

Big-data analysis points toward new drug discovery method

July 12 2017

A research team led by scientists at UC San Francisco has developed a computational method to systematically probe massive amounts of open-access data to discover new ways to use drugs, including some that have already been approved for other uses.

The method enables scientists to bypass the usual experiments in biological specimens and to instead do computational analyses, using open-access data to match FDA-approved drugs and other existing compounds to the molecular fingerprints of diseases like cancer. The specificity of the links between these drugs and the diseases they are predicted to be able to treat holds the potential to target drugs in ways that minimize side effects, overcome resistance and reveal more clearly how both the drugs and the diseases are working.

"This points toward a day when doctors may treat their patients with drugs that have been individually tailored to the idiosyncracies of their own disease," said first author Bin Chen, assistant professor with the Institute for Computational Health Sciences (ICHS) and the Department of Pediatrics at UCSF.

In a paper published online on July 12, 2017, in *Nature Communications*, the UCSF team used the method to identify four drugs with cancer-fighting potential, demonstrating that one of them—an FDA-approved [drug](#) called pyrvinium pamoate, which is used to treat pinworms—could shrink hepatocellular carcinoma, a type of [liver cancer](#), in mice. This cancer, which is associated with underlying liver disease and cirrhosis, is

the second-largest cause of cancer deaths around the world—with a very high incidence in China—yet it has no effective treatment.

The researchers first looked in The Cancer Genome Atlas (TCGA), a comprehensive map of genomic changes in nearly three dozen types of cancer that contains more than two petabytes of data, and compared the [gene expression](#) signatures in 14 different cancers to the gene expression signatures for normal tissues that were adjacent to these tumors. This enabled them to see which genes were up- or down-regulated in the cancerous tissue, compared to the normal tissue.

Once they knew that, they were able to search in another open-access database, called the Library of Integrated Network-based Cellular Signatures (LINCS) L1000 dataset, to see how thousands of compounds and chemicals affected cancer cells. The researchers ranked 12,442 small molecules profiled in 71 cell lines based on their ability to reverse abnormal changes in gene expression that lead to the production of harmful proteins. These changes are common in cancers, although different tumors exhibit different patterns of abnormalities. Each of these profiles included measurements of gene expression from 978 "landmark genes" at different drug concentrations and different treatment durations.

The researchers used a third database, ChEMBL, for data on how well biologically active chemicals killed specific types of cancer cells in the lab—specifically for data on a drug efficacy measure known as the IC50. Finally, Chen used the Cancer Cell Line Encyclopedia to analyze and compare molecular profiles from more than 1,000 cancer cell lines.

Their analyses revealed that four drugs were likely to be effective, including pyrvinium pamoate, which they tested against liver cancer cells that had grown into tumors in laboratory mice.

"Since in many cancers, we already have lots of known drug efficacy data, we were able to perform large-scale analyses without running any biological experiments," Chen said.

He and colleagues developed a ranking system, which he calls the Reverse Gene Expression Score (RGES), a predictive measure of how a given drug would reverse the gene-expression profile in a particular disease—tamping down genes that are over-expressed, and ramping up those that are weakly expressed, thus restoring gene expression to levels that more closely match normal tissue.

After using open-access databases to determine that RGES was correlated with drug efficacy in liver cancer, breast cancer and colon cancer. Chen focused on liver cancer cell lines, but since they have not been investigated as much as breast and colon cancer cell lines, there was far less data available to study them. So, he used RGES scores for drugs and other biologically active molecules that had been tested on non-liver cancer cell types. The RGES scores were powerful enough that he could still predict which molecules might kill liver cancer [cells](#).

Chen's collaborators from the Asian Liver Center at Stanford University examined four candidate molecules with known mechanisms of drug action. They found that all four killed five distinct liver cancer cell lines grown in the lab. Pyrvinium pamoate was the most promising drug, shrinking liver tumors grown beneath the skin in mice.

Cancer researchers usually target individual genetic mutations, but Chen said drugs that are targeted in this way often are less effective than anticipated and generate drug resistance. He said a broader measure such as RGES might lead to better drugs and also help researchers identify new drug targets.

Because RGES is based on the molecular characteristics of real tumors,

Chen said it also may be a better predictor of a drug's true clinical promise than high-throughput screening of large panels of drugs and other small molecules, which are based on drug activity in lab-grown cell lines.

"As costs come down and the number of gene expression profiles in diseases continues to grow, I expect that we and others will be able to use RGES to screen for drug candidates very efficiently and cost-effectively," Chen said. "Our hope is that ultimately our computational approach can be broadly applied, not only to [cancer](#), but also to other diseases where molecular data exist, and that it will speed up drug discovery in diseases with high unmet needs. But I'm most excited about the possibilities for applying this approach to individual patients to prescribe the best drug for each."

More information: Bin Chen et al. Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets, *Nature Communications* (2017). DOI: 10.1038/ncomms16022 , www.nature.com/articles/ncomms16022

Provided by University of California, San Francisco

Citation: Big-data analysis points toward new drug discovery method (2017, July 12) retrieved 24 April 2024 from <https://phys.org/news/2017-07-big-data-analysis-drug-discovery-method.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.