

Figure 2. Assessing ChromaLive's non-toxicity.
a) Cells were plated at a density of 1,000 cells / well in a 384 well plate, in presence of ChromaLive. Adherent cells were additionally labeled for 30 minutes with 1µM Hoechst 33342. Cells were imaged and counted on an Opera Phenix at 20x magnification.
b) MCF-7 breast cancer cells were plated at 1,000 cells / well in a 96-well plate and cultivated with

the indicated dilution of ChromaLive for 48h. Cell proliferation was assessed with the RealTime Glo assay from Promega.
c) Principal component analysis (PCA) representation of RNAseq assay. MCF-7 cells were plated at 500-1,000 cells / well in a 96-well plate and grown for 24h at 37°C, 5% CO2. Medium was then changed for medium corresponding to experimental conditions (control: DMSO 0.1%, ChromaLive (1x), Green CMFDA (10µM) and DRAQ5 (5µM)) and cells were incubated for a further 24h. Cells were then harvested to extract RNA and samples were analysed using the Illumina Nextseq500.
d) Volcano plot representations of genes exhibiting a significant increase (green) or decrease (red) in their expression as compared to the control, in presence of ChromaLive (left graph), Green CMFDA (middle graph) or DRAQ5 (right graph). Genes displayed in red or green have an adjusted p-value lower than 0.05

Imaging ChromaLive

Spectral properties and instrument requirements

ChromaLive is a unique **multi-chromatic dye** with multiple excitation and emission wavelengths, offering exceptional capabilities for high-content screening (HCS). Acquiring images in those different emission channels enable the collection of different intensity profiles, along with morphological and texture features that, altogether, ensure accurate differentiation of cellular states or cellular types. Specifically, it is recommended to image ChromaLive in minimally two different channels, and optionally three. These two channels are ChromaLive561, and either ChromaLive488_Yellow or ChromaLive488_Red. Figure 1a illustrates the fluorescence excitation and emission spectra of ChromaLive. The option to select between ChromaLive488_Yellow and ChromaLive488_Red is because of their high degree of similarity, which is sufficient for differentiating between cellular phenotypes, even when phenotypes only have subtle differences. However, these two channels can still provide slightly different information, hence the option to use both ChromaLive488_Yellow and ChromaLive488_Red (refer to Table 1). In addition, the large Stokes shift observed when exciting the dye at 488nm allows multiplexing with additional green dyes (excited at 488nm and imaged around 520nm), as well as blue dyes (e.g., Hoechst) or red dyes excited at 647nm (e.g., DRAQ5). However, it is important to carefully design multiplexed fluorescence experiments and verify fluorescence bleed-through, as it can vary depending on the imaging equipment and settings used for image acquisition (see Table 1).

In addition, the large Stokes shift observed when exciting the dye at 488nm allows multiplexing with additional green dyes (excited at 488nm and imaged around 520nm), as well as blue dyes (e.g., Hoechst) or red dyes excited at 647nm (e.g., DRAQ5). However, it is important to carefully design multiplexed fluorescence experiments and verify fluorescence bleed-through, as it can vary depending on the imaging equipment used for image acquisition (see Table 1).

Channel	Excitation wavelength	Emission wavelength
ChromaLive488-Yellow	488nm	550nm - 630nm
ChromaLive488-Red	488nm	630nm - 750nm
ChromaLive561	561nm	575nm - 630nm
Blue dye (e.g. Hoechst)	*405nm (370nm - 440nm)*	*410nm - 470nm*
Red dye (e.g. DRAQ5)	*647nm*	*660nm - 750nm*
Green dye	*488nm*	*500nm - 535nm*

Table 1: Recommended excitation and emission settings for ChromaLive and additional stains. This is specific to every imaging instrument and proper testing is recommended before adding additional stains as well as to define standard imaging settings.

In addition to its multi-chromatic properties, ChromaLive is non-fluorescent in the cell culture medium and exhibits fluorescence only upon entering cells. Inside the cell, ChromaLive differently stains multiple cellular structures and compartments and is highly sensitive to environmental and physiological changes within the cell, resulting in quantifiable spectral changes. These features necessitate multiple wavelength excitations and emissions, ensuring the acquisition of multi-parametric information crucial for precise phenotypic profiling. As an example, Figure 1b depicts a typical image of differentiated HepaRG cells cultured for 72 hours in the presence of ChromaLive, along with additional staining using Hoechst 33342. To minimise its impact on cell physiology, Hoechst 33342 was added at a concentration of 100 ng/mL, three hours prior to imaging.

As an example, Figure 1b depicts a typical image of differentiated HepaRG cells cultured for 72 hours in the presence of ChromaLive, along with additional staining using Hoechst 33342. To minimise its impact on cell physiology, Hoechst 33342 was added at a concentration of 100 ng/mL, three hours prior to imaging

Cell culture with ChromaLive

ChromaLive is a highly efficient solution designed to optimise cell culture and imaging processes. Unlike other additives, ChromaLive is non-toxic and does not require additional washing steps, making it an ideal choice for continuous use throughout various stages of cell culture. To illustrate its effectiveness, we provide an experimental design for imaging MCF-7 breast cancer cells in the presence of increasing doses of a compound of interest.

Day 0: Cell Seeding in ChromaLive-Complemented Medium. Before starting, allow the provided ChromaLive solution (1,000x in DMSO) to stabilise at room temperature, ensuring there is no condensation. Afterward, gently spin the tube to collect all the sample at the bottom. Next, dilute 10µL of ChromaLive (1,000x) in 10mL of RPMI 1640 supplemented with 10% FBS and 1% PenStrep, mixing thoroughly. This will serve as the standard culture medium for the entire assay. Finally, seed MCF-7 cells at a density of 7,000 cells per well (100µL per well) in a clear bottom, black 96-well plate using the prepared culture medium.

Day 1: Treating with Compounds of Interest. Prepare a dose-response of the compound of interest. As an example, staurosporine was sequentially diluted from 50µM to 5pM with 10-fold dilutions. Add 11µL of the prepared compounds to the wells, resulting in final concentrations ranging from 5µM to 0.5pM. It is important to equalise DMSO content across conditions. For example, adjust to 0.2% DMSO (0.1% DMSO from ChromaLive and 0.1% DMSO from the compound of interest). Imaging can be started immediately and repeated at the desired time points throughout the experiment to gather kinetic information. Of note, imaging was performed over a period of 72 hours (3 days) in assays presented in the following sections.

It is also possible to add a nuclear stain to the experiment at this stage (first day of imaging). For instance, Hoechst 33342 can be diluted to 1µg/mL in culture medium, and 12.5µL of this solution can be added per well to achieve a final concentration of 100ng/mL. Incubate at least 3 hours at 37°C before imaging, for sufficient staining at this low concentration. (In this case, increase compound volume to 12.5µL per well, for a 1:10 dilution in a final volume of 125µL per well).

Image analysis and feature extraction

High-content screening is also referred to as highcontent analysis as the precision of the phenotypic screen is directly related to the richness of the information extracted from the images and the quality of the data analysis (Simm et al., 2018). Performing image analysis and extracting information from these images (often referred to as feature extraction) can be achieved by any image analysis software. Microscopy and HCS equipment manufacturers often offer their own software with their imaging equipment. In the next use-cases, we will be mostly relying on CellProfiler, a widely used and freely available software (Stirling et al., 2021). CellProfiler simplifies image analysis by automating the process and enables the extraction of multiple features, which can be stored for later downstream data analysis. To learn more about CellProfiler, visit https://cellprofiler.org.

CellProfiler allows users to create custom-built image analysis pipelines tailored to their specific requirements (Figure 3). A typical pipeline consists of three parts: input modules, analysis

modules, and saving/exporting modules to store images and extracted features for future reference and analysis.

The first step in the pipeline is to load images by dragging and dropping them into the Images module. Images or image folders can be conveniently added using the "Drop files and folders here" pane. The Metadata module enables users to connect relevant information to the imported image files. This information can be extracted from the image filenames or folder names. Properly setting up file naming conventions in advance simplifies this process. Another option is to link a metadata file to the imported images, which can contain information about cell type, treatment, compounds added, and the dosage for each condition. The NamesAndTypes module is where the real magic happens! By assigning rules based on the extracted metadata, users can assign specific channels to each image, creating image sets that contain all the imaged channels. This capability proves extremely valuable in subsequent analysis steps, as each channel can be used to extract specific features from the image sets. Lastly, the Groups module allows for dataset splitting into independent groups. This feature comes in handy when dealing with larger datasets or when performing time-lapse imaging, as it enables the analysis of each time point or group separately.

Figure 3 provides a preview of the analysis modules that can be added to the pipeline's analysis section through the Adjust modules using the "+" button. Notably, the Identify Primary Objects and Identify Secondary Objects modules deserve special attention. These modules perform image segmentation, allowing users to create custom-named objects like "Nucleus" or "Cell". Specific features can then be extracted from these objects. Standard metrics such as image and object intensity, as well as shape and size information, can be extracted. Moreover, more subtle changes in images, such as texture or intensity distribution, can provide valuable insights into cell phenotypes.

Finally, it is essential to include modules that save generated images and export extracted features to a spreadsheet or database. Spreadsheets are useful for smaller datasets or when sharing results with third-party data analysis solutions. Exporting to a database, along with its properties file, is crucial for compatibility with CellProfiler Analyst software, which takes over the data analysis phase. An important feature of CellProfiler is the presence of help ("?") buttons throughout the interface, providing additional information and examples for each module and its settings. These resources are invaluable for users to familiarise themselves with the software and its capabilities. When adding analysis modules, it is crucial to extract as many relevant features as possible, as this directly impacts the precision and validity of the multiparametric analysis for phenotypic screening.

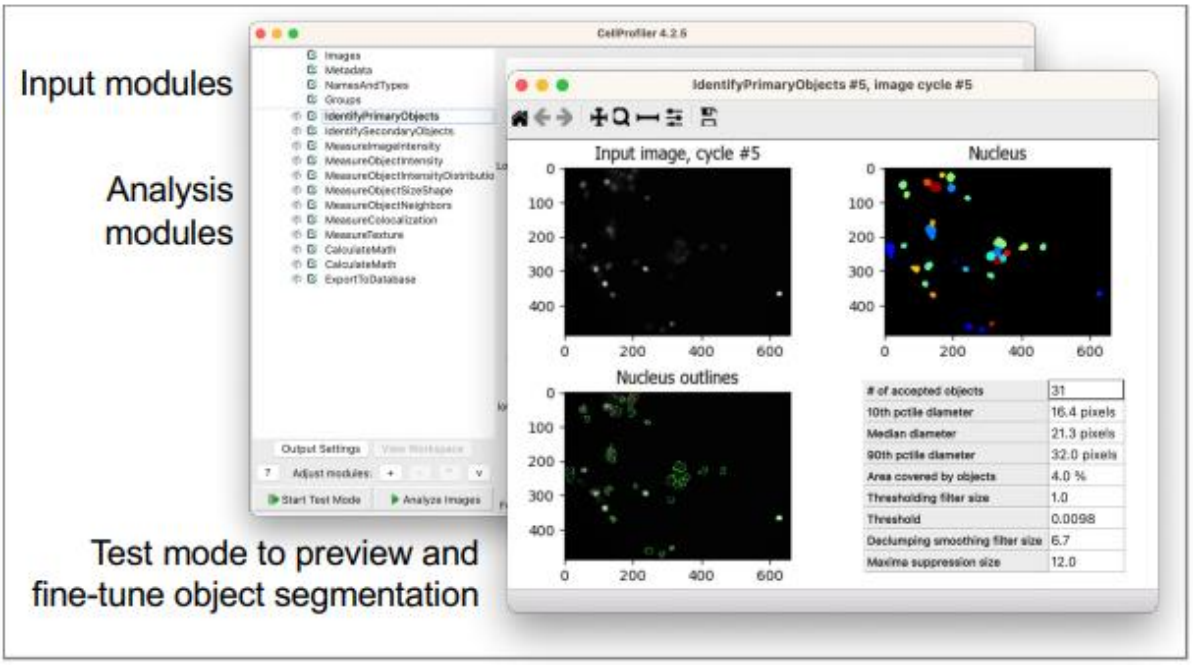

Figure 3: CellProfiler interface and pipeline overview. CellProfiler interface, showing the input modules (top 4 modules), followed by analysis modules, constituting a typical pipeline used for ChromaLive feature extraction. The superimposed window shows nucleus segmentation and is useful when fine-tuning the "Identify PrimaryObjects" module in Test Mode, used here to segment cell nucleus based on Hoechst 33342 staining.

Data analysis: supervised and unsupervised approaches

If the goal of image analysis is to extract as much information as possible from the assay, the aim of data analysis is to process this information, usually hundreds of descriptors for thousands of objects, into interpretable results. Data analysis can be performed in a supervised or an unsupervised manner. In this section, we will explore both approaches. Unsupervised data analysis aims to reduce the complexity of the information while preserving its integrity, making it comprehensible. Typically, this involves reducing the dimensionality of the data (to be able to represent the data in two or three-dimensional graphs for example) and identifying relationships between conditions to generate clusters with similar phenotypes. The advantage of this approach is its unbiased nature, as it doesn't rely on user input. However, the absence of expert insight into the experimental conditions makes it challenging to connect the obtained clusters with underlying biology (Omta et al., 2020).

On the other hand, supervised data analysis involves training algorithms to recognise similar conditions and classify cells into discrete classes (Dürr and Sick, 2016; Kraus et al., 2016). With the correct controls and the proper labeling, this means that it is ultimately possible to have an expert user create models and provide the necessary training to establish a relationship between imaged samples and known mechanisms of action. While this approach relies on user knowledge and is prone to user bias, it offers the advantage of leveraging expert insight.

In the following section, we will present examples of both approaches. First, an unsupervised analysis performed on the StratoMineR™ platform provided by Core Life Analytics, which demonstrates how this approach can identify different cell phenotypes (Omta et al., 2016). Afterwards, we will explore the use of supervised classification tools like CellProfiler Analyst to quantify compound-induced phenotypic changes.

Unsupervised analysis to cluster cell phenotypes

Dose-response assays for multiple compounds were performed following the protocol described in the previous section. In summary, MCF-7 cells were plated at 1,000 cells per well in a 384-well plate, with cell culture medium containing ChromaLive. After overnight adherence, cells were stained with 100ng/mL Hoechst 33342 for 3 hours at 37°C and 5% CO2. Subsequently, the cells were treated with various doses of compounds before imaging at regular time intervals for up to 72h on a PerkinElmer Opera QEHS system using the acquisition settings outlined in Table 1. Following imaging, CellProfiler pipelines, similar to the one described earlier, were used to extract features. The information related to the "cell" objects at the 24-hour time point was exported to a spreadsheet and uploaded to the StratoMineR™ platform.

This web-based data-analytics platform facilitated data normalization and data reduction in a holistic fashion. We performed data reduction using principal component analysis (PCA) and generated 6 components from over 300 features. As a result, compound-specific clusters were visualised, with intermediate phenotypes observed at lower compound doses (Figure 4). As an example, Figure 4a illustrates staurosporine or thapsigargin-treated MCF-7 cells as distinct clusters relative to DMSO-treated controls. Focusing on staurosporine treatment (Figure 4b), an intermediate phenotype cluster corresponding to 5-50nM doses (green and light green points) can be distinguished, which is clearly distinct from a final phenotype cluster observed at higher 0.5-5µM doses of staurosporine (orange and red points). This preliminary result validates the classification of staurosporine-induced phenotypes into "intermediate" and "final" classes, which will be further explored using supervised training tools in the next section.

Example of supervised classification focused on staurosporine-induced apoptosis

The phenotypic changes induced by staurosporine, a known inducer of apoptosis (Belmoktar et al., 2001), were analysed. With the same dataset as presented earlier, we utilised the supervised analysis capabilities of CellProfiler Analyst with the Classifier module (Figure 5a). Images from dose-response experiments of various compounds and time points were previously analysed with CellProfiler, and the extracted features were exported to a database with the relevant properties file. Importantly, compound dosage was linked to each image through metadata, providing a platemap and treatment information for each condition.

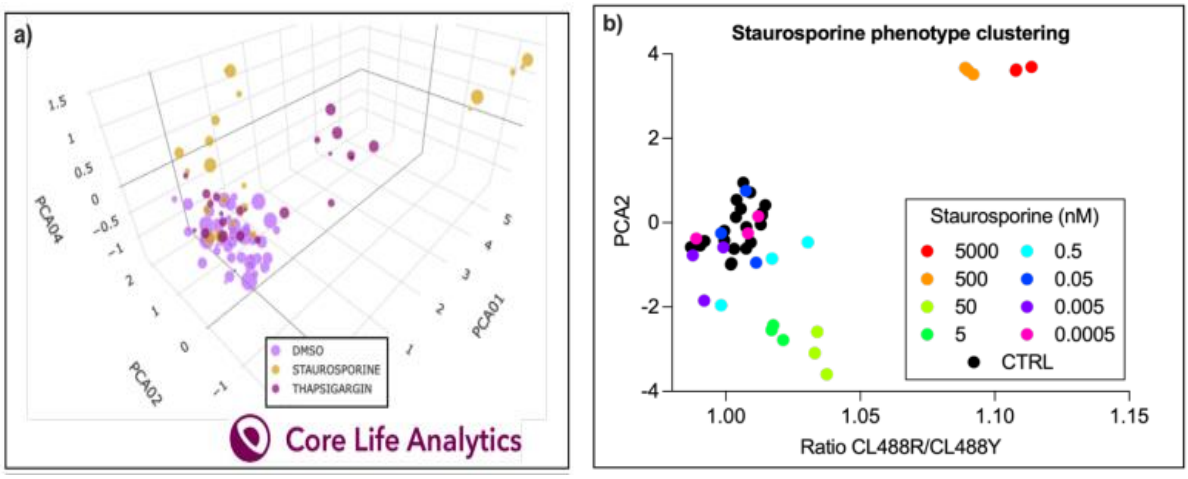

Figure 4: Detecting phenotypes with unsupervised clustering. Analysis performed on the H.C. StratoMineR™ platform developed by Core Life Analytics, and represented here after normalization and data reduction. a) 3D representation of principal component analysis (PCA), of ChromaLive-stained MCF-7 cells treated for 24h with varying doses of staurosporine or thapsigargin. b) 2D representation of staurosporine-induced clusters, the PCA01 axis of Figure 4a) has been replaced by the CL488R/CL44Y intensity ratio (see Note)

In this analysis, and following the earlier results obtained through unsupervised analysis, cells were separated into three different classes based on appearance (Figure 5a, columns from left to right):

1. Control: Cells presenting a healthy morphology and staining. Cells treated with a control amount of DMSO were used to train this class.
2. Apoptosis (intermediate): This intermediate phenotype class comprised cells treated with nanomolar range doses of staurosporine. Overall, this phenotype corresponded to elongated cells with fewer contacts with neighboring cells and less homogeneous ChromaLive staining.
3. Final: This class corresponds to dead cells characterised by an intensely labeled condensed nucleus and concentrated ChromaLive488 signal, paired with an absence of ChromaLive561 signal.

After training the classifier and evaluating the model using a confusion matrix (Figure 5a), the overall experiment was scored, and p(Enriched) values were obtained for each image (Jones, 2009). These values were then plotted against different concentrations of staurosporine, and EC50 and IC50 values were determined using GraphPad Prism software (Figure 5b). The "p(Enriched) Control" curve showed an IC50 of 4nM, indicating a decrease in the population of "Control" cells in favor of intermediate and final phenotype cells. In comparison, the "p(Enriched) Final" curve exhibited a higher E C50 of 70nM, indicating an increase in the population of "Final" dead cells. To explain the discrepancy between these doses, we stipulate that the "Intermediate" population appears at lower staurosporine concentrations. However, the curve corresponding to "p(Enriched) Intermediate" was more challenging to fit, as this population increased with the dosage of staurosporine before decreasing in favor of dead cells. A tentative fit of this curve (ignoring the two higher doses of staurosporine) yielded an EC50 of 3nM, consistent with the fit of "p(Enriched) Control".

Dose-response curves serve as robust internal controls and aid in classifier verification. They help identify possible effects of "over-training," where classifier scores may reflect more user-driven training input than objective truths from the imaging experiment. For instance, having the basal level of the dose-response curve correspond to the DMSO controls used to train "Control" phenotypes, objectively validates similar phenotypes for lower concentrations of treatment. Similarly, if specific conditions are used to train "Intermediate" or "Final" phenotypes, the ability to correctly fit the data with a typical sigmoid dose-response curve confirms the presence of these phenotypes across all tested concentrations.

To complement this assay, an orthogonal assay was performed to assess cell apoptosis using the CellEvent™ Caspase-3/7 assay (ThermoFisher). The experimental design and cell seeding in a 96-well plate were identical. Cells were imaged according to the manufacturer's guidelines, and intensity measurements were plotted against concentrations of staurosporine (Figure 6b). After fitting, an EC50 of 95nM was calculated, which is in good accordance with the "Final" phenotype observed with ChromaLive.

Additionally, a new dose-response tool currently being developed on the StratoMineR™ platform provides similar EC50 values (Figure 6a). For example, the "Final" phenotype can be accurately represented using the ChromaLive488- Red to ChromaLive488-Yellow ratio1 (EC50: 120nM), while the "Intermediate" phenotype can be represented with PCA04 obtained from principal component analysis (EC50: 8nM). To facilitate comparison, the different dose-response curves are presented together in Figure 6c, allowing for a better evaluation of the relationships between apoptosis phenotypes and doses of staurosporine.

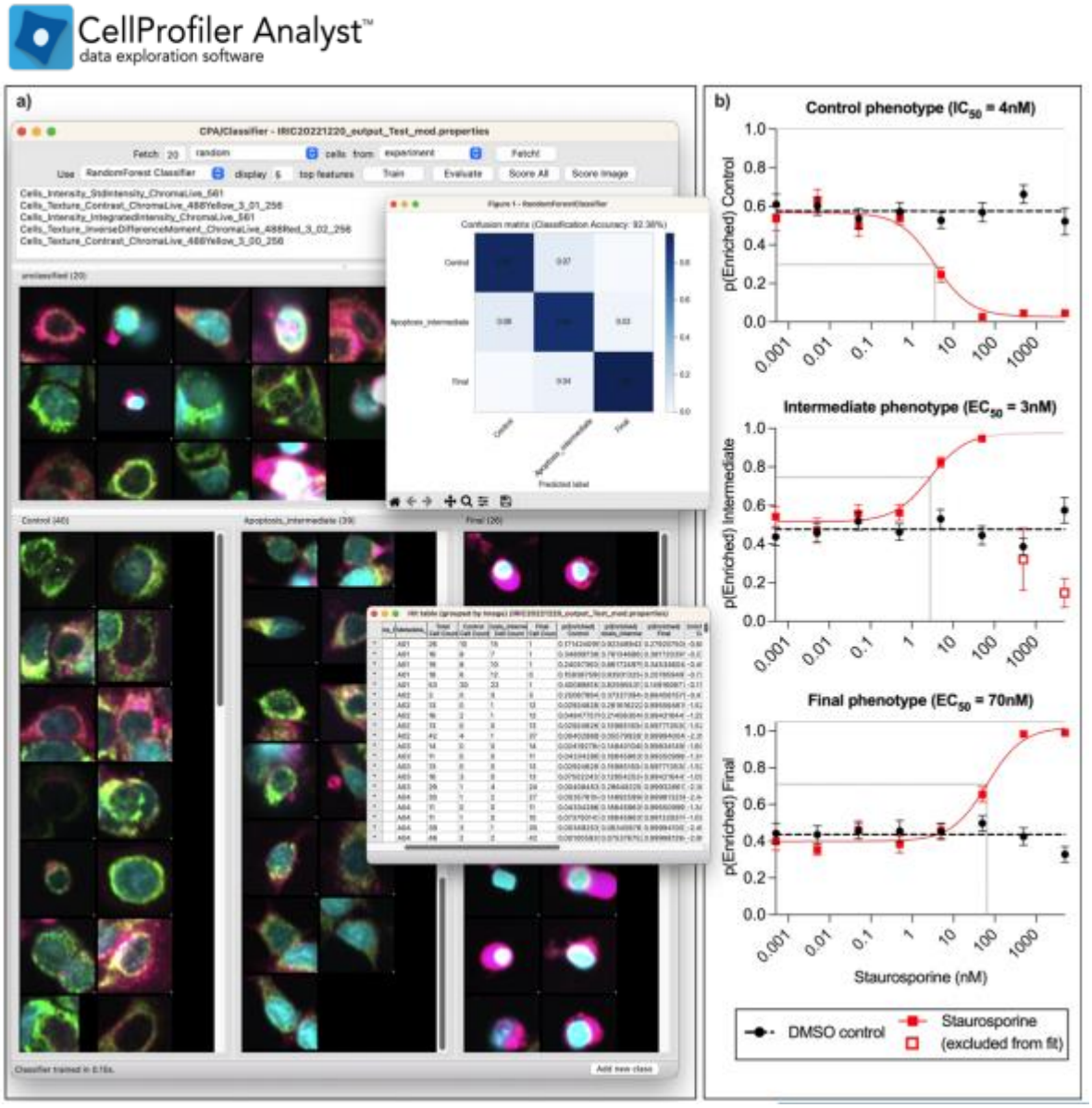

Figure 5: CellProfiler Analyst and staurosporine dose-response. a) Interface of the Classifier module of CellProfiler Analyst (CPA), showing objects (cells here) classified into 3 classes: "Control", "Apoptosis_intermediate" and "Final". Inset windows show a confusion matrix evaluating the trained classifier, as well as a hit table obtained after scoring the experiment. b) Plotted p(Enriched) values for each class, and fitted IC50 ("Control" phenotype, top graph) or EC50 ("Intermediate" and "Final" phenotypes, respectively middle and bottom graph). For each graph, various DMSO control triplicates are plotted along the xaxis to represent phenotype scoring throughout the plate.

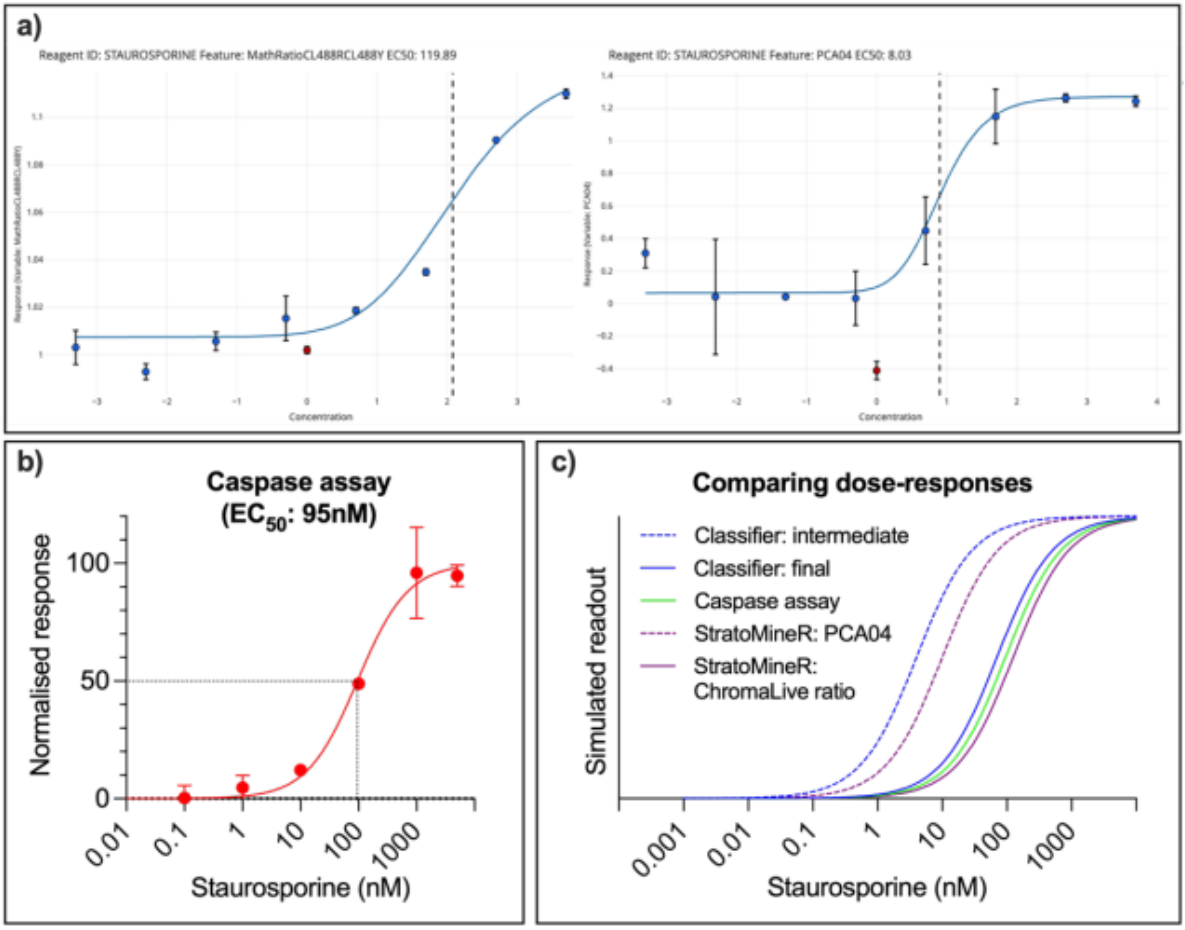

Figure 6: Comparing staurosporine dose-responses. a) Dose-response curves obtained from the StratoMineR™ platform, after data reduction and PCA. Left graph plotted with CL488R/CL488Y ratio (EC50: 120nM), right graph with PCA04 (EC50: 8nM). DMSO control replicates are represented in red. b) Staurosporine-induced apoptosis measured with the CellEvent™ Caspase-3/7 assay. Normalised intensity measurements are plotted against staurosporine concentrations. Fitted EC50: 95nM. c) Simulated curves with all measured EC50 across assays and analysis methods. Dashed lines correspond to "Intermediate" phenotypes (EC50: 3- 8nM), and full lines correspond to "Final" phenotypes (dead cells) (EC50: 70-120nM).

Together, these results show the potential of ChromaLive staining to discriminate between different phenotypes, even when induced by the same compound. Indeed, an "Intermediate" phenotype that could be attributed to cells undergoing apoptosis is measurable in the nanomolar range of staurosporine after 24 hours of treatment. On the other hand, a "Final" phenotype, corresponding to dead cells as evidenced by the caspase assay, is only predominant at higher doses around 100nM

Example of temporal phenotypes: thapsigargin-induced ER stress and cell death

A similar approach was taken to investigate phenotypic changes induced by thapsigargin, a known inducer of ER stress and cell death (Sehgal et al., 2017). Live-cell imaging with ChromaLive enabled the observation of cells at different time points after treatment. MCF-7 cells were imaged at 3h, 6h, 12h, and 24h following thapsigargin treatment. Similarly to the previous section, after image segmentation and feature extraction with CellProfiler, cells were classified into three different classes using the Classifier tool in CellProfiler Analyst: Control, ER stress (intermediate), and Final (Figure 7a). The intermediate phenotype corresponded to cells presenting a distinctive brighter, ring-like ChromaLive561 staining, while dead cells classified in the "Final" phenotype exhibited a condensed nucleus and a brighter ChromaLive488 staining.

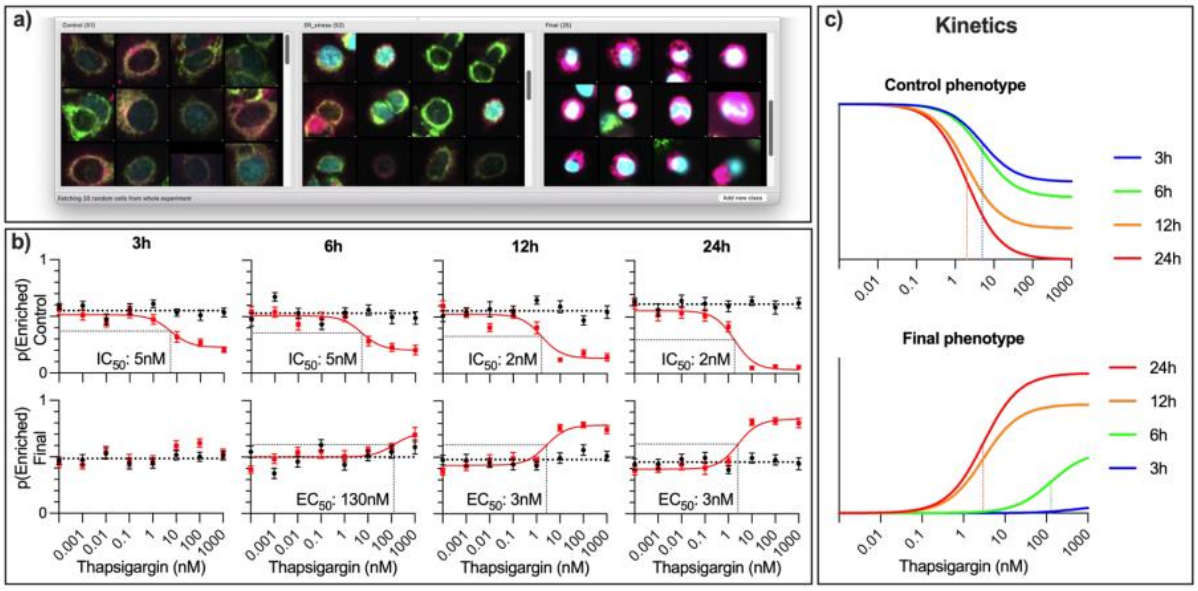

Figure 7: Thapsigargin treatment kinetics. a) Screenshot of CellProfiler Analyst (Classifier module) interface showing different classified phenotypes: "Control", "ER_stress" and "Final" (dead cells). b) Representing p(Enriched) values against thapsigargin concentrations for "Control" phenotype (top row) and "Final" phenotype (bottom row) at different time points (3h, 6h, 12h, 24h, from left to right). DMSO controls are plotted in black, and thapsigargin-treated conditions in red. c) Superposition of simulated curves for the "Control" phenotype (left graph) and "Final" phenotype (right graph). Curves are simulated based on measured IC50 and EC50 values as well as response amplitudes.

Training the classifier over multiple time points generated graphs for the "Control" and "Final" phenotypes (Figure 7b). Fitting dose-response curves revealed variations in IC50 and EC50 values as well as the amplitude of phenotypic changes over time. Notably, an IC50 of 5nM was

measured after 3 and 6 hours of thapsigargin treatment, which decreased to 2nM after 12 hours (remaining stable at 24 hours). In contrast, no significant enrichment in the "Final" phenotype was observed after 3 hours, and after 6 hours of thapsigargin treatment,the measured EC50 was much higher at around 130nM. However, after 12 hours of treatment, the EC50 decreased to 3nM (stable at 24 hours). These different curves are represented together in Figure 7c to illustrate the evolution of IC50 and EC50 values, as well as the amplitude of phenotypic changes over time. This demonstrates how the "Final" phenotype occurs more rapidly at higher doses, and how the ER stress "Intermediate" phenotype eventually leads to cell death.

Conclusion

The presented results demonstrate the remarkable versatility of ChromaLive as a non-toxic live cell dye, opening up numerous possibilities for experimental and screening design. Our findings indicate that ChromaLive has minimal impact on cell proliferation and gene expression in cell culture, allowing for the growth of cell models in the presence of the dye. This feature is particularly advantageous in progenitor or stem cell differentiation studies. Furthermore, the extended incubation time offered by ChromaLive proves valuable in achieving stable and homogeneous staining in 3D cell models (refer to Figure 8). Beyond its simple addition to the culture medium, ChromaLive's fluorogenic properties, which enable its fluorescence only when integrated into cells,further simplify experimental design by eliminating the need for washing steps

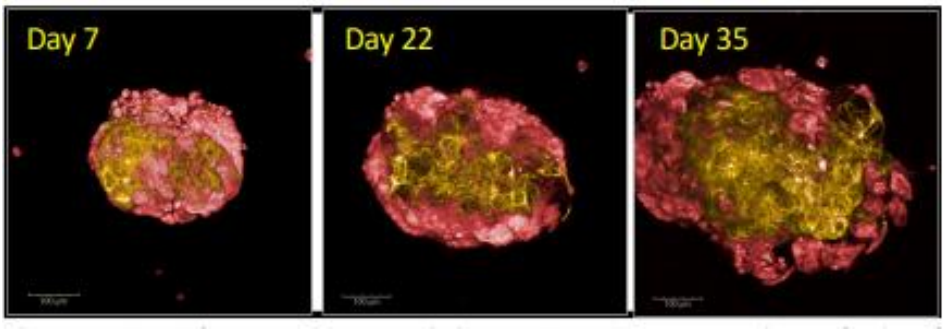

Figure 8: ChromaLive staining of cancer patient-derived organoids. From left to right, ovarian cancer patient-derived organoids were grown for 7, 22 and 35 days respectively, in custom hydrogel and with culture medium containing ChromaLive, imaged at 20x magnification. In yellow: ChromaLive561; in red: ChromaLive488_Red. Scale bar: 100nm. (Images courtesy of Betty Li and Alla Buzina from the Sunnybrook Research Institute).

. The "mix-and-read" protocol highlights how seamlessly ChromaLive can be incorporated into existing screening assays. By adding ChromaLive to the culture medium and performing the

subsequent live imaging at the desired incubation time, researchers can easily integrate this dye into their experiments. Moreover, this streamlined experimental design facilitates repeated imaging of the same conditions, introducing a kinetic dimension to high-volume screens. Notably, our studies on 2D cell cultures demonstrate the possibility of conducting kinetic screens over 72 hours without significant impact on cellular health or loss of cell labeling intensity.

Importantly, our validation demonstrates that while ChromaLive relies on non-specific cellular staining, it is remarkably well-suited for agnostic phenotypic profiling. A single assay with ChromaLive can detect a multitude of cellular phenotypes, although we have not delved into the specific details for the sake of brevity. Simple intensity measurements can distinguish between live and dead cells, while subtle image descriptors can reveal more nuanced phenotypes such as early apoptosis, ER stress, autophagy, and more. Leveraging the progress in data science and high content screening tools, ChromaLive emerges as an exceptionally versatile and straightforward tool for exploring cell phenotypes in the context of drug discovery.

Acknowledgments

Special thanks are due to Raphaelle Lambert and Patrick Gendron at IRIC for performing and analysing the RNA sequencing assay. Thanks are due as well to Betty Li and Alla Buzina at the **Sunnybrook Research Institute** for their 3D imaging of organoids, and to Yuniel Fernandez-Marrero for his study on the impact of ChromaLive on cell proliferation. We would also like to thank Beth Cimini and Barbara Diaz-Rohrer at the **BROAD Institute** for their support on CellProfiler and CellProfiler Analyst . Finally, thanks to Jacques Corbeil and Emanuel Paré from the department of Molecular Medicine at Université Laval for the staining of HepaRG cells.

References

Belmokhtar, C. A., Hillion, J., & Ségal-Bendirdjian, E. (2001). Staurosporine induces apoptosis through both caspase-dependent and caspase-independent mechanisms. Oncogene, 20(26), 3354–3362.

Bickle, M. (2010). The beautiful cell: High-content screening in drug discovery. Analytical and Bioanalytical Chemistry, 398(1), 219– 226.

Bray, M. A., Singh, S., Han, H., Davis, C. T. & Carpenter, A. E. (2016). Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. Nature Protocols, 11(9), 1757–1774.

Danuser, G. (2011). Computer vision in cell biology. Cell, 147(5), 973–978.

Dürr, O., & Sick, B. (2016). Single-cell phenotype classification using deep convolutional neural networks. Journal of Biomolecular Screening, 21(9), 998–1003.

Hughes, J. P., Rees, S. S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. British Journal of Pharmacology, 162(6), 1239–1249.

Kraus, O. Z., Ba, J. L., & Frey, B. J. (2016). Classifying and segmenting microscopy images with deep multiple instance learning. Bioinformatics, 32(12), i52–i59.

Lee, J. A., & Berg, E. L. (2013). Neoclassic drug discovery: The case for lead generation using phenotypic and functional approaches. Journal of Biomolecular Screening, 18(10), 1143–1155.

Lin, S., Schorpp, K., Rothenaigner, I., & Hadian, K. (2020). Image-based high-content screening in drug discovery. Drug Discovery Today, 25(8), 1348–1361.

Mari, P. O., Verbiest, V., Sabbioneda, S., Gourdin, A. M., & Giglia-Mari, G. (2010). Influence of the live cell DNA marker DRAQ5 on chromatin-associated processes. DNA Repair, 9(7), 848–855.

Omta, W. A., van Heesbeen, R. G., Pagliero, R. J., van der Velden, L. M., & Egan, D. A. (2016). HC StratoMineR: A Web-Based Tool for the Rapid Analysis of High-Content Datasets. Assay and Drug Development Technologies, 14(8), 439–452

Omta, W. A., van Heesbeen, R. G., Shen, I., de Nobel, & Egan, D. A. (2020). Combining Supervised and Unsupervised Machine Learning Methods for Phenotypic Functional Genomics Screening. SLAS Discovery, 25(6), 655–664.

Oppermann, S., Ylanko, J., Shi, Y., Hariharan, S., & Andrews, D. W. (2016). High-content screening identifies kinase inhibitors that overcome venetoclax resistance in activated CLL cells. Blood, 128(7), 934–947.

Richard, E., Causse, S., Spriet, C., Fourré, N. & Heliot, L. (2011). Short exposure to the DNA intercalator DRAQ5 dislocates the transcription machinery and induces cell death. Photochemistry and Photobiology, 87(1), 256–261.

Sehgal, P., Szalai, P., Olesen, C., Praetorius, H. A., & Møller, J. v. (2017). Inhibition of the sarco/endoplasmic reticulum (ER) Ca2- ATPase by thapsigargin analogs induces cell death via ER Ca2 depletion and the unfolded protein response. Journal of Biological Chemistry, 292(48), 19656–19673.

Simm, J., Klambauer, G., Arany, A., Steijaert, M., & Ceulemans, H. (2018). Repurposing High-Throughput Image Assays Enables Biological Activity Prediction for Drug Discovery. Cell Chemical Biology, 25(5), 611-618.e3.

Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A., & Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics, 22(1).

Swinney, D. C., & Anthony, J. (2011). How were new medicines discovered? Nature Reviews Drug Discovery, 10(7), 507–519.

Vincent, F., Nueda, A., Lee, J., Schenone, M., Prunotto, M., & Mercola, M. (2022). Phenotypic drug discovery: recent successes, lessons learned and new directions. Nature Reviews Drug Discovery, 21(12), 899–914.

Waring, M. J., Arrowsmith, J., Leach, A. R., Leeson, & Weir, A. (2015). An analysis of the attrition of drug candidates from four major pharmaceutical companies. Nature Reviews Drug Discovery, 14(7), 475–486