AutoHCS: Automated AI-based scoring and phenotype clustering in high-content screens

Ilya Goldberg[1], *Teresa Findley*[1], *John Delaney*[1] *& Rupert Dodkins*[1]

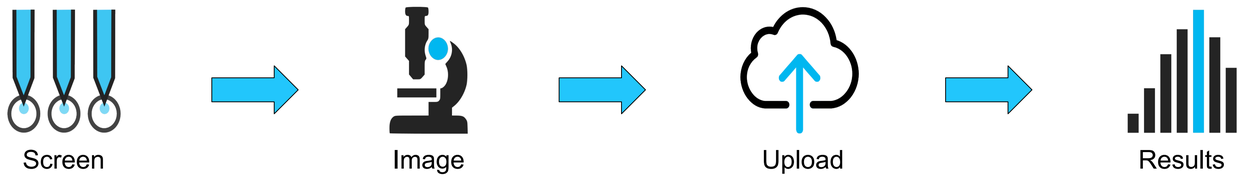
[1]ViQi Inc.

ABSTRACT

Modern drug development increasingly depends on high-content compound screens where automation is the key to rapid discoveries. ViQi, Inc.'s AutoHCS™ is an AI-based system that automatically detects and scores dose-dependent phenotypic responses to drugs in high-content screens. This platform has extensive applications within drug discovery including toxicity screens, therapeutics, and antiviral screens. With only images acquired on an automated plate imager and a plate map specifying compound concentrations, experimental replicates, and controls, AutoHCS scores phenotypic responses to each compound concentration within hours. Because the system does not depend on segmentation, it works non-parametrically with multichannel fluorescence, a combination of fluorescence and brightfield, or brightfield alone. Any cell line that can be experimentally set up to display any target phenotype can be used in a screen for compounds, or other manipulations that mimic a desired phenotypic response. For example, AutoHCS can be used on a screen for compounds that change the localization of proteins tracked with fluorescent tags, cells infected with virus in a screen for antiviral compounds, screens for chemoprotectants from environmentally or genetically induced negative phenotypes, toxicity screens for specific cell lines, phenotypic responses to gene inactivation, etc. Cells are plated on 96 or 384 well plates treated with different concentrations of compounds, as well as controls that mimic the target phenotypes, and untreated cells. For individual compounds, the system scores the phenotypic response for each concentration relative to each of the positive controls. In addition, phenotypes at each concentration are compared to all of the controls at once using a dendrogram. AutoHCS also scores phenotypic change across compound concentrations independently of the controls, permitting the discovery of novel phenotypes. This has many possible applications including the discovery of unwanted or unexpected cellular responses. Finally, the report uses a dendrogram to compare phenotype similarities between all of compounds in the screen. Similar compound-induced phenotypes are grouped in this dendrogram, which can indicate which compounds induce similar effects such as functional analogs, or compounds acting on different parts of the same pathway. AutoHCS relies on using the experimental controls alone rather than relying on user inputs to guide the analysis. This eliminates subjective parameter selection that may bias phenotype scoring. This also makes it extremely easy to use, as the analysis depends only on images and a description of the experimental plate layout. The analysis platform is cloud-based, so there is no software or specialized computing hardware to install locally. Image management and storage are also provided automatically. AutoHCS harnesses the pattern recognition abilities of modern AIs to precisely score high-content screens in an automated and objective manner.

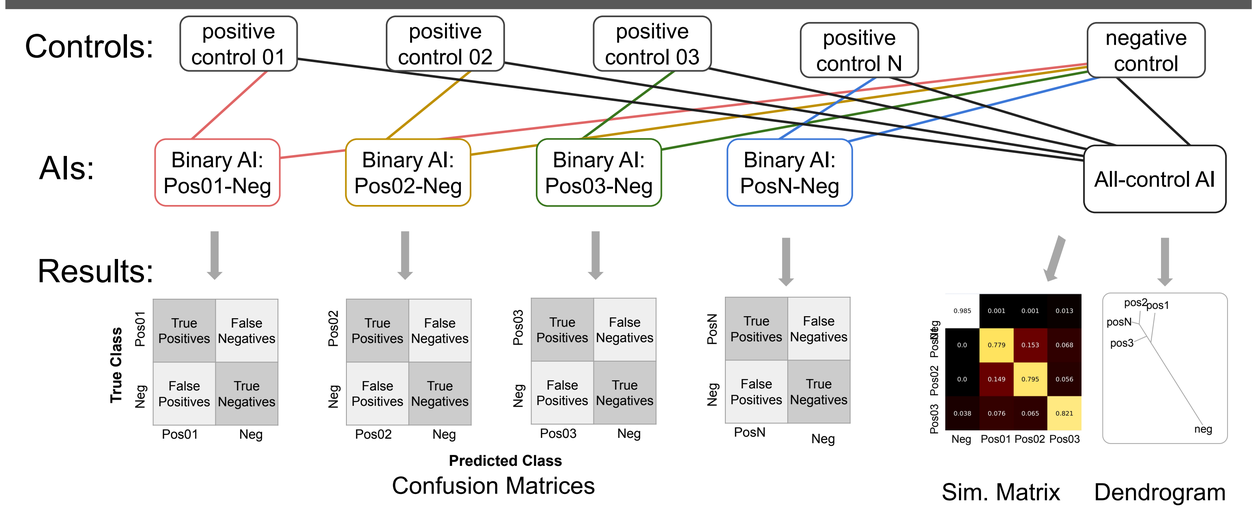
WORKFLOW & EXPERIMENT SET UP

AutoHCS workflow



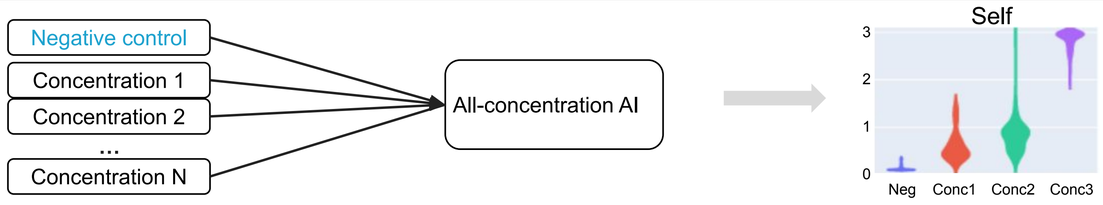
Using autoHCS is simple and user-friendly. 1) Plate cells configured to display a target phenotype for the screen, compounds at varying dilutions, and positive and negative controls. 2) Capture images using one of the many high-throughput microplate imaging devices that can automatically image plates at high resolution. 3) Upload the images to ViQi Inc. servers along with a plate map describing the location of drugs, concentrations, and positive and negative controls. 4) Receive a complete analysis report and quantitative assay readout for each well.

TRAINING A BINARY AI FOR EACH CONTROL



For each positive control, we train a binary AI to distinguish that control from the negative control resulting in a simple confusion matrix of true or false positives and negatives. We also train an AI on all controls resulting in a similarity matrix and dendrogram depicting similarities between positive controls.

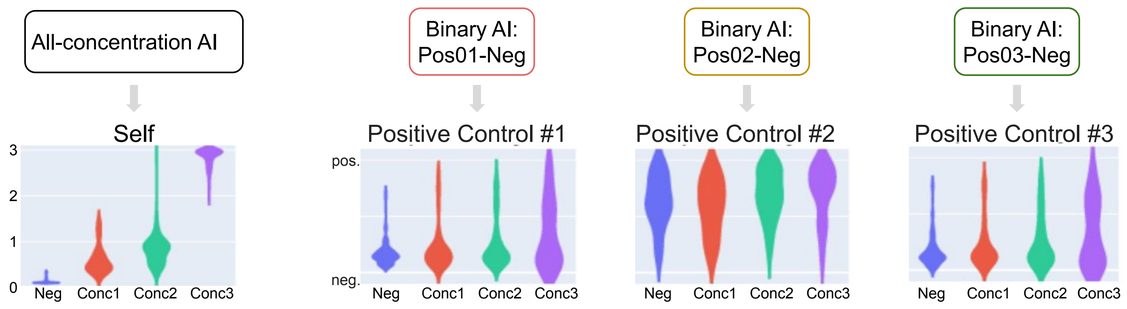
TRAINING A DOSE-RESPONSE AI FOR EACH COMPOUND



For each compound, an AI is trained on all compound concentrations independently of positive controls. Scoring concentration-dependent effects independently of positive controls allows for the discovery of new phenotypes.

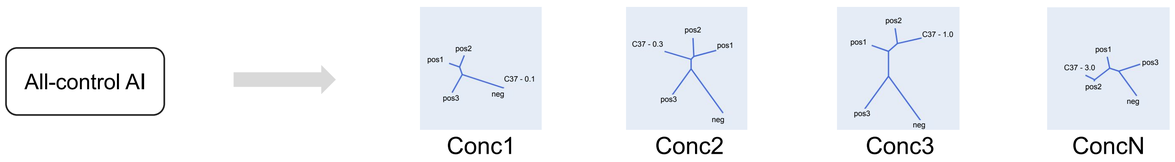
AUTOMATED COMPOUND SCORING

Scoring each compound using the dose-response AIs



Each compound is scored using the AI trained on its compound concentrations (first left) and using each binary control AI (right). By presenting scores using violin plots, users can tell at a glance which control (positive or negative) the phenotypic responses to compounds are more similar to.

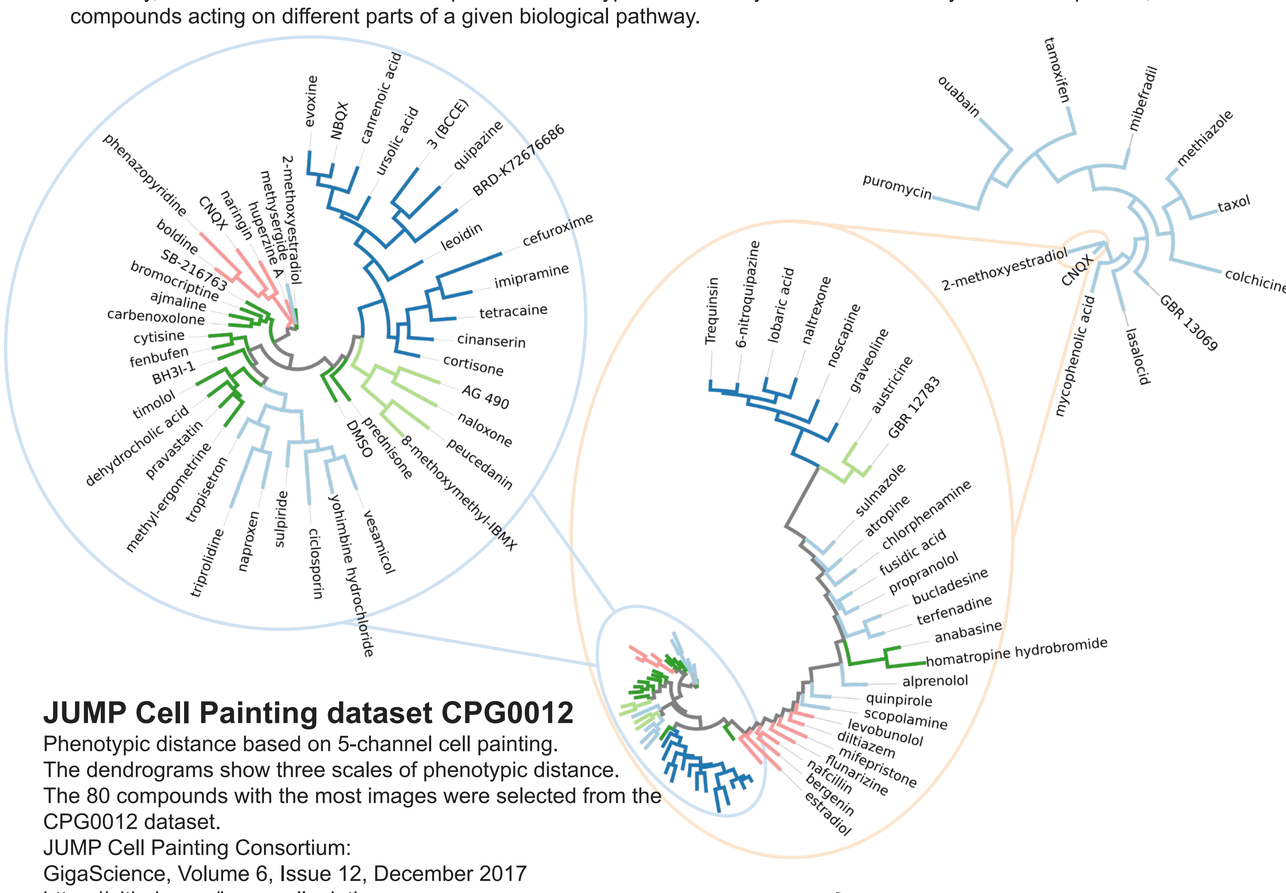
Scoring each concentration using the all-control AI



Each concentration is scored using the all-control Control AI and compound concentrations are then compared with controls for phenotypic similarity.

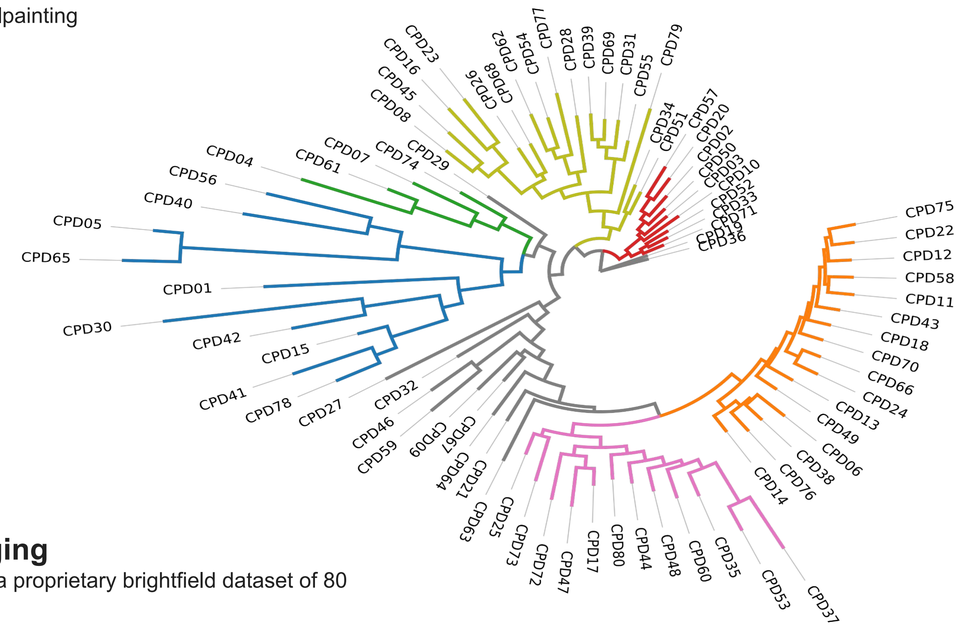
CLUSTERING COMPOUNDS BY PHENOTYPE SIMILARITY

Images from the highest concentrations for each compound are used to train an AI to produce estimates of phenotypic similarity, which are used to cluster compounds. Phenotype clusters may indicate functionally similar compounds, or compounds acting on different parts of a given biological pathway.



JUMP Cell Painting dataset CPG0012

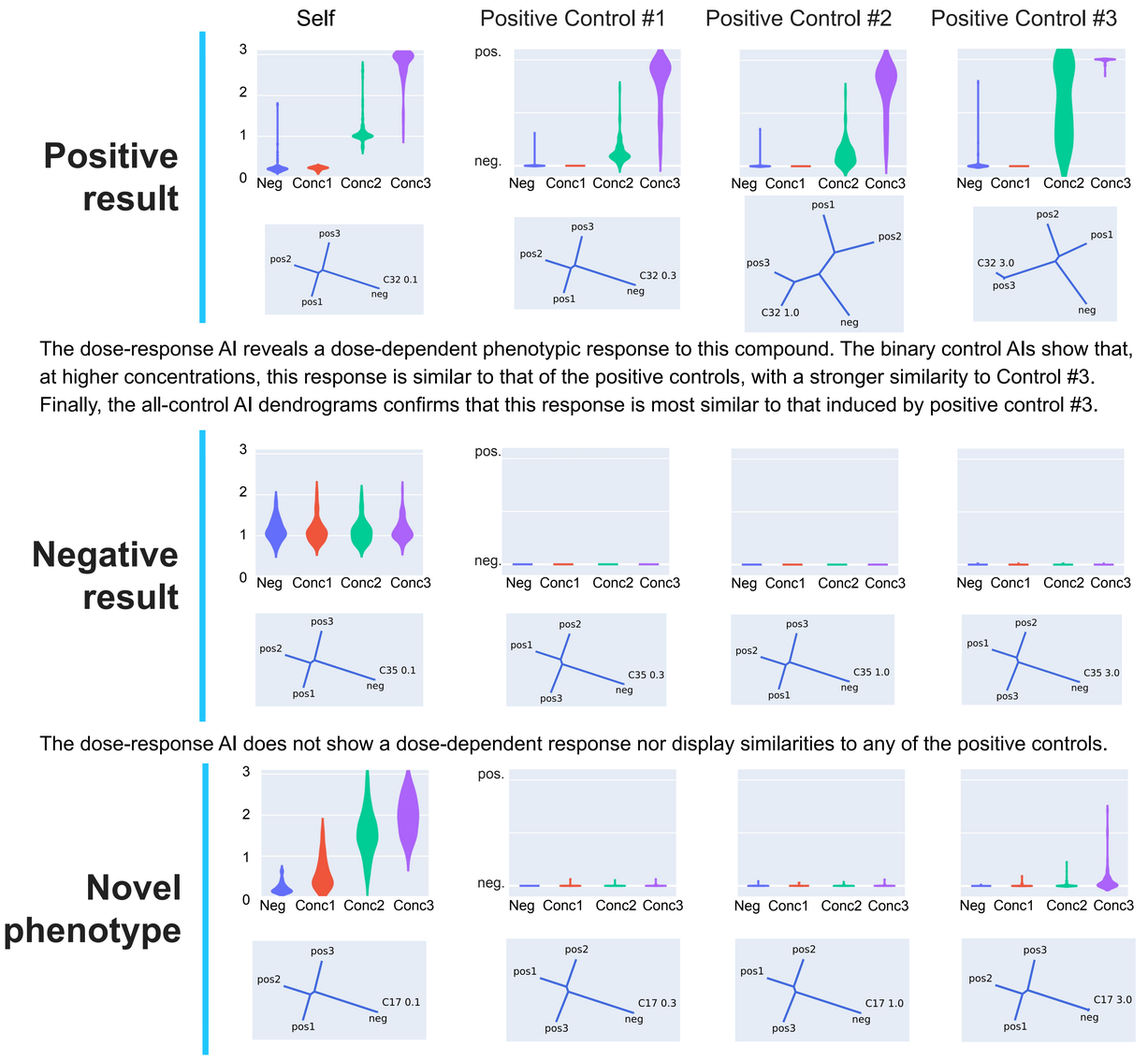
Phenotypic distance based on 5-channel cell painting.
The dendrograms show three scales of phenotypic distance.
The 80 compounds with the most images were selected from the CPG0012 dataset.
JUMP Cell Painting Consortium:
GigaScience, Volume 6, Issue 12, December 2017
https://github.com/jump-cellpainting



Brightfield imaging

Phenotypic distance in a proprietary brightfield dataset of 80 compounds.

AUTO-HCS IDENTIFIES MULTIPLE COMPOUND TYPES



The dose-response AI reveals a dose-dependent phenotypic response to this compound. The binary control AIs show that, at higher concentrations, this response is similar to that of the positive controls, with a stronger similarity to Control #3. Finally, the all-control AI dendrograms confirms that this response is most similar to that induced by positive control #3.

The dose-response AI does not show a dose-dependent response nor display similarities to any of the positive controls.

While the binary AIs show no similarities in phenotype with the positive controls, the dose-response AI for this compound which is trained only across concentrations (and not controls) shows a dose-dependent response indicating that an unanticipated phenotype arises in response to this compound.

CONCLUSION & NEXT STEPS…

- AutoHCS can rapidly score compounds in an automated and objective manner for any target phenotype in any cell line.
- We are looking for collaborators to test this approach in different types of screens!

Learn More at www.viqiai.com/autohcs