

From pixels to phenotypes: Integrating image-based profiling with cell health data as BioMorph features improves interpretability

Srijit Seal^{1,2,3}, Jordi Carreras-Puigvert^{1,3}, Shantanu Singh^{1,2}, Anne E. Carpenter^{1,2}, Ola Spjuth^{1,3,*}, and Andreas Bender^{1,2,*}

¹Imaging Platform, Broad Institute of MIT and Harvard, Cambridge MA 02142; ²Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; ³Department of Pharmaceutical Biosciences and Science for Life Laboratory, Uppsala University, 752 37 Uppsala, Sweden

ABSTRACT Cell Painting assays generate morphological profiles that are versatile descriptors of biological systems and have been used to predict *in vitro* and *in vivo* drug effects. However, Cell Painting features extracted from classical software such as CellProfiler are based on statistical calculations and often not readily biologically interpretable. In this study, we propose a new feature space, which we call BioMorph, that maps these Cell Painting features with readouts from comprehensive Cell Health assays. We validated that the resulting BioMorph space effectively connected compounds not only with the morphological features associated with their bioactivity but with deeper insights into phenotypic characteristics and cellular processes associated with the given bioactivity. The BioMorph space revealed the mechanism of action for individual compounds, including dual-acting compounds such as emetine, an inhibitor of both protein synthesis and DNA replication. Overall, BioMorph space offers a biologically relevant way to interpret the cell morphological features derived using software such as CellProfiler and to generate hypotheses for experimental validation.

Monitoring Editor
Alex Mogilner
New York University

Received: Aug 1, 2023
Revised: Dec 7, 2023
Accepted: Dec 22, 2023

 New Methods

INTRODUCTION

Cell Painting profiles (Gustafsdottir *et al.*, 2013) can be used to study the morphological characteristics of cells treated with chemical or genetic perturbations and provide valuable information about the function of a biological system (Simm *et al.*, 2018; Chandrasekaran *et al.*, 2021). The Cell Painting assay involves labelling eight relevant cellular components or organelles with six fluorescent dyes, imaging them in five channels (Bray *et al.*, 2016), and analysing images (Stirling *et al.*, 2021) to provide thousands of morphological features such as shape, area, intensity, texture, correlation, etc. Cell Painting data

have been used to successfully predict drug effects on many aspects of cell health (Way *et al.*, 2021), such as cytotoxicity (Seal *et al.*, 2021), mitochondrial toxicity (Seal *et al.*, 2022), proteolysis targeting chimera (PROTAC) phenotypic signatures (Trapotsi *et al.*, 2022), cardiotoxicity (Seal *et al.*, 2023a), and other types of bioactivities (Trapotsi *et al.*, 2021; Seal *et al.*, 2023b). Further, Cell Painting data can be used to cluster together compounds with various mechanisms of action based on the similarity of resulting morphological features they induce (Nyffeler *et al.*, 2020; Way *et al.*, 2022). Thus, Cell Painting

This article was published online ahead of print in MBoc in Press (<http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E23-08-0298>) on January 3, 2024.

Conflicts of Interest: The authors declare no conflict of interest.

Contributions: S. Seal designed and performed data analysis, constructed the BioMorph space organisation, and implemented, and trained the models. S. Seal, J.C.P. and O.S. analysed the biological interpretation of morphological features. S. Seal and J.C.P. contributed to analysing the results of models. S. Seal wrote the manuscript with extensive discussions with O.S., J.C.P., S. Singh, and A.E.C. A.B. and O.S. supervised the project. All the authors (S. Seal, J.C.P., S. Singh, A.E.C., O.S., and A.B.) reviewed, edited, and contributed to discussions on the manuscript.

*Address correspondence to: Ola Spjuth (ola.spjuth@farmbio.uu.se); Andreas Bender (ab454@cam.ac.uk).

Abbreviations used: AUC, area under the curve; BLA, beta-lactamase activity; CLOOME, contrastive learning framework for image-structure pairs; CRISPR, clustered regularly interspaced short palindromic repeats; DNA, deoxyribonucleic acid; ER, estrogen receptor; HDAC, histone deacetylase; InChI, international

chemical identifier; JAK, Janus kinase; JUMP, joint undertaking in morphological profiling; MAD, mean absolute deviation; mAP, mean average precision; MAPK, Mitogen-activated protein kinase; MCC, Matthews Correlation Coefficient; MOA, mechanism of action; PAEA, Principal Angle Enrichment Analysis; PROTAC, Proteolysis Targeting Chimera; RNA, ribonucleic acid; ROS, reactive oxygen species; RTK, receptor tyrosine kinase; WS-DINO, weakly supervised form of self-distillation with no labels; WNT, WNT signaling pathway; US EPA, United States Environmental Protection Agency; UPR, unfolded protein response; STAT, signal transducer and activator of transcription; SRB, sulforhodamine B; SMILES, simplified molecular-input line-entry system; SHAP, SHapley Additive explanations; SD, standard deviation.

© 2024 Seal *et al.* This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution 4.0 International Creative Commons CC-BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

“ASCB®,” “The American Society for Cell Biology®,” and “Molecular Biology of the Cell®” are registered trademarks of The American Society for Cell Biology.

features serve as a tool for investigating the chemical space and enabling the prediction of a compound's biological activities (Liu *et al.*, 2023; Pruteanu and Bender, 2023).

In general, Cell Painting features are obtained using classical image processing software, such as CellProfiler (Stirling *et al.*, 2021). After establishing the threshold for distinguishing signal from the background noise, classical image processing software identifies all signal-containing pixels and their intensity, and groups neighbouring pixels into objects using object-based correlations (Help! How does the Robust Background method work? | Carpenter-Singh Lab). The measured morphological features are then extracted from each object (cell or subcellular structure). Given this image processing pipeline, Cell Painting features primarily represent numerical data from image analysis (often aggregated to the treatment level for machine learning tasks), rather than directly reflecting the underlying biological processes or molecular interactions. (Help! How does the Robust Background method work? | Carpenter-Singh Lab) Therefore, interpreting the Cell Painting data and making informed decisions about drug safety, toxicity, efficacy, or the underlying mechanisms and cellular processes based on such data remains challenging. This suggests that integrating Cell Painting features with some a priori knowledge about the biological effects of different chemical or genetic perturbations may result in improved predictive power of models derived from Cell Painting data.

An orthogonal strategy that considers a priori knowledge about the biological effects is the Cell Health assay, a set of two image-based assays (Chessel and Carazo Salas, 2019) that collectively capture a broad range of biological pathways. The Cell Health assay thus records measurable characteristics from cellular responses to different treatments (or environmental conditions, pathological states, etc.; Markowitz, 2010; Szalai *et al.*, 2019) which determine the overall condition, functionality, and viability of cells (Riss *et al.*, 2016), including the different stages of the cell cycle. Following a similar premise, a study by Way *et al.* (2021) used the Cell Health assay and CRISPR/Cas9 to genetically perturb a small subset of 118 gene perturbations across three cell lines. Recording the effects of these genetic perturbations using carefully chosen reagents for specific cellular processes (e.g., apoptosis, DNA damage, etc.) allowed them to define 70 Cell Health readouts that can be used to quantify and model cellular responses to different treatments (Way *et al.*, 2021). The Cell Health readouts are directly related to mechanisms and cellular function and can be used to predict the mechanism of action (MOA) of the perturbation and derive functional conclusions. However, unlike the hypothesis-free Cell Painting assay, the Cell Health assay requires specifically targeted reagents focused on individual measurement and is difficult to scale for high throughput applications.

Recent advancements in data integration methodologies have demonstrated the potential of connecting distinct data modalities to enhance interpretability. This is common in gene set enrichment analysis where methods such as the χ^2 test have been used which combine a set of gene expression features connected by annotations to a common pathway into a gene-set level statistic (Hung *et al.*, 2012). Another example is the Gene Ontology transformed gene expression profiles of small molecule perturbations developed using Principal Angle Enrichment Analysis (PAEA; Clark *et al.*, 2015; Wang *et al.*, 2016). Other studies have combined prior knowledge of pathways and gene expression data to identify latent variables (inferred using models) to elucidate underlying patterns in gene sets that are unique compared with the input gene expression data (Basili *et al.*, 2022). The application of contrastive learning has also emerged, such as CLOOME, aiming to bridge the gap between image-based representations and chemical structures by embed-

ding them into the same representation space (Sanchez-Fernandez *et al.*, 2023). In this context, our work introduces a method specifically tailored for classical features derived from the Cell Painting using softwares such as CellProfiler, with an emphasis on data-based feature grouping. Unlike extant approaches that predominantly aim to improve target prediction, our methodology aims to establish a novel interpretative space, facilitating a deeper comprehension of cellular biology phenomena.

Here, we address the limitations of both Cell Painting and Cell Health assays by integrating their capabilities. We propose a new feature space, called the BioMorph space, that provides a function-informed framework for interpreting Cell Painting features in the cell biology context. We used publicly available Cell Painting data and Cell Health data (Way *et al.*, 2021) to define this BioMorph space. To demonstrate the use of the BioMorph space, we used the Cell Painting features from chemical perturbations (Bray *et al.*, 2017) to predict a range of nine broad biological activities from ToxCast, such as apoptosis, cytotoxicity, oxidative stress, and ER stress. We then mapped important Cell Painting features from these models into BioMorph terms. Identifying the BioMorph terms that contribute most strongly to model performance helped generate MOA hypotheses, some in agreement with the existing literature and some novel. Taken together, our proposed method offers several potential advantages, including improved interpretability of cell morphology features, enhanced understanding of cellular mechanisms and MOA, and more interpretable predictions of drug toxicity and efficacy. All BioMorph datasets generated from this study are available at <https://broad.io/BioMorph>.

RESULTS AND DISCUSSION

We developed a structured framework for mapping Cell Painting features to a more biologically synthesized BioMorph space. We used feature selection, linear regression, and Random Forest classifiers on the publicly available Cell Painting and Cell Health datasets (Way *et al.*, 2021) for a set of 119 CRISPR perturbations (for further details see *Materials and Methods*). This mapping was then used to interpret models predicting biological activity using a dataset containing morphological profiles of 30,000 small molecules produced using the Cell Painting assay (Bray *et al.*, 2017). Mapping those Cell Painting features that contribute the most to the performance to the BioMorph space led to an improvement in interpretability and allowed us to generate hypotheses on the cause of cellular effects.

Development of the BioMorph space through the integration of Cell Painting and Cell Health assays

We mapped the groups of Cell Painting features into five levels within the BioMorph space as shown in Figure 1 (see *Materials and Methods* for technical details and Supplemental Table S1 and Supplemental Figure S1 for all terms). These levels were chosen to leverage the maximum information from the Cell Health assay and include the Cell Health assay type (Level 1), Cell Health measurement type (Level 2), specific Cell Health phenotypes (Level 3), Cell process affected (Level 4), and the subset of Cell Painting features (Level 5). The first level, the Cell Health assay type, represents results from one of the two screening assays used to measure the Cell Health parameters, for example, the viability assay or the cell cycle assay. The second level, Cell Health measurement type, describes the various aspects of Cell Health measured in that assay, such as cell death, apoptosis, reactive oxygen species (ROS), and shape for viability assays, and cell viability, DNA damage, S phase, G₁ phase, G₂ phase, early mitosis, mitosis, late mitosis, and cell cycle count for cell cycle

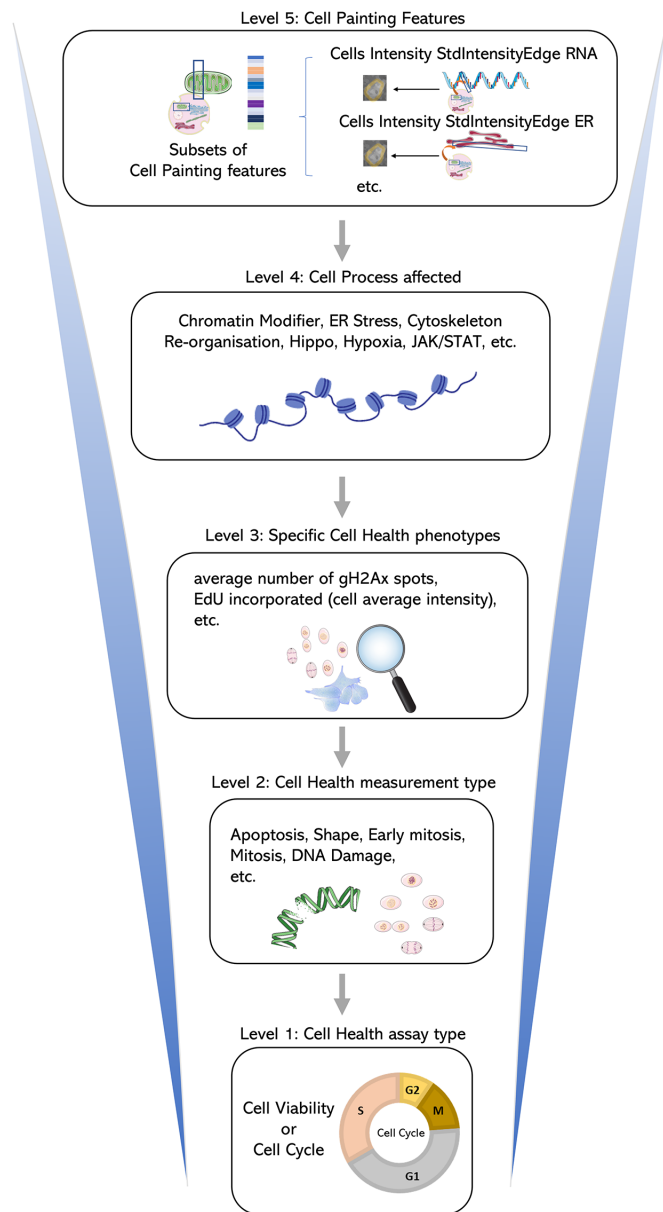


FIGURE 1: A map of the BioMorph space. A general representation of the hierarchy of levels is shown (with examples) for each BioMorph term that is organised from Cell Painting features (level 5) and containing information on Cell Health (level 3) associated with measurement type (level 2) under an assay type (level 1) associated with Cell process affected (level 4). Further terms in Supplemental Figure S1 with all terms listed in Supplemental Table S1.

and DNA damage assays. The third level, specific Cell Health phenotypes, describes specific assay readouts that capture different aspects of the phenotype, such as the fraction of cells in G₁, G₂ or S-phase cells. The fourth level, the Cell process affected, contains information on the type of Cell process affected that caused the change in morphological characteristics, for example, effects of chromatin modifier, DNA damage, metabolism, etc. Finally, the fifth level, Cell Painting features, is the subset of Cell Painting image-based features that map to the combination of the previous four levels. These five levels formed the basis of the BioMorph space.

To build the BioMorph space we focused on the overlap of perturbations between Cell Painting and Cell Health assay containing

827 Cell Painting features and 70 continuous Cell Health endpoints. We used an all-relevant feature selection method Borutapy (Kursa and Rudnicki, 2010; Figure 2, step A) to detect a subset of Cell Painting features that contain information important for predicting each of the 70 Cell Health labels. Further, we trained a baseline Linear Regression model (Figure 2, step B) and determined which subsets of Cell Painting features are relatively better predictors for each of the 70 Cell Health labels. Meaningful models were built for 34 Cell Health labels which resulted in corresponding 34 subsets of Cell Painting features. Next, for each of the Cell Health labels, we used Borutapy to select subsets of Cell Painting features that could distinguish a particular CRISPR perturbation from the negative controls (Figure 2, step C). Lastly, we trained a baseline Random Forest Classifier (Figure 2, step D) to predict which of the sets of selected Cell Painting features perform better at differentiating negative controls from respective CRISPR perturbations with a Matthews Correlation Coefficient (MCC) >0.50. This led to 412 subsets (combinations of the various levels above) of informative Cell Painting features which were used to define 412 BioMorph terms (Supplemental Figure S1; Supplementary Table S1 lists all the terms and their description). Thus, each BioMorph term integrates a unique combination of information derived from the perturbations and Cell Health labels in the Cell Health assay and a subset of Cell Painting features.

For example, the BioMorph term “viability_apoptosis_vb_percent_dead_only_Chromatin Modifiers” records a morphological change that includes information about the “fraction of caspase negative in dead cells” (level 3) associated with apoptosis (level 2), cell viability (level 1), and the effect of CRISPR knockout of a gene associated with a chromatin modifier benchmarked against the negative control (level 4) for which a particular set of Cell Painting features (level 5) contained a signal to distinguish from negative control. This multilevel approach allows for a more nuanced understanding of cellular health and its relation to specific biological mechanisms. In the example given above, the caspase-negative dead cells are a readout for cells that have undergone nonapoptotic cell death (Tait and Green, 2008). Furthermore, the term associates this form of cell death with the effects of the CRISPR knockout of a gene associated with a chromatin modifier, which is consistent with existing evidence that certain inhibitors that affect chromatin modifications, such as histone deacetylase (HDAC) inhibitors, can initiate nonapoptotic cell death mechanisms (Shao *et al.*, 2004). Therefore, this specific BioMorph term captures signals associated with these biological characteristics and MOA.

BioMorph space retains all information for biological activity from the original Cell Painting features

We first ensured that BioMorph space contains all information from the original Cell Painting readouts, which we found to be the case as shown in Supplemental Figure S2. We used Random Forest classifiers using 398 BioMorph terms directly as features (p values from a χ^2 test on the groups of Cell Painting features; although there were 412 terms defined, only 398 terms out of these were noninfinite and continuous and used for modelling). We compared these classifiers to the models trained on all 827 Cell Painting features. Supplemental Table S2 shows the mean Area Under Curve-Receiver Operating Characteristic (AUC) and mean balanced accuracy from the 20 internal test sets of the repeated nested cross-validation (Parvande *et al.*, 2020) for all nine biological activities. Overall, models using Cell Painting features (mean AUC = 0.60) achieved a similar performance compared with models using BioMorph terms (mean AUC = 0.61; as shown in Supplemental Figure S2 with a

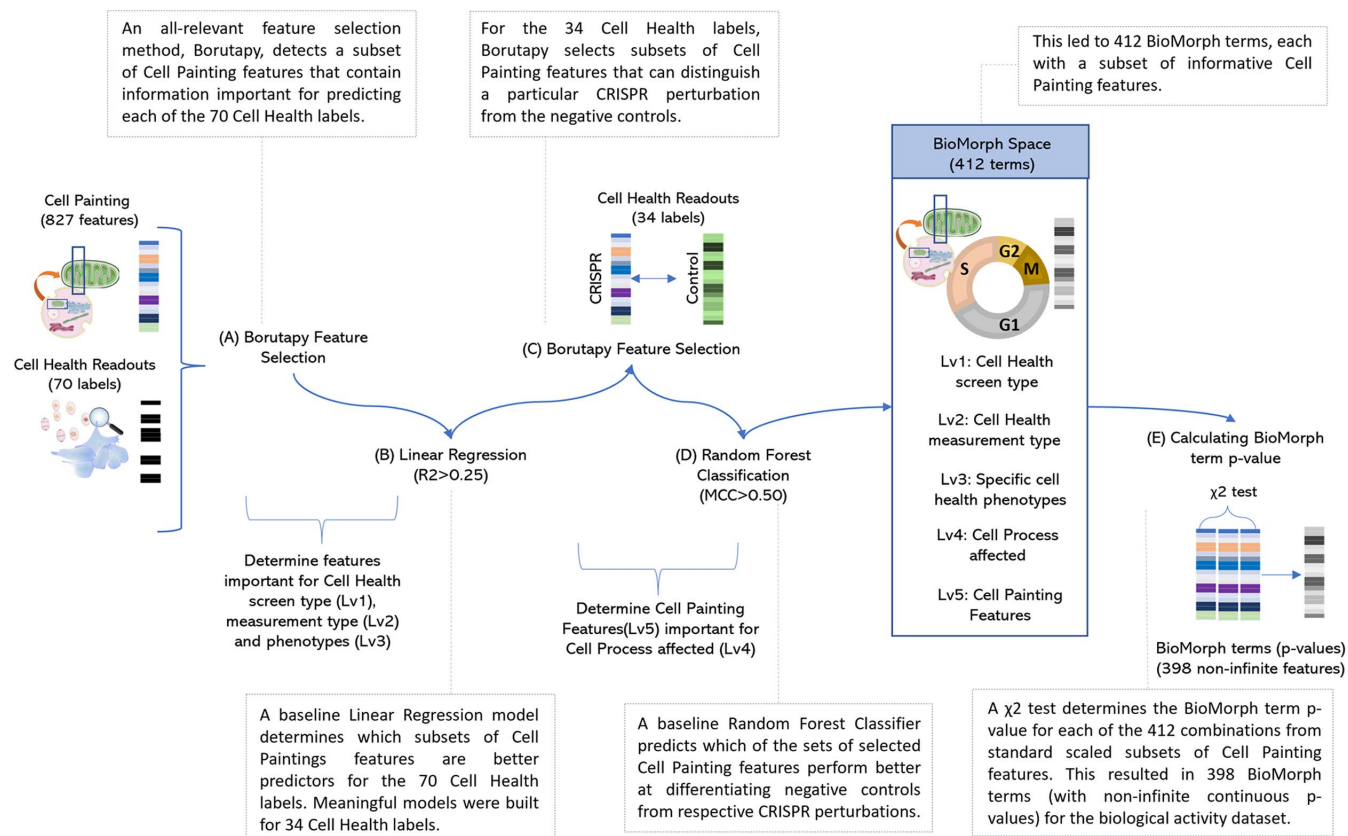


FIGURE 2: Schematic representation of methodology to generate BioMorph terms mapped from CRISPR perturbations measured by the Cell Painting assay and Cell Health assay. Further details on all terms are in Supplemental Figure S1 with all BioMorph Space listed in Supplemental Table S1.

paired t test). Thus, transforming important Cell Painting features from models into the BioMorph space made these models more interpretable without any loss in performance compared with models using BioMorph terms directly.

Incorporating information about phenotypic characteristics (Cell Health phenotype; level 3) enhances the ability to connect Cell Painting features (level 5) to biological activity from ToxCast

To compare the ability of Cell Painting features alone, or when integrated with Cell Health phenotypes (level 3), to predict biological activity, we used 56 cytotoxicity and cell stress response assays from a public dataset called ToxCast (Exploring ToxCast Data | US EPA). We generated predictions for nine biological activities (for the mapping 56 assays into nine activity labels see Judson *et al.*, 2016): (1) upregulation of apoptosis (apoptosis up); (2) cytotoxicity as measured using beta-lactamase activity as a viability reporter (Riss *et al.*, 2016; cytotoxicity BLA); (3) cytotoxicity measured using SulfoRhodamine B assays that quantify cellular density based on the protein content (Riss *et al.*, 2016; cytotoxicity SRB); (4) ER stress; (5) heat shock; (6) microtubule upregulation; (7) upregulation of mitochondrial disruption; (8) upregulation of oxidative stress; and (9) decrease in proliferation. We cross-referenced these nine biological activities with public Cell Painting profiles to focus on a dataset of 658 structurally unique compounds. For each of the nine biological activities, we trained Random Forest classifiers using 827 Cell Painting features to build predictive models and calculated feature importance for each Cell Painting feature. For eight out of nine biological activities (mitochondrial disruption was excluded because its models recorded

$AUC < 0.50$ and were not interpreted), the Cell Painting features most contributing to the eight models were mapped into BioMorph terms revealing interesting details about the associations between morphological features, phenotypic characteristics and cellular processes, as shown for the endpoint “ER stress” in Figure 3 for illustrative purposes. In this example, the BioMorph space terms that contain the highest percentage overlap with the Cell Painting features associated with the ER stress revealed potential secondary mechanisms of “ER stress” biological activity, such as G₂ cell cycle arrest (level 3) and the JAK/STAT signalling pathway (level 4), both in agreement with the literature (Bourougaa *et al.*, 2010; Meares *et al.*, 2014).

At the level of phenotypic characteristics, the five most-contributing Cell Health phenotypes (level 3) for the eight biological processes are shown in Figure 4 (with a comprehensive analysis across various levels of BioMorph terms given in Supplemental Table S3). For the biological process of apoptosis, the most-contributing Cell Health phenotype (level 3) was the fraction of cells containing more than three γ H2AX spots per cell, indicating DNA damage (Figure 4). This finding is consistent with our understanding of apoptosis as a coordinated response to DNA damage (Wang, 2001). In terms of cytotoxicity predictions, we observed that the performance of predicting results of BLA assays was improved when the BioMorph terms that incorporate Cell Health phenotypes (level 3) related to DNA damage for cells in S and G₂ phases (Figure 4), in agreement with the well-established effect of DNA damage on cell cycle arrest. On the other hand, SRB assays measure protein content, which is affected by overall cell death, including nonapoptotic cell death, and we observed that Cell Painting features contributing to model performance here incorporated caspase-negative death Cell Health

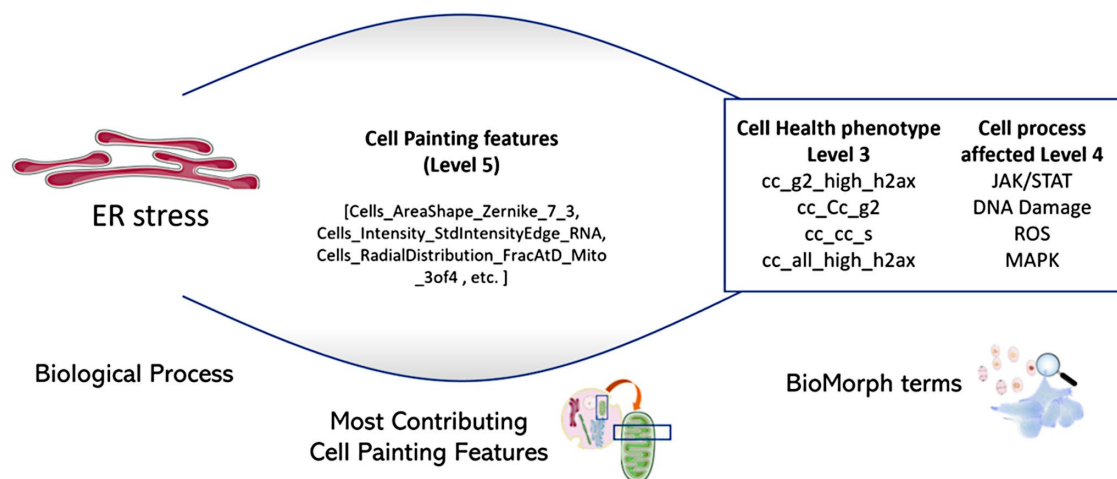


FIGURE 3: The subset of most-contributing Cell Painting features (level 5) for the model predicting ER stress and the BioMorph terms enriched from this subset (BioMorph terms that contain the highest percentage overlap with these Cell Painting features). This revealed potential secondary mechanisms of biological activity such as G₂ cell cycle arrest and the JAK/STAT signalling pathway was the most enriched Cell Health phenotype (level 3) and Cell process (level 4), respectively, for ER stress.

phenotypes (Figure 4). The Cell Health phenotypes (level 3) that contributed the most to the biological activities of ER stress, heat shock, and proliferation decrease were related to high γ H2AX activity (based on the feature related to the fraction of G₂ cells with >3 γ H2Ax spots within nuclei, Figure 4), indicating DNA damage. This is consistent with previously reported observations that ER stress and heat shock cause cell cycle arrest at both G₁/S and G₂/M phases (Brewer *et al.*, 1999; Kühl and Rensing, 2000; Bourougaa *et al.*, 2010). For the biological activity of microtubule upregulation, the most-contributing Cell Health phenotypes (level 3) were the overall DNA damage and the fraction of caspase-negative dead cells, in agreement with their roles in cell death (Kim, 2022). Finally, for the biological activity of oxidative stress, the most contributing Cell Health phenotype (level 3) was the average nucleus roundness, which is consistent with the significant crosstalk between DNA damage, oxidative stress, and nuclear shape alterations (Barascu *et al.*, 2012). Taken together, we found that the BioMorph space (level 3 Cell Health phenotypes) effectively captured biologically relevant information, allowing for a more nuanced understanding of how biological processes overall affect specific cellular processes. This is particularly advantageous compared with using Cell Painting features directly where no measurements on cell cycle phase or cell processes are made directly.

Integrating information about the Cell process affected (level 4) enhances insights into mechanisms of biological activity

In addition to the information about phenotypic characteristics, the BioMorph space also includes information about specific cellular processes responsible for the alterations in cell morphology, which in turn can help to identify potential targets and biological pathways that, when modulated, could lead to desired phenotypic changes. Therefore, we examined information from affected cellular processes (level 4 of the BioMorph Space) for each of the eight biological activities. The top five enriched Cell processes associated with each of the eight biological activities are shown in Figure 5, with a comprehensive analysis across various levels of BioMorph terms given in Supplemental Table S3. For each of the eight endpoints, we found consistent agreement between the top enriched Cell pro-

cesses and the existing literature. For example, in the case of apoptosis endpoint, the top three enriched processes were ROS, receptor tyrosine kinase (RTK) and mitogen-activated protein kinase (MAPK) pathways (Figure 5), which agrees with the existing literature (Howard *et al.*, 2003; Redza-Dutordoir and Averill-Bates, 2016; Yue and López, 2020). The JAK/STAT signalling pathway was the most enriched Cell process for ER stress (Figure 5), aligning with its role in ER stress-induced inflammation (Meares *et al.*, 2014). Similarly, the most enriched processes for the other endpoints (Figure 5), that is, Hippo signaling pathway for Cytotoxicity BLA, cyclosporine binding protein for Cytotoxicity SRB, DNA damage for heat shock response, apoptosis and hypoxia for oxidative stress, and Hippo pathways for proliferation, are all in agreement (Zaghloul *et al.*, 1987; Yu and Guan, 2013; Wang *et al.*, 2015; Kantidze *et al.*, 2016; McGarry *et al.*, 2018). Collectively, these findings illustrate the high level of agreement between BioMorph terms and well-established biological knowledge. They also highlight how integrating information about biological processes (level 4 in BioMorph space) allows for more mechanistic interpretations and predictions.

BioMorph terms can be used to generate hypotheses for a compound's mechanisms of action

We next investigated how BioMorph terms can reveal more specific mechanisms of action of a compound causing a particular biological activity. To this end, we analysed 56 predicted true positive compounds across nine biological activities and analysed the SHapley Additive exPlanations (SHAP; Scott Lundberg, 2018) values of Cell Painting features (a positive SHAP value for a feature indicates a positive impact on prediction, leading the model to predict toxicity in this case). These contributing Cell Painting features were mapped to the BioMorph terms, along with the two most-contributing Cell Health phenotypes (level 3 of the BioMorph) and Cell process affected (level 4 of the BioMorph). We were able to identify relationships between specific compounds and their impact on cellular health (see Table 1 for a selection of illustrative compounds discussed below; and Supplemental Table S4 for the complete set of 54 compounds analysed). For example, for melatonin, an “apoptosis up” compound, we noted that the most contributing Cell Painting features were related to BioMorph terms for DNA damage

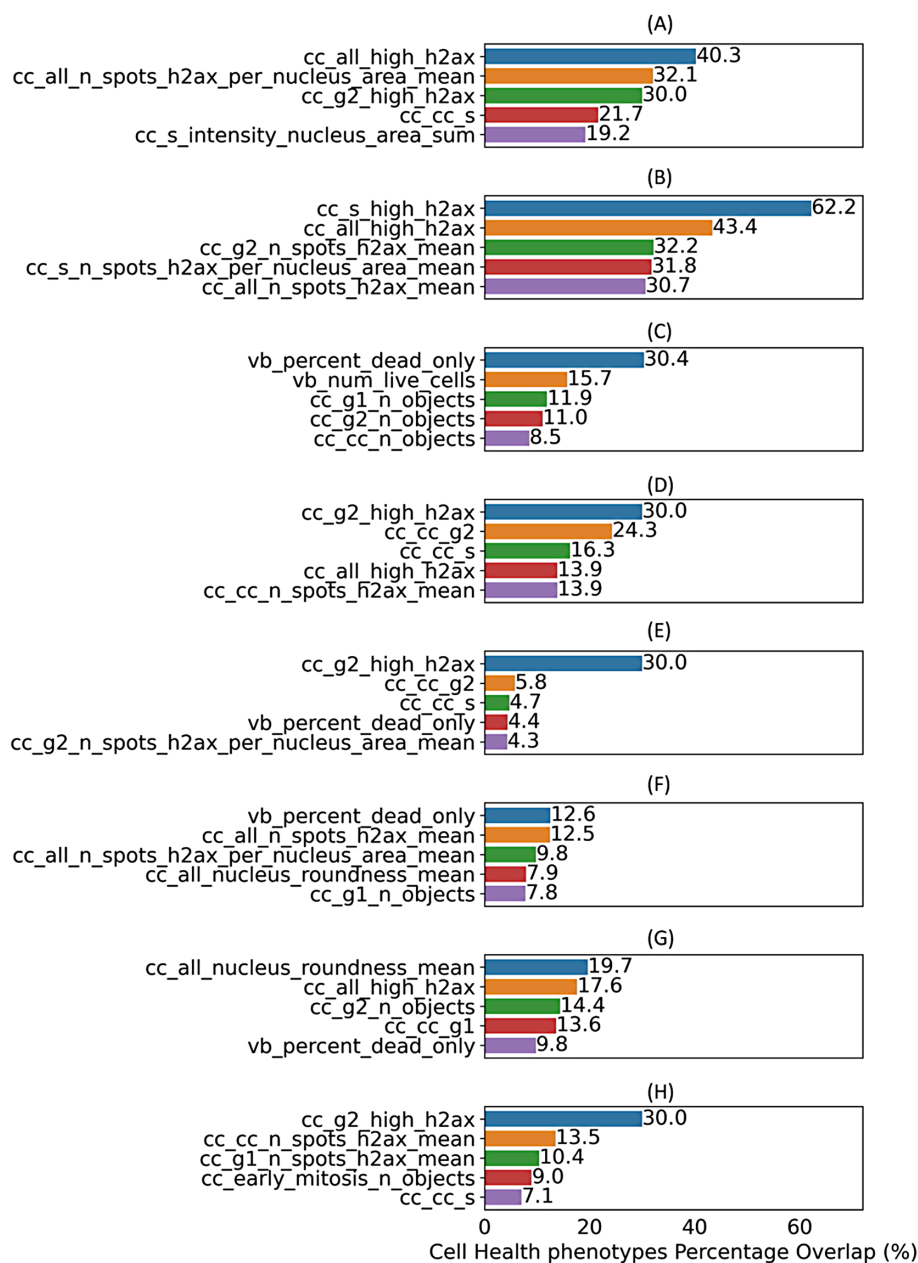


FIGURE 4: Top five specific Cell Health phenotypes (level 3) enriched by contributing Cell Painting features (as per feature importance) for each Random Forest model for eight different biological activities (a) apoptosis up, (b) cytotoxicity BLA, (c) cytotoxicity SRB, (d) ER stress, (e) heat shock, (f) microtubule upregulation, (g) oxidative stress, and (h) proliferation decrease. Models for mitochondrial disruption recorded AUC < 0.50 and were not interpreted.

(as indicated by the presence of more than three γ H2AX spots within the cells) and the fraction of cells arrested in the S phase, which is most likely due to increased ROS. In the case of melatonin, the effects on the cell cycle via ROS generation have been previously reported (Song *et al.*, 2018). In general, we observed that BioMorph space can help generate hypotheses to uncover secondary effects that might otherwise be overlooked, and examples listed in Table 1 and shown in Figure 6 speak to the granularity of the BioMorph space information. In the case of ER stressors, piromidic acid, clozapine, bisphenol A diglycidyl ether, and emetine, the top two most contributing Cell Health phenotypes and top two Cell processes affected were mostly different. This highlights that each compound

may exhibit the same bioactivity (e.g., “ER stress”) but cause it by affecting different targets/pathways and having distinct MOAs. The most contributing BioMorph terms for piromidic acid are related to cell viability (such as the number of cells and roundness of living cells); whereas emetine, a protein synthesis inhibitor, was linked to the fraction of cells in the S-phase of the cell cycle, which agrees with the secondary activity in early S-phase related to inhibition of DNA replication (Schweighoffer *et al.*, 1991). On the other hand, compounds linked to heat shock responses (alfadolone acetate, suxibuzone, and diflorasone) exhibited the same features and were associated with Hippo pathway-related terms, the roundness of the nucleus, and DNA damage in the S phase. This agrees with the established role of the Hippo pathway in promoting cell survival in response to various stressors (Di Cara *et al.*, 2015) while the shape of the nucleus (senescent cells can be characterized by flattened, enlarged or irregular-shape nuclei as shown by Zhao and Darzynkiewicz, 2013 and Heckenbach *et al.*, 2022) and vulnerability of early S-phase cells to mild genotoxic stress are common mechanisms of heat stress effects (Verbeke *et al.*, 2001; Velichko *et al.*, 2015). We also noted similarities among the level 3 and level 4 BioMorph space terms associated with compounds that cause proliferation decrease (raclopride, nimodipine, and ketanserin). These compounds are associated with hypoxia and apoptosis, suggesting that these compounds may act via increasing levels of ROS, which leads to oxidative stress (McGarry *et al.*, 2018). For the compounds causing an upregulation of microtubules, bifemelane was linked to BioMorph terms related to cell death as well as chromatin modifiers and DNA damage in the S phase consistent with its known role in enhancing the synthesis of cytoskeletal proteins (Asanuma *et al.*, 1993) and regulating dynamic chromosome organization (Spichal and Fabre, 2017). Taken together, we showcase how identifying the BioMorph terms having the greatest contribution to predicting

a compound’s biological activity, we can gain insights into not only primary but secondary biological processes affected by the compounds as well. These predictions can then be used to formulate mechanistic hypotheses and inform drug discovery and development efforts.

Limitations of mapping Cell Painting into BioMorph terms

This proof-of-concept study demonstrates the potential benefits of mapping Cell Painting features into BioMorph terms to address a serious challenge for the field of image-based profiling: making sense of complex combinations of image-based features that are not readily interpretable. We find that BioMorph does provide a more

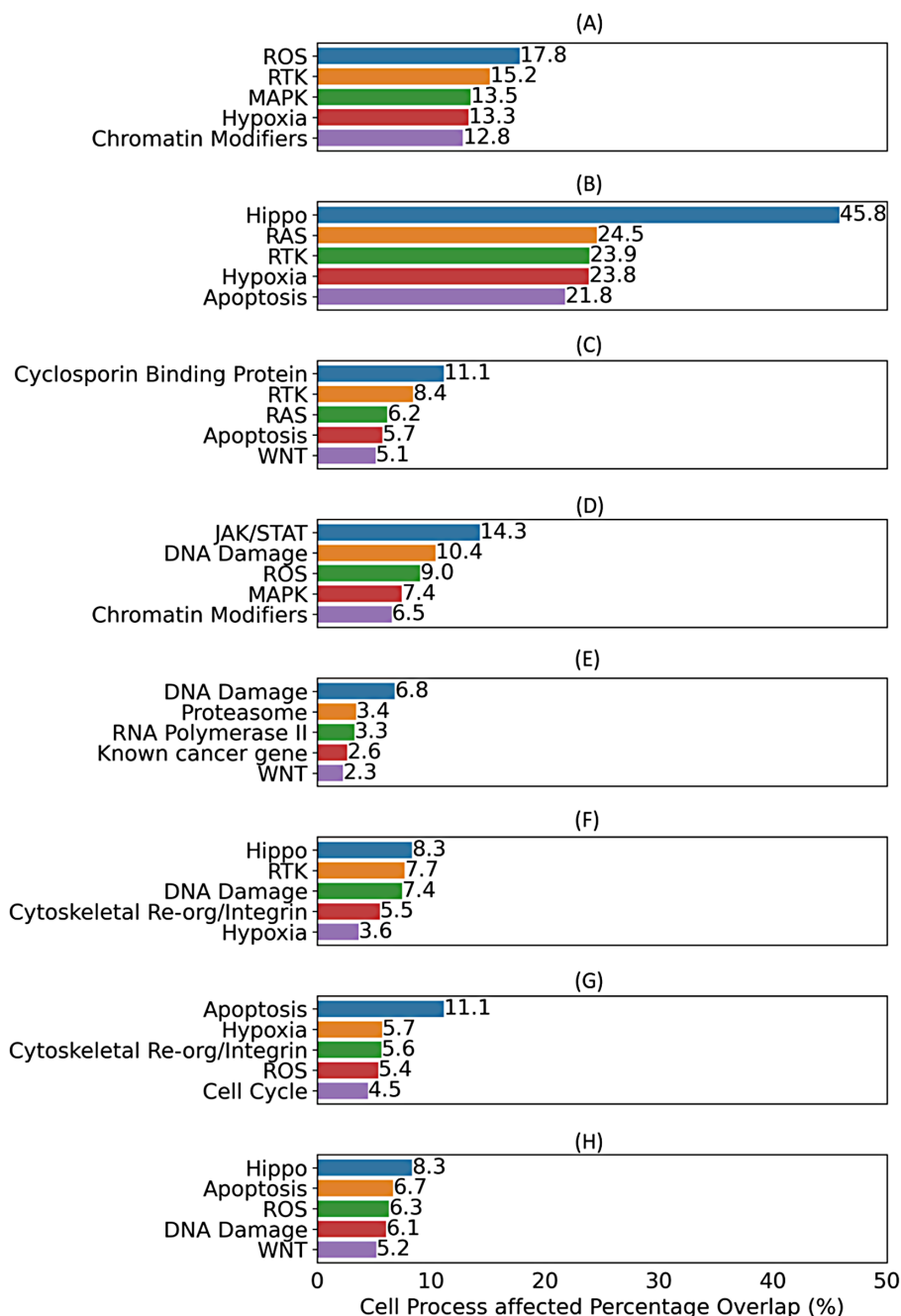


FIGURE 5: Top five Cell process affected (level 4) terms enriched by contributing Cell Painting features (as per feature importance) for each Random Forest model for eight different biological activities (a) apoptosis up, (b) cytotoxicity BLA, (c) cytotoxicity SRB, (d) ER stress, (e) heat shock, (f) microtubule upregulation, (g) oxidative stress, and (h) proliferation decrease. Models for mitochondrial disruption recorded AUC < 0.50 and were not interpreted.

interpretable and biologically relevant representation of data. However, there are several limitations relevant to this iteration of BioMorph space. BioMorph space was built using robust but limited data; therefore, using larger datasets of CRISPR perturbations and Cell Painting/Cell Health datasets would improve the organization of BioMorph space. Additionally, the associations between Cell Painting features and BioMorph terms are not absolute; these would need to be updated if alternative feature extraction strategies are used, for example, updated versions from CellProfiler (Stirling *et al.*, 2021) or

deep learning-based feature extraction (Pawlowski *et al.*, 2016; Caicedo *et al.*, 2022) such as in the JUMP-Cell Painting dataset (Chandrasekaran *et al.*, 2023), and we advise caution against using these groupings directly if the feature extractions differ from the current study. Finally, we evaluated our BioMorph space for nine broad biological activities; generalization to other cellular mechanisms and biological processes would require assays that are focused on other readouts, such as those related to particular types of toxicity, or tailored to particular cell types like neurons or cardiomyocytes. Despite these limitations, the study introduces an algorithm to map Cell Painting features into BioMorph terms and explores the application of this new BioMorph space in interpreting predictive models, generating hypotheses for small molecule biological activity, MOA, and toxicity.

SIGNIFICANCE

In this work, we demonstrated a strategy to map Cell Painting features into BioMorph terms to enable a better understanding of the relationships between compound-induced cellular perturbations and nine different biological activities. We could correctly identify potential secondary mechanisms of biological activities such as ER stress and cell cycle arrest at the G₂ phase (Bourougaa *et al.*, 2010) as well as mechanisms of action of dual-function compounds such as emetine, which is a well-known protein synthesis inhibitor, but also acts at an early S-phase to inhibit DNA replication (Schweighoffer *et al.*, 1991). These are biological effects that can often be overlooked; however, the BioMorph space allows for a more comprehensive understanding of these mechanisms, for uncovering hidden relationships and generating new hypotheses by connecting them to specific phenotypes and cellular processes.

Recently, overwhelming evidence has accumulated for the strong performance of deep learning methods over classical features for computer vision tasks, such as microscopy image segmentation and classification (Lafarge *et al.*, 2019; Chow *et al.*, 2022; Moshkov *et al.*, 2022; Wong *et al.*, 2023). Hofmarcher *et al.* (2019) demonstrated that for bioactivity prediction,

CNNs trained directly on image data outperformed fully connected neural networks that relied on computed CellProfiler features. The study attributed this improvement to better cell segmentation, sparse signal detection, and single-cell level image analysis when using CNN models directly on imaging data compared with CellProfiler features that rely on aggregate statistics. In the more specific task of identifying relationships among reagents using image-based profiling, extracting features using deep learning has recently begun to pull ahead of classically defined features.

Common name	Biological Process	Specific Cell Health phenotypes (most impacted)	Specific Cell Health phenotypes (2nd most impacted)	Cell process affected (most)	Cell process affected (2nd most)
Melatonin	apoptosis up	The fraction of cells containing more than three gH2AX spots within all cells:	Fraction of EdU positive cells (S-phase of the cell cycle)	ROS	MAPK
Piromidic acid	ER stress	Total number of cells	Cell Roundness	RTK	ER Stress/UPR
Clozapine	ER stress	Fraction of G1 cells	Fraction of caspase negative in dead cells	Chromatin Modifiers	ER Stress/UPR
Bisphenol A diglycidyl ether	ER stress	Fraction of G2 cells	Total number of cells	JAK/STAT	WNT
Emetine	ER stress	Fraction of EdU positive cells (S-phase of the cell cycle)	The fraction of cells containing more than three gH2AX spots within all cells	ROS	Chromatin Modifiers
Alfadolone acetate	heat shock	Average nucleus roundness	The fraction of >3 γ H2Ax spots in S phase cells	Hippo	Chromatin Modifiers
Suxibuzone	heat shock	Average nucleus roundness	The fraction of >3 γ H2Ax spots in S phase cells	Hippo	Chromatin Modifiers
Diflorasone diacetate	heat shock	Average nucleus roundness	The fraction of >3 γ H2Ax spots in S phase cells	Hippo	Chromatin Modifiers
Bifemelane	microtubule up	Fraction of caspase negative in dead cells	EdU incorporated (average intensity per cell) in S phase cells	Hippo	Chromatin Modifiers
Raclopride	proliferation decrease	Fraction of caspase negative in dead cells	Width/Length	Hypoxia	Apoptosis
Nimodipine	proliferation decrease	Fraction of caspase negative in dead cells	Width/Length	Hypoxia	Apoptosis
Ketanserin	proliferation decrease	Number of G1 cells	Fraction of caspase negative in dead cells	Hypoxia	Apoptosis

TABLE 1: Top two contributing Cell Health phenotypes (level 3) and Cell process affected (level 4) from BioMorph space for a selection of illustrative true positives predicted by the models for biological activity. See Supplemental Table S4 for the complete set of 54 compounds.

Recent studies employed innovative training strategies that have further boosted deep learning performance by as much as 29% compared with CellProfiler features when evaluated based on mean average precision (mAP) for classifying chemical perturbations (Kim *et al.*, 2023). Most recently, because our study was completed, both convolutional neural networks and vision transformers-based masked autoencoders were shown to outperform weakly supervised models (Kraus *et al.*, 2023; Wong *et al.*, 2023). Remarkably, some of these models achieved performance improvements of up to 28% in deducing established biological relationships in image-based data based on ground truth annotations from databases like StringDB and Reactome (Kraus *et al.*, 2023). Interpretability in machine learning models using image data is often crucial for biologists who use such models to understand the cause of these predictions such as in understanding mechanisms of compound toxicity in drug discovery (Dara *et al.*, 2022). Interpreting deep learning-extracted features is an active area of research (Selvaraju *et al.*, 2016; Wong *et al.*, 2022). We recognized the potential challenges in interpreting biological meaning for CellProfiler features (Lundberg *et al.*, 2021), and in this work, we aimed to improve the interpretability of these features by defining a BioMorph space for them. BioMorph is applied to enhance the clarity and comprehensibility of classical image features derived from CellProfiler (Stirling *et al.*, 2021; the most commonly available features in

public and private Cell Painting data) while retaining the potential to be applied to other features, such as those extracted by deep learning. Currently, there are several deep learning-based feature extractor protocols such as CNN-based feature extraction (Steigele *et al.*, 2020), DeepProfiler (cytomining/DeepProfiler: Morphological profiling using deep learning), and WS-DINO (Cross-Zamirski *et al.*, 2022) among others. In the future, as a standardized protocol/software for extracting features via deep learning becomes more established across industry and academia, these features might be integrated with Cell Health assays to form a BioMorph space to enhance the comprehension of the biological insights embedded within deep learning-derived features.

Mapping Cell Painting features into BioMorph terms offers several advantages over using CellProfiler-derived Cell Painting features directly. First, we improved interpretability by using a more biologically interpretable feature space; we identified relationships between compound mechanisms of action and their impact on cell morphology. For example, the use of BioMorph space identified relevant pathways such as the JAK/STAT signalling pathway's prominence in ER stress (Mearns *et al.*, 2014). These insights are not possible with Cell Painting features alone, which have no information on biological pathways. Second, we could pinpoint the specific cell processes and stages of the cell cycle affected by a compound, a task not possible with the Cell Painting features, which do not

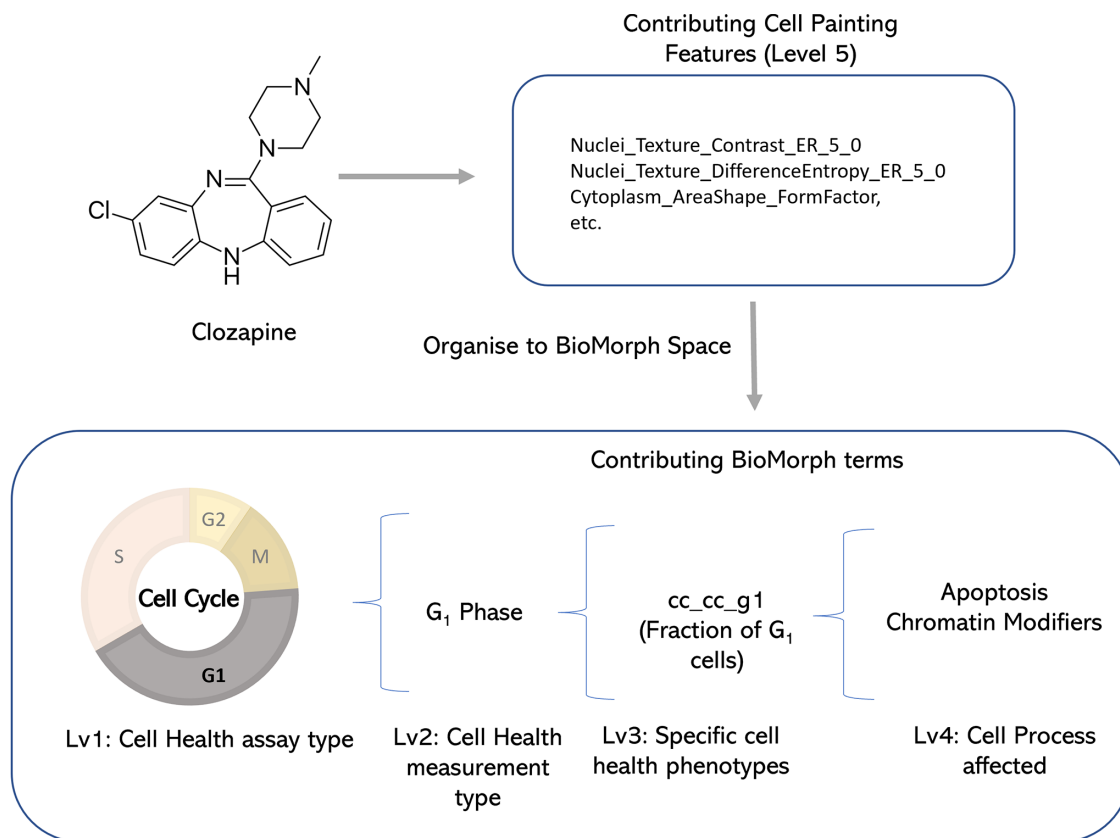


FIGURE 6: For the compound clozapine, which is an ER stressor, SHAP values indicate a list of the most-contributing Cell Painting features (Level 5) to model performance for ER stress. Organising this to BioMorph terms allows interpretation: clozapine can induce cell cycle arrest in the G₀/G₁ phase.

contain direct information on which cell cycle stage is impacted. Finally, we could facilitate hypothesis generation by identifying the BioMorph terms that contribute most significantly to compound activity. These targeted hypotheses can guide the future validation of compounds. Taken together, the BioMorph space represents a more integrative and comprehensive method for analysing cellular MOA and can enable the development of more effective strategies for identifying and mitigating toxic effects.

MATERIALS AND METHODS

Cell Painting Dataset for CRISPR Perturbations

We used the Cell Painting pilot dataset of (CRISPR) knockout perturbations from the Broad Institute (Way *et al.*, 2021). Here, the authors used a Cell Painting assay for three different cell lines (A549, ES2, and HCC44) and each cell line used 357 perturbations representing 119 clustered regularly interspersed short palindromic repeats (CRISPR) knockout perturbations (further details in Supplemental Table S5). They further generated median consensus signatures for each of the 357 perturbations. This led to a dataset of 949 morphology features (and metadata annotations) for 357 consensus profiles (119 CRISPR perturbations × 3 cell lines). Among these, only 827 Cell Painting features were in intersection with the Cell Painting dataset for compound perturbations (described below) used in this proof-of-concept study. The Cell Painting dataset for CRISPR Perturbations is released publicly at <https://zenodo.org/records/10011861>.

Cell Health assays for CRISPR Perturbations

We used the Cell Health assay developed by the Broad Institute containing 70 specific Cell Health phenotypes (Way *et al.*, 2021).

The authors used seven reagents in two Cell Health panels to stain cells for the same 119 CRISPR perturbations for three different cell lines (A549, ES2, and HCC44). We used median consensus signatures for the 357 consensus profiles (119 CRISPR perturbations × 3 cell lines) as above. This dataset is released publicly at <https://zenodo.org/records/10011861>.

Cell Painting Dataset for Compound Perturbations

The Cell Painting assay used in this proof-of-concept study, from the Broad Institute, contains cellular morphological profiles of more than 30,000 small molecule perturbations (Bray *et al.*, 2017). The morphological profiles in this dataset are composed of a wide range of feature measurements (shape, area, size, correlation, texture, etc.). The authors in this study normalized morphological features to compensate for variations across plates and further excluded features having a zero median absolute deviation (MAD) for all reference cells in any plate. Following the procedure from Lapins and Spjuth (2019), we subtracted the average feature value of the neutral DMSO control from the compound perturbation average feature value on a plate-by-plate basis. We standardised the InChI (International Chemical Identifier) [Goodman *et al.*, 2021] using RDKit (RDKit) and for each compound and drug combination, we calculated a median feature value. Where the same compound was replicated for different doses, we used the median feature value across all doses that were within one SD of the mean dose. Finally, we obtained 1783 median Cell Painting features for 30,404 unique compounds. This dataset is publicly released at <https://broad.io/biomorph>. Among these, only 827 Cell Painting features were common with the dataset for CRISPR Perturbations which were used in this proof-of-concept study.

Biological activity from ToxCast assay with Cell Painting annotations

Toxicity and biological activity-related data were collected from 56 cytotoxicity and cell stress response assays from 56 ToxCast (Exploring ToxCast Data | US EPA; Wu *et al.*, 2018) for nine broad biological processes (for the mapping between 56 ToxCast assays and nine biological processes see Judson *et al.*, 2016): apoptosis up, cytotoxicity BLA, cytotoxicity SRB, ER stress, heat shock, microtubule upregulation, mitochondrial disruption up, oxidative stress up, and proliferation decrease (Judson *et al.*, 2016). Compound SMILES were converted to standardised InChI using RDKit (RDKit). To generate consensus endpoint labels, the presence of positive activity (toxicity) in at least one assay related to the biological activity was considered sufficient to mark the compound active in the consensus endpoint. Thus, consensus endpoints for each of the nine biological activities were generated from the 56 ToxCast assays. We calculated the intersection of the Cell Painting profiles for compound perturbations (above) and nine biological activity (ToxCast) assays using the standardised InChI. Cell Painting features were standardised by removing the mean and scaling to unit variance. This resulted in a complete dataset of 658 structurally unique compounds with 827 Cell Painting features and nine biological activity consensus hit calls that were used in this proof-of-concept study. The dataset, referred to as containing biological activities in this study, is publicly released at <https://zenodo.org/records/10011861>.

Mapping Cell Painting terms into BioMorph space

The overlap of Cell Painting and Cell Health assay for gene perturbations (Way *et al.*, 2021) contained 827 Cell Painting features (that were also present in the Cell Painting experiments on compound perturbations from Bray *et al.*, 2017) and 70 continuous Cell Health endpoints (e.g., the number of late polynuclear cells, which measures the shape in a cell cycle assay) for 354 consensus profiles (118 CRISPR perturbations \times 3 cell lines, the empty well was removed). As shown in Figure 2 step A, for feature selection, we used an all-relevant feature selection method, Borutapy (Kursa and Rudnicki, 2010) implemented using the Python package Boruta (Boruta · PyPI) with a Random Forest Classifier estimator of maximum depth 5 and the number of estimators determined automatically based on the size of the dataset using “auto”. Using Borutapy, we detected a subset of Cell Painting features that contain information for each of the 70 Cell Health regression labels. Further, we trained a baseline Linear Regression model as implemented in scikit-learn (scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation; Figure 2, step B) with an 80–20 random train-test split to predict which subsets of Cell Painting features are relatively better predictors of Cell Health phenotype. Thirty-seven of the 70 Cell Health models (with $R^2 > 0.25$) were selected for further analysis. This results in 34 subsets of Cell Painting features (one set for each of the 34 Cell Health labels). Next, for each of the 34 Cell Health labels and the 354 consensus profiles, we separated subsets of Cell Painting data for the negative control CRISPR Perturbation (which consisted of 30 datapoints of LacZ, Luc, and Chr2 CRISPR perturbations) and other CRISPR perturbations affecting various known cell processes (such as chromatin modifiers, ER Stress/UPR, metabolism, etc.). For each of these pairs (negative control and CRISPR perturbations), we used Borutapy (Boruta · PyPI; Figure 2, step C) to detect a further subset from the subset of Cell Painting features which contained a signal on whether the datapoint is a negative control or the CRISPR Perturbation. We train a baseline Random Forest Classifier (Figure 2, step D), as implemented in scikit-learn (scikit-learn: machine learning in Python – scikit-learn

1.2.0 documentation), with an 80–20 random train test split to predict which sets of selected Cell Painting features perform relatively better at differentiating negative controls from the CRISPR perturbation (MCC > 0.50). This led to 412 subsets of informative Cell Painting features which are then indicators of 412 BioMorph terms. We used a χ^2 test to determine the BioMorph term p value for each of the 412 combinations (Figure 2, step E) from standard scaled subsets of Cell Painting features.

Further using this mapping, any dataset with Cell Painting features can be mapped into BioMorph terms. The dataset of biological activities with 827 Cell Painting features were grouped into these 412 combinations and their BioMorph term p value was calculated. We then standardised these BioMorph terms using a standard scalar (as implemented in scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation), and only columns with noninfinite continuous p values were retained (with other columns dropped). This resulted in 398 BioMorph terms for the biological activity dataset. The dataset is now released at <https://zenodo.org/records/10011861>. For the BioMorph dataset for all 30,000 compounds, please see <https://broad.io/BioMorph>.

Comparing models using Cell Painting and BioMorph terms as features

To ensure that the BioMorph terms contain all information from the original Cell Painting readouts, we compared models using only 827 Cell Painting features and models using the 398 BioMorph terms directly as features (although there were 412 terms defined, only 398 terms out of these were non-infinite and continuous and used for modelling). For each of the nine biological activities, we used five times repeated fourfold nested cross-validation and a Random Forest Classifier (as implemented in scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation). First, the data was split into four folds using a stratified split on biological activity labels where 25% of the data was reserved for the test set and 75% remaining used for training. Using this training data, we trained two models, one using the 827 Cell Painting features, and the other using 398 BioMorph terms (p values from subsets of Cell Painting features; although there were 412 terms defined, only 398 terms out of these were noninfinite and continuous and used for modelling). We optimised these models using a fivefold cross-validation with stratified splits and a random halving search algorithm (with hyperparameter space given in Supplemental Table S6 and as implemented in scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation). The optimised model was fit on the entire training data and cross-validation predictions are used to determine the optimal threshold using the J statistic value (Youden, 1950). We then used this threshold to determine the predictions for the test set predictions. A single loop of nested cross-validation results in four test sets, which are repeated five times thus giving 20 individual test set predictions.

Model training with Cell Painting features

To evaluate the use of BioMorph space, we now used a fixed held-out test set. For each of the nine biological activities, we used a stratified split on biological activity labels such that 75% of the data was used in cross-validation training and 25% as held-out test data. We trained Random Forest classifiers (as implemented in scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation) using 827 Cell Painting features and a random halving search algorithm (as implemented in scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation) to optimise the hyperparameters (with the hyperparameter space given in Supplemental Table S6). Similar to above, the optimised model was fit on the

entire training data and cross-validation predictions are used to determine the optimal threshold using the J statistic value that considers both true and false positive rates. This optimal threshold is then used on the predicted probabilities of the held-out test data to obtain the final held-out test data predictions.

Feature importance and interpretation in BioMorph terms

First, we used feature importance from the Random Forest classifier (as implemented in scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation) to determine the features that contributed the most to model importance. This gave us important features per biological activity (at an endpoint/biological activity level). Second, we evaluated SHAP values (Lundberg and Lee, 2017), as implemented in the shap (Scott Lundberg, 2018) python package, for each compound predicted as true positive in the held-out test set. We used true positives only, as these are the predictions for which the feature importance value (from SHAP) is valid. This gave us the important features per toxic compounds in the held-out test set for each biological activity (at a compound level). We then selected the Cell Painting features (from model importance values at the endpoint level or SHAP values at a compound level) that were greater than two standard deviations of all features as the most important or contributing features. These features were mapped into the BioMorph space by determining whether the features related to the individual levels of the BioMorph term were present among the important features selected above. At the level of Cell process affected (level 4), the percentage enrichment was determined as the percentage of Cell Painting features that were present among the defined subset of Cell Painting features (level 5). For an overall enrichment value (used for Figure 3 and Figure 4) for each specific Cell Health phenotype term (level 4) or Cell process affected (level 5), we used the mean of enrichment of all BioMorph terms where the corresponding level 3 or level 4 term appeared. For detailed enrichment analysis, we determined enrichment of the level (lv_x) to be the percentage of the immediate lower level (lv_{x-1}) with enrichment $\geq 10\%$ progressively from specific Cell Health phenotypes (level 3) to Cell Health assay type (level 1). This is released per biological activity in Supplemental Table S3.

Evaluation Metrics

To evaluate models in this proof-of-concept study we used Balanced Accuracy which considers both sensitivity and specificity, the AUC-Receiver Operating Characteristic and Mathew's correlation constant (MCC) as implemented in scikit-learn (scikit-learn: machine learning in Python – scikit-learn 1.2.0 documentation).

Statistics and Reproducibility

We have released the datasets used in this proof-of-concept study which are publicly available at <https://broad.io/biomorph> and <https://zenodo.org/records/10011861>. We released the Python code for the models which are publicly available at https://github.com/srijitseal/BioMorph_Space.

ACKNOWLEDGMENTS

S.S. acknowledges funding from the Cambridge Centre for Data-Driven Discovery (C2D3) Accelerate Programme for Scientific Discovery. A.E.C. acknowledges funding from the National Institutes of Health (R35 GM122547). O.S. acknowledges funding from the Swedish Research Council (grants 2020-03731 and 2020-01865), FORMAS (grant 2022-00940), Swedish Cancer Foundation (22 2412 Pj 03 H), and Horizon Europe grant agreement #101057014 (PARC) and #101057442 (REMEDIA4ALL).

REFERENCES

- Asanuma M, Ogawa N, Hirata H, Chou HH, Kondo Y, Mori A (1993). Ischemia-induced changes in α -tubulin and β -actin mRNA in the gerbil brain and effects of bifemelane hydrochloride. *Brain Res* 600, 243–248.
- Barascu A, Le Chalony C, Pennarun G, Genet D, Imam N, Lopez B, Bertrand P (2012). Oxidative stress induces an ATM-independent senescence pathway through p38 MAPK-mediated lamin B1 accumulation. *EMBO Journal* 31, 1080–1094.
- Basili D, Reynolds J, Houghton J, Malcomber S, Chambers B, Liddell M, Muller I, White A, Shah I, Everett L-J, et al. (2022). Latent variables capture pathway-level points of departure in high-throughput toxicogenomic data. *Chem Res Toxicol* 35, 670–683.
- Bourougaa K, Naski N, Boularan C, Mlynarczyk C, Candeias MM, Marullo S, Fähræus R (2010). Endoplasmic reticulum stress induces G2 cell-cycle arrest via mRNA translation of the p53 Isoform p53/47. *Mol Cell* 38, 78–88.
- Bray MA, Gustafsdottir SM, Rohban MH, Singh S, Ljosa V, Sokolnicki KL, Bittker JA, Bodycombe NE, Dančik V, Hasaka TP, et al. (2017). A dataset of images and morphological profiles of 30 000 small-molecule treatments using the Cell Painting assay. *Gigascience* 6, 1–5.
- Bray M-A, Singh S, Han H, Davis CT, Borgeson B, Hartland C, Kost-Alimova M, Gustafsdottir SM, Gibson CC, Carpenter AE (2016). Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat Protoc* 11, 1757–1774.
- Brewer JW, Hendershot LM, Sherr CJ, Diehl JA (1999). Mammalian unfolded protein response inhibits cyclin D1 translation and cell-cycle progression. *Proc Natl Acad Sci USA* 96, 8505–8510.
- Caicedo JC, Arevalo J, Piccioni F, Bray MA, Hartland CL, Wu X, Brooks AN, Berger AH, Boehm JS, Carpenter AE, Singh S (2022). Cell Painting predicts impact of lung cancer variants. *Mol Biol Cell* 33, ar49.
- Di Cara F, Maile TM, Parsons BD, Magico A, Basu S, Tapon N, King-Jones K (2015). The Hippo pathway promotes cell survival in response to chemical stress. *Cell Death Differ* 22, 1526–1539.
- Chandrasekaran SN, Ackerman J, Alix E, Ando DM, Arevalo J, Bennion M, Boisseau N, Borowa A, Boyd JD, Brino L, et al. (2023). JUMP Cell Painting dataset: morphological impact of 136,000 chemical and genetic perturbations. *BioRxiv*, <https://doi.org/10.1101/2023.03.23.534023>.
- Chandrasekaran SN, Ceulemans H, Boyd JD, Carpenter AE (2021). Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nat Rev Drug Discov* 20, 145–159.
- Chessel A, Carazo Salas RE (2019). From observing to predicting single-cell structure and function with high-throughput/high-content microscopy. *Essays Biochem* 63, 197–208.
- Chow YL, Singh S, Carpenter AE, Way GP (2022). Predicting drug polypharmacology from cell morphology readouts using variational autoencoder latent space arithmetic. *PLoS Comput Biol* 18, e1009888.
- Clark NR, Szymkiewicz M, Wang Z, Monteiro CD, Jones MR, Ma'ayan A (2015). Principle Angle Enrichment Analysis (PAEA): Dimensionally reduced multivariate gene set enrichment analysis tool. *Proceedings (IEEE Int Conf Bioinformatics Biomed)* 2015, 256–262.
- Cross-Zamirski JO, Williams G, Mouchet E, Schönlieb C-B, Turkki R, Wang Y (2022). Self-Supervised Learning of Phenotypic Representations from Cell Images with Weak Labels. *arXiv*, <https://doi.org/10.48550/arXiv.2209.07819>.
- Dara S, Dhamecherla S, Jadav SS, Babu CM, Ahsan MJ (2022). Machine learning in drug discovery: A review. *Artif Intell Rev* 55, 1947–1999.
- Goodman JM, Pletnev I, Thiessen P, Bolton E, Heller SR (2021). InChI version 1.06: now more than 99.99% reliable. *J Cheminform* 13, 40.
- Gustafsdottir SM, Ljosa V, Sokolnicki KL, Wilson JA, Walpita D, Kemp MM, Seiler KP, Carrel HA, Golu TR, Schreiber SL, et al. (2013). Multiplex cytological profiling assay to measure diverse cellular states. *PLoS One* 8, e80999.
- Heckenbach I, Mkrtchyan GV, Ezra MB, Bakula D, Madsen JS, Nielsen MH, Oró D, Osborne B, Covarrubias AJ, Idda ML, et al. (2022). Nuclear morphology is a deep learning biomarker of cellular senescence. *Nat Aging* 2, 742–755.
- Hofmarcher M, Rumetshofer E, Clevert D-A, Hochreiter S, Klambauer G (2019). Accurate prediction of biological assays with high-throughput microscopy images and convolutional networks. *J Chem Inf Model* 59, 1163–1171.
- Howard PL, Chia MC, Del Rizzo S, Liu FF, Pawson T (2003). Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proc Natl Acad Sci USA* 100, 11267–11272.
- Hung JH, Yang TH, Hu Z, Weng Z, DeLisi C (2012). Gene set enrichment analysis: Performance evaluation and usage guidelines. *Brief Bioinform* 13, 281–291.

- Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I, Little S, Wambaugh J, Setzer RW, Koehya P, et al. (2016). Analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space. *Toxicological Sciences* 152, 323–339.
- Kantidze OL, Velichko AK, Luzhin AV, Razin SV (2016). Heat stress-induced DNA damage. *Acta Naturae* 8, 75–78.
- Kim JM (2022). Molecular link between DNA damage response and microtubule dynamics. *Int J Mol Sci* 23, 6986.
- Kim V, Adaloglou N, Osterland M, Morelli FM, Zapata PAM (2023). Self-supervision advances morphological profiling by unlocking powerful image representations. *BioRxiv*, <https://doi.org/10.1101/2023.04.28.538691>.
- Kraus O, Kenyon-Dean K, Saberian S, Fallah M, McLean P, Leung J, Sharma V, Khan A, Balakrishnan J, Celik S, et al. (2023). Masked autoencoders are scalable learners of cellular morphology. *arXiv*, <https://doi.org/10.48550/arXiv.2309.16064>.
- Kühl NM, Rensing L (2000). Heat shock effects on cell cycle progression. *Cell Mol Life Sci* 57, 450–463.
- Kursa MB, Rudnicki WR (2010). Feature selection with the boruta package. *J Stat Softw* 36, 1–13.
- Lafarge MW, Caicedo JC, Carpenter AE, Pluim JPW, Singh S, Veta M (2019). Capturing Single-Cell Phenotypic Variation via Unsupervised Representation Learning. *Proc Mach Learn Res* 102, 315–325.
- Lapins M, Spjuth O (2019). Evaluation of Gene Expression and Phenotypic Profiling Data as Quantitative Descriptors for Predicting Drug Targets and Mechanisms of Action. *BioRxiv*, <https://doi.org/10.1101/580654>.
- Liu A, Seal S, Yang H, Bender A (2023). Using chemical and biological data to predict drug toxicity. *SLAS Discovery* 28, 53–64.
- Lundberg E, Funke J, Uhlmann V, Gerlich D, Walter T, Carpenter A, Coelho LP (2021). Which image-based phenotypes are most promising for using AI to understand cellular functions and why? *Cell Syst* 12, 384–387.
- Lundberg SM, Lee SI (2017). A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017, 4766–4775.
- Markowitz F (2010). How to understand the cell by breaking it: network analysis of gene perturbation screens. *PLoS Comput Biol* 6, e1000655.
- McGarry T, Biniacka M, Veale DJ, Fearon U (2018). Hypoxia, oxidative stress and inflammation. *Free Radic Biol Med* 125, 15–24.
- Meares GP, Liu Y, Rajbhandari R, Qin H, Nozell SE, Mobley JA, Corbett JA, Benveniste EN (2014). PERK-dependent activation of jak1 and stat3 contributes to endoplasmic reticulum stress-induced inflammation. *Mol Cell Biol* 34, 3911–3925.
- Moshkov N, Bornholdt M, Benoit S, Smith M, McQuin C, Goodman A, Senft RA, Han Y, Babadi M, Horvath P, et al. (2022). Learning representations for image-based profiling of perturbations. *BioRxiv*, <https://doi.org/10.1101/2022.08.12.503783>.
- Nyffeler J, Willis C, Lougee R, Richard A, Paul-Friedman K, Harrill JA (2020). Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling. *Toxicol Appl Pharmacol* 389, 114876.
- Parvande S, Yeh HW, Paulus MP, McKinney BA (2020). Consensus features nested cross-validation. *Bioinformatics* 36, 3093–3098.
- Pawlowski N, Caicedo JC, Singh S, Carpenter AE, Storkey A (2016). Automating Morphological Profiling with Generic Deep Convolutional Networks. *BioRxiv*, <https://doi.org/10.1101/085118>.
- Pruteanu LL, Bender A (2023). Using transcriptomics and cell morphology data in drug discovery: The long road to practice. *ACS Med Chem Lett* 14, 386–395.
- Redza-Dutordoir M, Averill-Bates DA (2016). Activation of apoptosis signalling pathways by reactive oxygen species. *Biochim Biophys Acta Mol Cell Res* 1863, 2977–2992.
- Riss TL, Moravec RA, Niles AL, Duellman S, Benink HA, Worzella TJ, Minor L (2016). *Cell Viability Assays*. Assay Guidance Manual.
- Sanchez-Fernandez A, Rumetshofer E, Hochreiter S, Klambauer G (2023). CLOOME: contrastive learning unlocks bioimaging databases for queries with chemical structures. *Nature Communications* 14, 7339.
- Schweighoffer T, Schweighoffer E, Apati A, Antoni F, Molnar G, Lapis K, Banfalvi G (1991). Cytometric analysis of DNA replication inhibited by emetine and cyclosporin A. *Histochem* 96, 93–97.
- Scott Lundberg (2018). Welcome to the SHAP documentation. Available at: <https://shap.readthedocs.io/en/latest/> (accessed 16 May 2023).
- Seal S, Carreras-Puigvert J, Trapotsi M-A, Yang H, Spjuth O, Bender A (2022). Integrating cell morphology with gene expression and chemical structure to aid mitochondrial toxicity detection. *Commun Biol* 5, 858.
- Seal S, Spjuth O, Hosseini-Gerami L, Garcia-Ortegon M, Singh S, Bender A, Carpenter AE (2023a). Insights into Drug Cardiotoxicity from Biological and Chemical Data: The First Public Classifiers for FDA DICTrank. *BioRxiv*, <https://doi.org/10.1101/2023.10.15.562398>.
- Seal S, Yang H, Trapotsi M-A, Singh S, Carreras-Puigvert J, Spjuth O, Bender A (2023b). Merging bioactivity predictions from cell morphology and chemical fingerprint models using similarity to training data. *J Cheminform* 15, 56.
- Seal S, Yang H, Vollmers L, Bender A (2021). Comparison of cellular morphological descriptors and molecular fingerprints for the prediction of cytotoxicity- and proliferation-related assays. *Chem Res Toxicol* 34, 422–437.
- Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2016). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int J Comput Vis* 128, 336–359.
- Shao Y, Gao Z, Marks PA, Jiang X (2019). Apoptotic and autophagic cell death induced by histone deacetylase inhibitors. *Proc Natl Acad Sci USA* 101, 18030–18035.
- Simm J, Unter Klambauer G, Arany A, Steijaert M, Wegner JK, Gustin E, Chupakhin V, Chong YT, Vialard J, Buijsters P, et al. (2018). Repurposing high-throughput image assays enables biological activity prediction for drug discovery cell chemical biology resource repurposing high-throughput image assays enables biological activity prediction for drug discovery. *Cell Chem Biol* 25, 611–618.e3.
- Song J, Ma SJ, Luo JH, Zhang H, Wang RX, Liu H, Li L, Zhang ZG, Zhou RX (2018). Melatonin induces the apoptosis and inhibits the proliferation of human gastric cancer cells via blockade of the AKT/MDM2 pathway. *Oncol Rep* 39, 1975–1983.
- Spichal M, Fabre E (2017). The emerging role of the cytoskeleton in chromosome dynamics. *Front Genet* 8, 60.
- Steigele S, Siegismund D, Fassler M, Kustec M, Kappler B, Hasaka T, Yee A, Brodte A, Heyse S (2020). Deep learning-based hcs image analysis for the enterprise. *SLAS Discovery* 25, 812–821.
- Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimino BA, Goodman A (2021). CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* 22, 1–11.
- Szalai B, Subramanian V, Holland CH, Alföldi R, Puskás LG, Saez-Rodriguez J (2019). Signatures of cell death and proliferation in perturbation transcriptomics data—from confounding factor to effective prediction. *Nucleic Acids Res* 47, 10010–10026.
- Tait SWG, Green DR (2008). Caspase-independent cell death: Leaving the set without the final cut. *Oncogene* 27, 6452–6461.
- Trapotsi MA, Mouchet E, Williams G, Monteverde T, Juhani K, Turkki R, Miljkovic F, Martinsson A, Mervin L, Pryde KR, et al. (2022). Cell Morphological profiling enables high-throughput screening for proteolysis targeting chimera (PROTAC) phenotypic signature. *ACS Chem Biol* 17, 1733–1744.
- Trapotsi MA, Mervin LH, Afzal AM, Sturm N, Engkvist O, Barrett IP, Bender A (2021). Comparison of chemical structure and cell morphology information for multitask bioactivity predictions. *J Chem Inf Model* 61, 1444–1456.
- Velichko AK, Petrova NV, Razin SV, Kantidze OL (2015). Mechanism of heat stress-induced cellular senescence elucidates the exclusive vulnerability of early S-phase cells to mild genotoxic stress. *Nucleic Acids Res* 43, 6309.
- Verbeke P, Fonager J, Clark BFC, Rattan SIS (2001). Heat shock response and ageing: Mechanisms and applications. *Cell Biol Int* 25, 845–857.
- Wang JYJ (2001). DNA damage and apoptosis. *Cell Death Differ* 8, 1047–1048.
- Wang W, Xiao ZD, Li X, Aziz KE, Gan B, Johnson RL, Chen J (2015). AMPK modulates Hippo pathway activity to regulate energy homeostasis. *Nat Cell Biol* 17, 490.
- Wang Z, Clark NR, Ma'ayan A (2016). Drug-induced adverse events prediction with the LINCS L1000 data. *Bioinformatics* 32, 2338–2345.
- Way GP, Kost-Alimova M, Shibue T, Harrington WF, Gill S, Piccioni F, Becker T, Shafiqat-Abbasi H, Hahn WC, Carpenter AE, et al. (2021). Predicting cell health phenotypes using image-based morphology profiling. *Mol Biol Cell* 32, 995–1005.
- Way GP, Natoli T, Adeboye A, Litichevskiy L, Yang A, Lu X, Caicedo JC, Cimino BA, Karhohs K, Logan DJ, et al. (2022). Morphology and gene expression profiling provide complementary information for mapping cell state. *Cell Syst* 13, 911–923.e9.
- Wong DR, Logan DJ, Hariharan S, Stanton R, Clevert DA, Kiruluta A (2023). Deep representation learning determines drug mechanism of action from cell painting images. *Digital Discov* 2, 1354–1367.
- Wong KS, Zhong X, Low CSL, Kanchanawong P (2022). Self-supervised classification of subcellular morphometric phenotypes reveals extracellular matrix-specific morphological responses. *Sci Rep* 12, 1–14.
- Wu Z, Ramsundar B, Feinberg EN, Gomes J, Geniesse C, Pappu AS, Leswing K, Pande V (2018). MoleculeNet: A benchmark for molecular machine learning. *Chem Sci* 9, 513–530.

- Youden WJ (1950). Index for rating diagnostic tests. *Cancer* 3, 32–35.
- Yu FX, Guan KL (2013). The Hippo pathway: Regulators and regulations. *Genes Dev* 27, 355–371.
- Yue J, López JM (2020). Understanding MAPK signaling pathways in apoptosis. *Int J Mol Sci* 21, 2346.
- Zaghloul I, Ptachcinski RJ, Burckart GJ, Van Thiel D, Starzel Th. E, Venkataramanan R (1987). Blood protein binding of cyclosporine in transplant patients. *J Clin Pharmacol* 27, 240–242.
- Zhao H, Darzynkiewicz Z (2013). Biomarkers of cell senescence assessed by imaging cytometry. *Methods Mol Biol* 965, 83–92.
- Boruta · PyPI. Available at: <https://pypi.org/project/Boruta/> (accessed 2 January 2023).
- cytomining/DeepProfiler: Morphological profiling using deep learning. Available at: <https://github.com/cytomining/DeepProfiler> (accessed 7 October 2023).
- Exploring ToxCast Data | US EPA. Available at: www.epa.gov/chemical-research/exploring-toxcast-data (accessed 8 July 2023).
- Help! How does the Robust Background method work? | Carpenter-Singh Lab. Available at: <https://carpenter-singh-lab.broadinstitute.org/blog/help-how-does-robust-background-method-work> (accessed 7 July 2023).
- RDKit. Available at: www.rdkit.org/ (accessed 17 April 2023).
- scikit-learn: machine learning in Python — scikit-learn 1.2.0 documentation. Available at: <https://scikit-learn.org/stable/index.html> (accessed 2 January 2023).