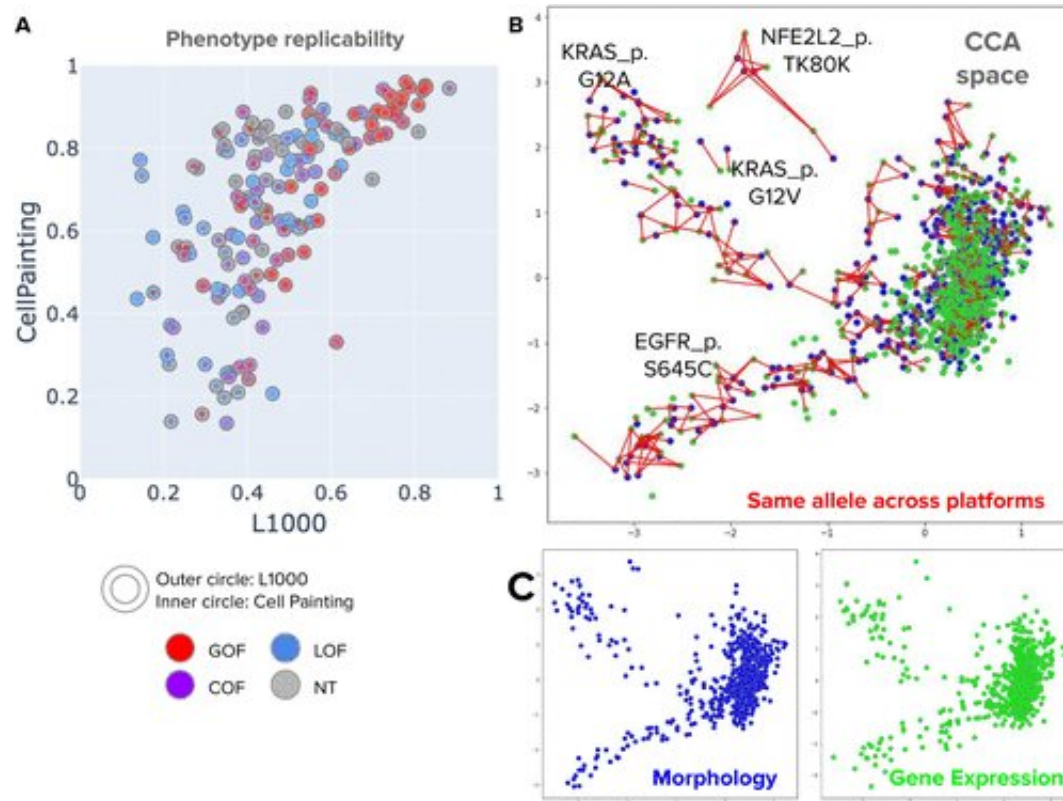


Scaling up cell imaging

August 3 2022, by Alex Viveros



Correlation between Cell Painting profiles and L1000 profiles for a common subset of 160 alleles. (A) Signal replicability, defined as the median pairwise correlation between replicates of the same allele, was calculated for each variant in the common subset in both profiling platforms. The x-axis corresponds to the signal strength in L1000 and the y-axis represents the signal strength in Cell Painting. The Spearman correlation coefficient is 0.69. (B) Canonical correlation analysis (CCA) in the multidimensional feature space for both profiling platforms at the perturbation level. CCA obtains a common latent space by finding the directions of maximal correlation between two multivariate datasets, allowing us to project data points from Cell Painting and L1000 in the same subspace. The axes in this plot are the first and second CCA directions. Points in

blue are morphology profiles and points in green are gene expression profiles. The red lines connect two points of different modalities that represent the same allele. (C) Same representation of Cell Painting profiles (morphology) and L1000 profiles (gene expression) in CCA space as in B, but using an independent plot for each platform. Credit: *Molecular Biology of the Cell* (2022). DOI: 10.1091/mbc.E21-11-0538

Scientists have learned a lot about human biology by looking at cells under a microscope, but they might not notice tiny differences between cells or even know what they're looking for. Researchers at the Broad Institute of MIT and Harvard, in the laboratories of Anne Carpenter and Stuart Schreiber, first started developing cell painting 13 years ago to take cell imaging to the next level. The method, further advanced by Carpenter, now senior director of the Broad's Imaging Platform and senior group leader Shantanu Singh, and colleagues, uses six colored dyes to stain eight different cell organelles. Machine learning models recognize subtle differences in the images—changes in cell morphology that might indicate disease or a drug or genetic perturbation—which allows researchers to predict the effects of a drug or mutation.

The Broad team has recently made strides in scaling up the method. They have spent the last several years building a consortium of drugmakers and academic institutions to create the world's largest public cell [painting](#) database, which drug developers hope will help accelerate their search for promising drug candidates.

Carpenter, Singh, and colleagues are also demonstrating more applications of cell painting beyond [drug discovery](#), such as using it to assess the impact of gene variants in [lung cancer](#). They are working to integrate cell painting with gene-expression profiling methods to show how it can complement and enhance gene-based methods, and finding

ways to extract even more information from cell painting images.

We spoke with Carpenter and Singh about their efforts to scale up cell painting and to learn how it has already made an impact.

What kinds of discoveries is cell painting now enabling?

Anne Carpenter: Cell painting is allowing us to accelerate multiple steps in the drug discovery pathway. For example, we can use it to identify what has gone wrong in cell structures for a certain disorder.

Once we figure out the disease's impact on cell structures, it helps us understand the mechanism of disease, and it also immediately provides a way to screen drugs. We simply try to find drugs that have the opposite morphological impact compared to the disease. The company Recursion (I serve on the company's Scientific Advisory Board) has used this strategy to put four potential medicines towards clinical trials so far. There are so many things you can do, and we're just scratching the surface of all the different kinds of applications that can hopefully improve human health. For example, the U.S. Environmental Protection Agency is beginning to use cell painting to identify toxic impacts of chemical compounds.

What challenges has your lab faced in working on cell painting?

Carpenter: Our biggest challenge has been the shortage of available data to make discoveries. The past three years, we've been leading the JUMP-Cell Painting Consortium, which is a collection of mostly [pharmaceutical companies](#) and a couple of nonprofits that are dedicated to making the world's largest public cell painting database of chemical

and genetic perturbations. We have in this past year just finished creating that large data set, and now it's available for us to do all kinds of experiments with. Like many labs at the Broad across different fields, our lab has really gained a reputation for making large datasets!

Shantanu Singh: I think one of the challenges has been scaling this up. I'm amazed how smoothly it has gone, despite the scale that we're trying to achieve so rapidly with our small but mighty team. You know, scientists need to be a little bit delusional, or they will never take on anything.

How are you using machine learning to mine these datasets?

Singh: The JUMP-Cell Painting Consortium's database, being released in November 2022, is huge, but I think of it as a seed to then build upon a much larger data set that others can contribute to. Once we have a very large lookup table of cell painting data, [machine learning](#) can help us find connections.

For example, we can find the function of genes that are currently unannotated, because we might find they group together with genes of known function. We might be able to predict whether a mutation that has been found to be associated with disease is in fact impactful or not.

Eventually, machine learning combined with these very large data sets will allow us to convert a lot of questions in biology into essentially linear algebra questions. That's the digital biology future I'm so excited about.

Carpenter: We're starting to use deep learning to extract more kinds of information and more subtle information than previous algorithms

allowed. The nice thing about machine learning is that it is not constrained by what the human brain or biologists think might be important to measure; it just aims to measure everything—all the variations in images that we might not notice or know to look for.

Can you talk about a recent project where machine learning helped you gain new insight?

Carpenter: One project uses deep learning networks to try to predict which drugs might be useful to treat particular disorders. Historically, if you want to do that kind of experiment, you'd design some kind of assay and then test thousands to millions of compounds to figure out which ones have the behavior you're looking for. This process is expensive but is the mainstay of most drugmakers.

By contrast, we took existing images of [cells](#) treated with different compounds and wondered if there was enough information in those images to predict whether the compounds might be useful for other diseases. We took a [deep learning](#) network and asked the computer to predict the outcome of a disease-focused assay, by asking the computer to find compounds whose cell images "match" cell images produced by perturbing disease-associated genes.

It doesn't work all the time, but in a small fraction of cases where it does, it eliminates the need to do a very large-scale chemical screen. It's a really inexpensive way to eliminate some of the lab work in trying to identify new drugs.

How have you used cell painting to study gene function?

Carpenter: One of the problems with cancer is that there are a lot of

mutations involved, and you don't know which ones are really causing the cancer and which ones have just occurred along the way. Figuring out which mutations are responsible is important for deciding what anti-cancer drugs a person should get.

We took cells growing in a dish and introduced the mutated forms of a bunch of proteins that were found in lung cancer. We then looked at the morphology of the cells that were expressing those proteins. If they had the normal versions of the protein, the cells looked a certain way; if they had a mutated version, they might look different.

By using cell painting to compare the normal versus mutant version of the protein, we were able to figure out whether that mutation was actually impacting the function of the gene, which can help guide clinical decisions.

Singh: Importantly, our lung carcinoma paper showed that you can come up with a hypothesis of whether a mutation is impactful or not just by looking at microscopic pictures of cells. In spite of seeing this so many times over, I still find it very uncanny when it works!

What impact will cell painting have on drug discovery?

Singh: It will be an invaluable tool in the toolbelt of drug hunters as we move into this new era where [drug](#) discovery is going to be a combination of boutique experiments combined with large, data-driven exercises. cell painting is such an easy way of getting so much information about biology that I'm confident it's going to be a key component of a lot of efforts out there.

More information: Juan C. Caicedo et al, Cell Painting predicts

impact of lung cancer variants, *Molecular Biology of the Cell* (2022).
[DOI: 10.1091/mbc.E21-11-0538](https://doi.org/10.1091/mbc.E21-11-0538)

Provided by Broad Institute of MIT and Harvard

Citation: Scaling up cell imaging (2022, August 3) retrieved 27 April 2024 from
<https://phys.org/news/2022-08-scaling-cell-imaging.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.