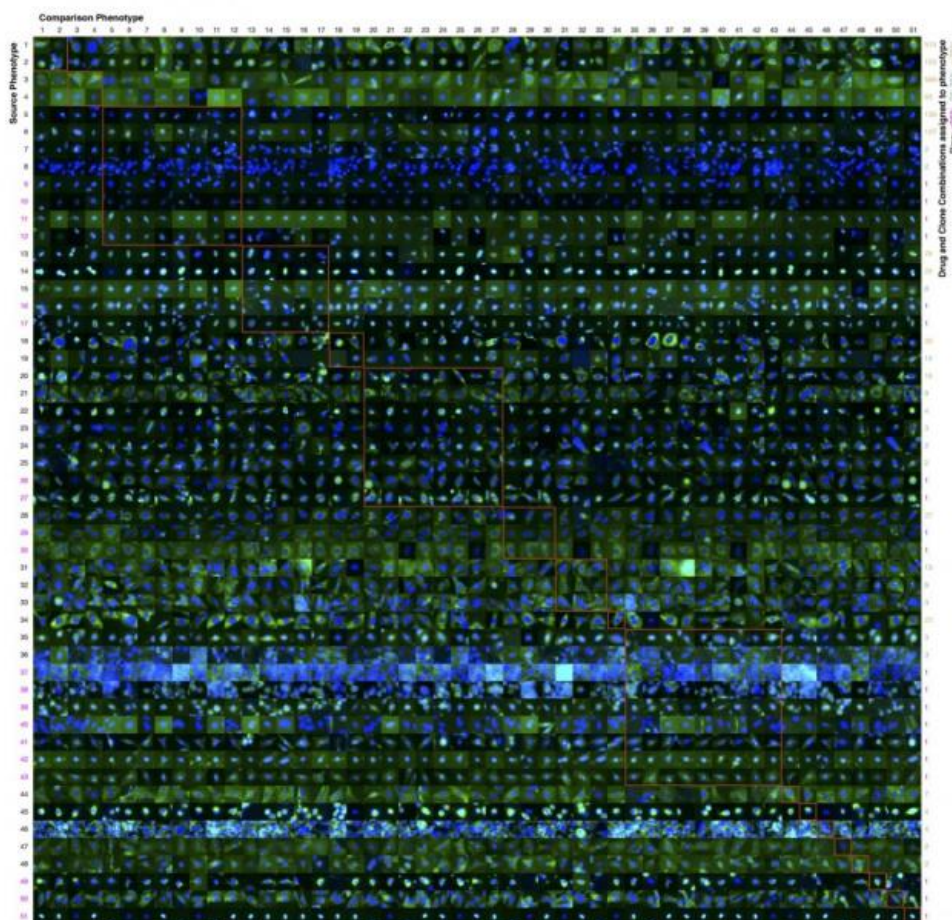


Robotically driven system could reduce cost of discovering drug and target interactions

February 9 2016



A comparison of 51 diverse phenotypes that an active learner identified by the final round of experiments it performed. The phenotypes have been roughly organized in the display by similarity; the learner in several cases identified phenotypes that were created by different drugs that visually were subtly different. Credit: Armaghan Naik and Robert F. Murphy

Researchers from Carnegie Mellon University (CMU) have created the first robotically driven experimentation system to determine the effects of a large number of drugs on many proteins, reducing the number of necessary experiments by 70%.

The model, presented in the journal *eLife*, uses an approach that could lead to accurate predictions of the interactions between novel drugs and their targets, helping reduce the cost of [drug](#) discovery.

"Biomedical scientists have invested a lot of effort in making it easier to perform numerous experiments quickly and cheaply," says lead author Armaghan Naik, a Lane Fellow in CMU's Computational Biology Department.

"However, we simply cannot perform an experiment for every possible combination of biological conditions, such as genetic mutation and cell type. Researchers have therefore had to choose a few conditions or targets to test exhaustively, or pick experiments themselves. The question is which experiments do you pick?"

Naik says that careful balance between performing experiments that can be predicted confidently and those that cannot is a challenge for humans, as it requires reasoning about an enormous amount of hypothetical outcomes at the same time.

To address this problem, the research team has previously described the application of a machine learning approach called "active learning". This involves a computer repeatedly choosing which experiments to do, in order to learn efficiently from the patterns it observes in the data. The team is led by senior author Robert F. Murphy, Professor at the Ray and Stephanie Lane Center for Computational Biology, and Head of CMU's

Computational Biology Department.

While their approach had only been tested using synthetic or previously acquired data, the team's current model builds on this by letting the computer choose which experiments to do. The experiments were then carried out using liquid-handling robots and an automated microscope.

The learner studied the possible interactions between 96 drugs and 96 cultured mammalian cell clones with different, fluorescently tagged proteins. A total of 9,216 experiments were possible, each consisting of acquiring images for a given cell clone in the presence of a given drug. The challenge for the algorithm was to learn how proteins were affected in each of these experiments, without performing all of them.

The first round of experiments began by collecting images of each clone for one of the drugs, totaling 96 experiments. Images were represented by numerical features that captured the [protein](#)'s location in the cell.

At the end of each round, all experiments that passed quality control were used to identify phenotypes (patterns in the location of a protein) that may or may not have related to a previously characterized drug effect.

A novelty of this work was for the learner to identify potentially new phenotypes on its own as part of the learning process. To do this, it clustered the images to form phenotypes. The phenotypes were then used to form a predictive model, so the learner could guess the outcomes of unmeasured experiments. The basis of the model was to identify sets of proteins that responded similarly to sets of drugs, so that it could predict the same prevailing trend in the unmeasured experiments.

The learner repeated the process for a total of 30 rounds, completing 2,697 out of the 9,216 possible experiments. As it progressively

performed the experiments, it identified more phenotypes and more patterns in how sets of proteins were affected by sets of drugs.

Using a variety of calculations, the team determined that the algorithm was able to learn a 92% accurate model for how the 96 drugs affected the 96 proteins, from only 29% of the experiments conducted.

"Our work has shown that doing a series of experiments under the control of a machine learner is feasible even when the set of outcomes is unknown. We also demonstrated the possibility of active learning when the robot is unable to follow a decision tree," explains Murphy.

"The immediate challenge will be to use these methods to reduce the cost of achieving the goals of major, multi-site projects, such as The Cancer Genome Atlas, which aims to accelerate understanding of the molecular basis of cancer with genome analysis technologies."

More information: Armaghan W Naik et al. Active machine learning-driven experimentation to determine compound effects on protein patterns, *eLife* (2016). [DOI: 10.7554/eLife.10047](https://doi.org/10.7554/eLife.10047)

Provided by eLife

Citation: Robotically driven system could reduce cost of discovering drug and target interactions (2016, February 9) retrieved 5 May 2024 from <https://phys.org/news/2016-02-robotically-driven-drug-interactions.html>

<p>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.</p>
--