Exploring relationships between gene functions using JUMP Cell Painting Consortium data

Alán F. Muñoz and Anne Carpenter, in behalf of the JUMP Cell-Painting Consortium

Broad Institute of Harvard and MIT

Abstract

With the Cell Painting assay we quantify cell morphology using six dyes to stain eight cellular components: Nucleus, mitochondria, endoplasmic reticulum, nucleoli, cytoplasmic RNA, actin, golgi aparatus, and plasma membrane. After high-throughput fluorescence microscopy, image analysis algorithms then extract thousands of morphological features from each single cell's image. By comparing of these "profiles" we can can uncover new relationships among genetic and chemical perturbations.

This year, the JUMP-CP Consortium (Joint Undertaking for Morphological Profiling-Cell Painting) released the first public high-throughput dataset with over 140,000 genetic and chemical perturbations (Chandrasekaran et al. 2023).

Here, we describe how this data can now be used to answer many biological questions. Researchers can pick any gene of interest and find what morphological phenotypes are induced when it is knocked-out or over-expressed and what genes produce a similar morphological profile when altered, uncovering functional relationships. Novel software tools developed for this dataset empower biologists to make discoveries of their own, and we show that mining this dataset can yield novel insights into current and relevant biological questions.

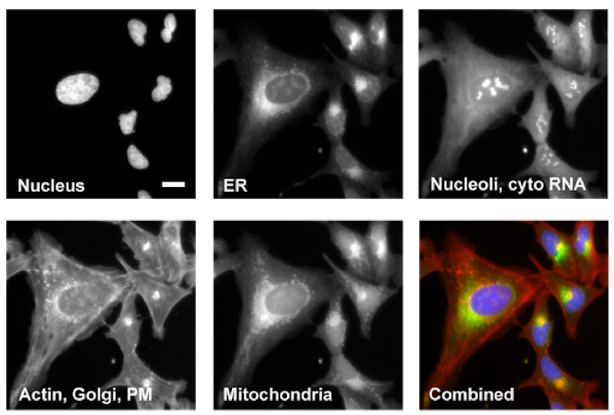
Goals

Device methods to interpret profile-based datasets to yield useful biological insight.

Develop a tool/workflow for biologists to discover genes that result in phenotypes similar to theirs.

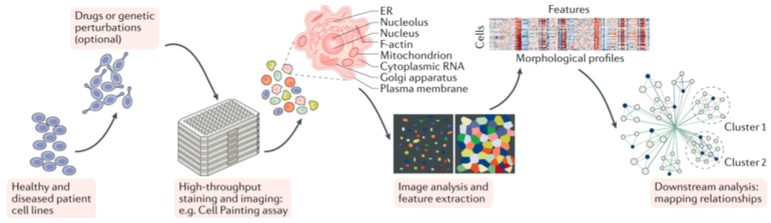
Build a stepping stone for a universal and accessible framework against which biologists can validate cell phenotypes.

We use data from the Cell Painting assay, in which eight cellular components are stained using six dyes and imaged in five channels



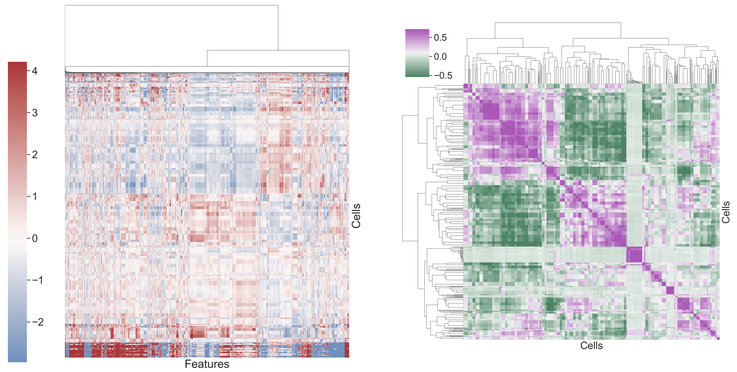
Morphological profiles were generated at a high-throughput scale

We generated and preprocessed a database composed of thousands of cell painting experiments.



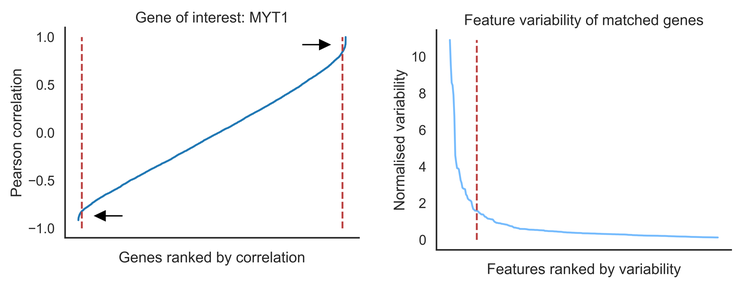
We generated a reference dataset for cells and features that indicates clustered groups of genes

It becomes possible to query individual genes and find similar profiles.



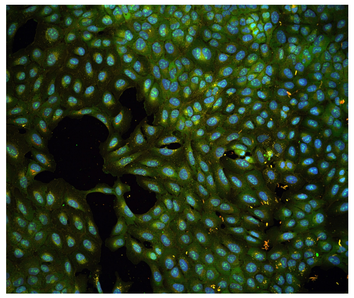
We then proceed to find morphological features that link genes closely

1. Find the most correlated and anticorrelated genes.
2. Find the features that show highest variance between these correlated/anticorrelated candidates.
3. Use these feature to guide comparisons between perturbed cells and negative controls.

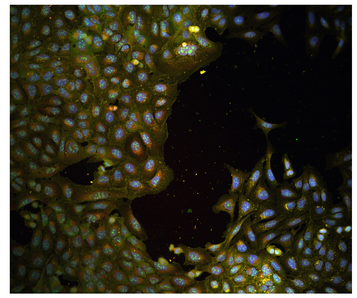


Using these morphological features eases discovering novel insights

We compare images using tools that decompose the channels to focus on the most important features obtained from data mining



MYT1 deletion



Control

Conclusions

The JUMP Cell Painting can serve as a resource to obtain candidate genes to find further insight on genes or proteins of interest.

A combination of computational and biological expertise can accelerate drug/mechanism discovery.

Even high throughput analyses require biological expertise to provide novel insights.

Next Steps

- Perform new experiments to confirm found candidates
- Incorporate compound data for bidirectional queries
- Explore further promising candidates, based on collaborators' work:

Gene/compound	Phenotype/disease
MYT1/RNF41	Neuronal fate
MUC1	Cancer
PDE Inhibitors	Cancer treatment
CTDNEP1	Nuclear structure
MMP9	Alzheimer

- Release a public tool for scientists to query their own profiles and access existing ones to design their experiments.

More info and updates at `broad.io/jump-cellpainting`

References

Chandrasekaran, Srinivas Niranj, Jeanelle Ackerman, Eric Alix, D. Michael Ando, John Arevalo, Melissa Bennion, Nicolas Boisseau, 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`.