

# Finding the repurpose

By Tracey Baas, Senior Editor

A **Stanford University** team has used a computational approach to analyze publicly available gene expression profiles and identified two drug repurposing opportunities.<sup>1,2</sup> The researchers found that the anticonvulsant topiramate could have use in inflammatory bowel disease and that the generic ulcer drug cimetidine could help treat lung adenocarcinoma.

The computational method is exclusively licensed to **NuMedii Inc.**, a bioinformatics spinout from Stanford that is using the technology to evaluate molecular data to repurpose drugs.

The Stanford group used previously published gene expression data from 100 diseases and 164 drugs to confirm already known therapeutic uses and to predict the potential of the drugs to treat new indications (see **Figure 1**, “Schematic workflow of drug–disease pairing using gene expression profile similarity”).

The computational technique showed that topiramate had a stronger therapeutic score for IBD than prednisolone and that cimetidine had a stronger therapeutic score for lung adenocarcinoma than **AstraZeneca plc**’s Iressa gefitinib.

Topiramate, a sulfamate-substituted monosaccharide that acts as a sodium channel blocker,  $\gamma$ -aminobutyric acid receptor (GABAR) agonist and AMPAR antagonist, is marketed to treat epilepsy, seizures and migraines. The drug is available as a generic and is marketed as Topamax by **Johnson & Johnson**.

Iressa, an epidermal growth factor receptor (EGFR) inhibitor, is marketed in the EU to treat non–small cell lung cancer (NSCLC).

In a second *in silico* study, the researchers specifically focused on IBD. Using only the gene expression data from IBD patient samples and those of the 164 drug compounds, they again confirmed that topiramate had a stronger IBD therapeutic score than prednisolone.

The IBD and lung cancer results also were borne out in placebo-controlled animal studies. In a rat model of IBD, topiramate lowered the incidence of diarrhea and IBD-related colitis compared with saline. In mice with human lung adenocarcinoma cells, cimetidine decreased tumor growth compared with saline control.

The corresponding author of the study, Atul Butte, is chief of the Division of Systems Medicine and an associate professor in medicine and pediatrics at the **Stanford University School of Medicine**. He is a cofounder and chairman of the scientific advisory board at NuMedii.

Data were published in *Science Translational Medicine*.

“Drug repurposing using gene expression information is not new,”

according to Guanghui Hu, associate director of informatics and analysis at **Merck & Co. Inc.** However, he added, “this work makes two major contributions: Butte and coworkers manually generated 100 high-quality disease signatures, many of which had been missed in previous studies that mostly relied on automatic data processing, and by far the most important contribution was to validate some of the findings *in vivo*.”

Oscar Puig, also associate director of informatics and analysis at Merck, added that “it is fairly quick to use these types of computational approaches to generate many potential new indications—the bottleneck is experimental validation, including both *in vitro* and *in vivo* testing. The lack of large-scale, cost-effective systems to speed up the validation process is a problem.”

Jeremy Jenkins, a senior investigator in chemical biology informatics at **Novartis AG**’s **Novartis Institutes for BioMedical Research**, added that “typically, transcriptional profiles have been used to correlate compounds to one another to predict mechanism. In Butte’s case, compound profiles are compared directly to disease signatures, bypassing the difficult requirement of compound–target and even compound–mechanism elucidation.”

## Beyond expression

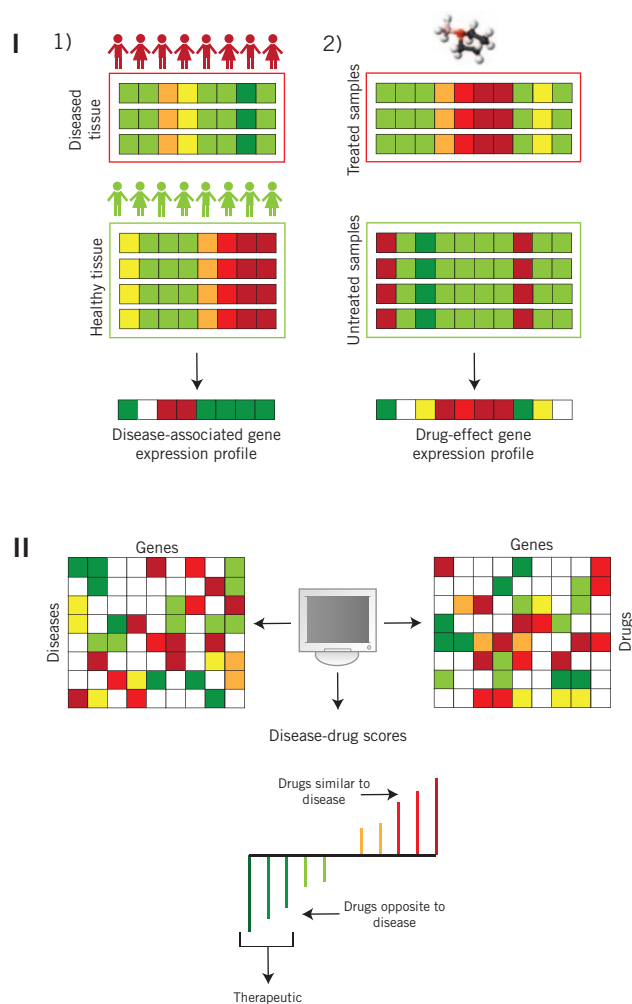
Butte thinks that with the wide range of publicly available data repositories, his team’s computational method can provide a systematic approach for repositioning established drugs to treat a wide range of diseases.

“Although we have shown proof of concept that our method works using microarray gene expression data available through public repositories such as the **National Center for Biotechnology Information** (NCBI)’s Gene Expression Omnibus and the **Broad Institute of MIT and Harvard**’s Connectivity Map, any type of experimentally obtained molecular signature data could be computationally analyzed and mined for possible drug repurposing,” he said.

“I’m tremendously excited by Butte’s effort, which represents leading-edge thinking in this important area,” said David Shaywitz. “Today, most medical research is fundamentally reductionist in nature, trying to cure disease by breaking it down into the component parts. There’s an urgent need—and opportunity—for complementary efforts that take a far more empirical approach and seek to capture and leverage the vast quantities of existing data and use this information to refine clinical treatment and improve patient health.”

Shaywitz is director of strategic and commercial planning at **Theravance Inc.** He also is cofounder of the Harvard PASTEUR program, a translational research initiative at **Harvard Medical School**, and a founding advisor of **Sage Bionetworks**, a not-for-profit medical research initiative focused on open innovation.

According to Butte, the computational analysis technique could be applied to datasets beyond gene expression information. “Multiple publically accessible data repositories exist and hold different forms of high throughput data based on systems chemical biology information,” he said. “One example of such a data repository is NCBI’s PubChem Bioassay, where hundreds of different bioactivity readouts are stored for millions



**Figure 1. Schematic workflow of drug-disease pairing using gene expression profile similarity.** Two types of gene expression collections are used to computationally evaluate a compound's therapeutic potential. **(I)** Disease-associated gene expression collections contain data that compare genes up- and downregulated in diseased tissue with genes in healthy tissue (**I.1**). Drug-effect gene expression collections contain data that compare genes up- and downregulated in drug-treated tissue with genes in untreated tissue (**I.2**).

**(II)** Using disease-associated and drug-effect gene expression collections, a modified Connectivity Map computational method<sup>4</sup> is used to query disease profiles against the drug-effect profiles and assign a disease-drug score to each comparison based on profile similarity. A drug that causes opposite gene expression changes from a specific disease (that is, the drug-effect and disease-associated profiles have a negative similarity score) is hypothesized to potentially have a therapeutic effect on that disease. (Figure based on Figure 1 in ref. 1.)

drug responses from whole-genome sequencing data in a family of four.<sup>3</sup>

The study, led by Euan Ashley, assistant professor in cardiovascular medicine at Stanford, and coauthored by Butte, estimated risk factors for multiple diseases but focused in particular on the family history of pulmonary embolism and bleeding risk when using therapeutic blood thinners.

Personalis was founded in August of this year by Ashley; Butte; Russ Altman, chairman of the bioengineering department at Stanford; Michael Snyder, chairman in the genetics department at Stanford; and John West, previously CEO of Solexa Ltd., which was acquired by **Illumina Inc.** Personalis plans to use genome sequencing technology and computational analysis to provide genome interpretation for use in academic studies and clinical medicine.

Stanford and NuMedii have filed for patents covering topiramate to treat IBD and cimetidine to treat lung cancer. The IP is available for licensing from NuMedii.

**Baas, T. *SciBX* 4(37); doi:10.1038/scibx.2011.1030**  
Published online Sept. 22, 2011

## REFERENCES

- Sirota, M. *et al. Sci. Transl. Med.*; published online Aug. 17, 2011; doi:10.1126/scitranslmed.3001318  
**Contact:** Atul J. Butte, Stanford University School of Medicine, Stanford, Calif. e-mail: [abutte@stanford.edu](mailto:abutte@stanford.edu)
- Dudley, J.T. *et al. Sci. Transl. Med.*; published online Aug. 17, 2011; doi:10.1126/scitranslmed.3002648  
**Contact:** Atul J. Butte, Stanford University School of Medicine, Stanford, Calif. e-mail: [abutte@stanford.edu](mailto:abutte@stanford.edu)
- Dewey, F.E. *et al. PLoS Genet.*; published online Sept. 15, 2011; doi:10.1371/journal.pgen.1002280  
**Contact:** Euan A. Ashley, Stanford University, Stanford, Calif. e-mail: [euana@stanford.edu](mailto:euana@stanford.edu)
- Lamb, J. *et al. Science* **313**, 1929–1935 (2006)

## COMPANIES AND INSTITUTIONS MENTIONED

**AstraZeneca plc** (LSE:AZN; NYSE:AZN), London, U.K.  
**Broad Institute of MIT and Harvard**, Cambridge, Mass.  
**Harvard Medical School**, Boston, Mass.  
**Illumina Inc.** (NASDAQ:ILMN), San Diego, Calif.  
**Johnson & Johnson** (NYSE:JNJ), New Brunswick, N.J.  
**Merck & Co. Inc.** (NYSE:MRK), Whitehouse Station, N.J.  
**National Center for Biotechnology Information**, Bethesda, Md.

of different chemical molecules. Just as we evaluated a compound's therapeutic potential by mining gene expression data, we can do a similar thing by mining bioassay readout data. And that is only one example."

Jenkins would like to see other data integrated into future computational analyses to further elucidate an emerging axis of drug targets presented in Butte's manuscript but not fully evaluated.

"Their drug-disease matrix contains an implicit third axis of drug targets," he said. "Effective mapping of compounds to targets through integrated bioassay data would allow a deeper understanding of how compound-target-disease transcriptional responses can be exploited."

Puig and Hu agreed. They noted that high throughput-based results such as *in vitro* screening, text mining or structural similarity data can be integrated for orthogonal profile analysis to further narrow down the best compounds for experimental validation.

Regardless of data type, Butte said NuMedii is interested in using computational analysis to find similarities between molecular signatures of orphan diseases and diseases with available therapeutics, as well as new first uses for compounds in development.

Work published last week in the *Public Library of Science Genetics* provided the basis for **Personalis Inc.**, another company cofounded by Butte. Using publically accessible databases, a group of Stanford University scientists published a computational analysis of genetic disease risk and

**Novartis AG** (NYSE:NVS; SIX:NOVN), Basel, Switzerland  
**Novartis Institutes for BioMedical Research**, Cambridge, Mass.  
**NuMedii Inc.**, Menlo Park, Calif.  
**Personalis Inc.**, Palo Alto, Calif.

**Sage Bionetworks**, Seattle, Wash.  
**Stanford University**, Stanford, Calif.  
**Stanford University School of Medicine**, Stanford, Calif.  
**Theravance Inc.** (NASDAQ:THRX), South San Francisco, Calif.