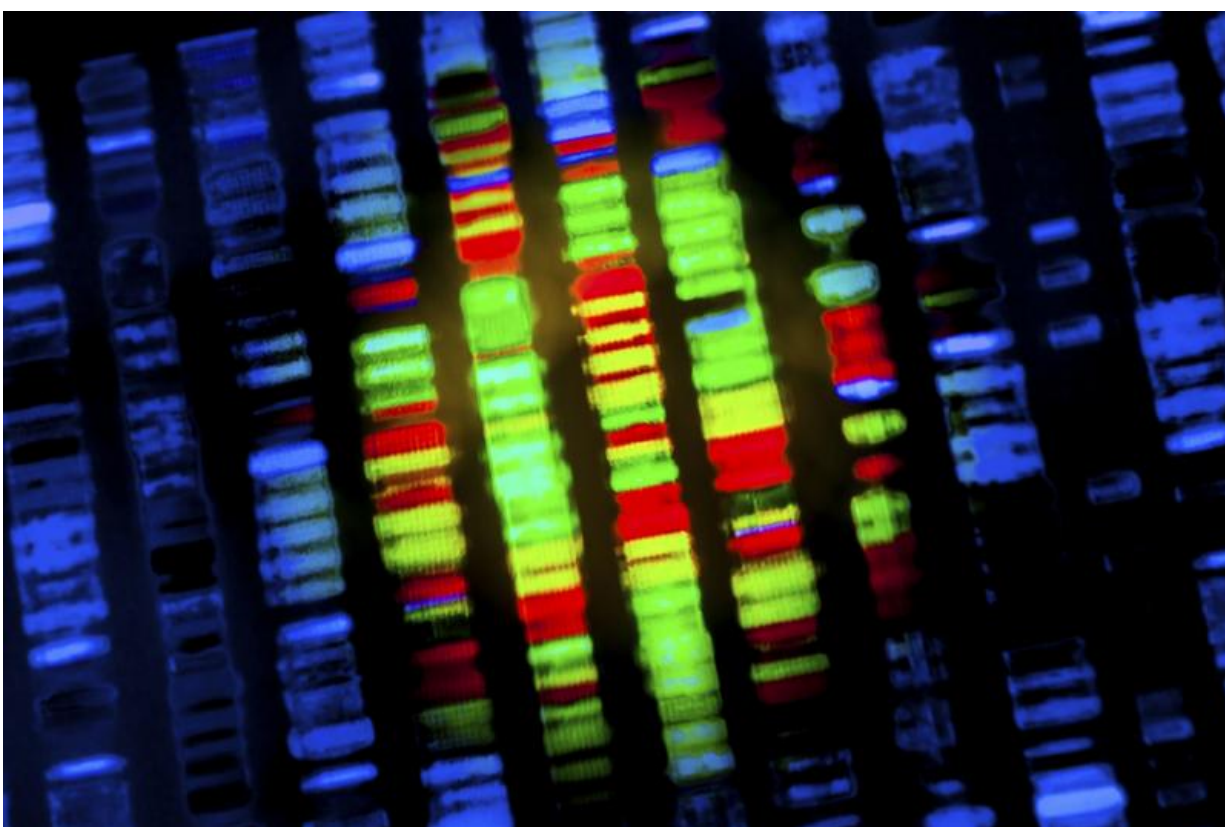# Faster drug discovery? Startup develops more cost-effective test for assessing how cells respond to chemicals

January 28 2016, by Rob Matheson

CEO Justin Lamb says Genometry aims to help pharmaceutical firms use gene-expression data (represented here as a heat map of gene expression values) to effectively digitize their compound collections.

For pharmaceutical firms, gene-expression profiling has become a valuable tool for drug discovery. This process involves measuring the activity of a cell's genes in response to drugs, to determine the compounds' effectiveness, toxicity, and other characteristics. Conventional profiling methods, however, are inefficient or expensive, sometimes costing millions of dollars.

Now Genometry has commercialized a high-throughput gene-expression assay developed at the Broad Institute of MIT and Harvard, which operates at a fraction of the cost of conventional methods. It does so by using measurements of 1,000 genes to accurately and quickly estimate the activity of all the 20,000 or so genes expressed in a cell.

The fast, low-cost assay allows for much larger experiments than previously possible, and for gene-expression profiling to be used much earlier in the drug-discovery process—which could speed things up, says Genometry co-founder, president, and CEO Justin Lamb, a former Broad Institute researcher. "Rather than profiling only a handful of compounds at the end of the search for a new drug to confirm that you got what you wanted, you can use profiling right at the start of the search, and hence do the search in much more efficient ways, because you have much more information," he says.

A dozen pharmaceutical firms and other companies are now using the assay. Last October, Genometry signed a multiyear contract with Janssen Pharmaceutica to generate gene-expression profiles for 250,000 compounds in the Belgian firm's small-molecule library. This represents the first time gene-expression profiling has been applied at such a large scale. The data will be used for drug screening and improving the selection of candidate drugs before clinical studies.

Lamb co-developed the assay, called L1000, with researchers in Broad Institute Chief Scientific Officer Todd Golub's group, including

Genometry co-founder Aravind Subramanian and David Peck, both currently researchers at the Broad.

Landmark genes

For gene-expression profiling, researchers have traditionally used microarrays, pieces of DNA arranged on silicon wafers, or polymerase chain reaction (PCR), which copies DNA fragments in test tubes. PCR is more accurate, but microarrays are faster: It takes the same amount of time to measure a few dozen genes using PCR as it does to measure an entire transcriptome—the full range of genes expressed in a cell —using microarrays.

But these experiments, Subramanian says, can cost up to $500 to generate the cellular response to a single drug or other "perturbagen." Matching signatures against libraries of a million or so drug candidates and different cell types, adds up.

L1000 takes advantage of the fact "that genes don't act independently of each other. They travel in clusters," according to Subramanian.

Instead of recording all the gene expression in a cell, L1000 measures the expression of approximately 1,000 so-called landmark genes that have been selected because they have special qualities, such as being minimally redundant across the genome or widely expressed in different cell types. According to Subramanian, because of these properties, the landmark genes together contain around 80 percent of the information in the entire transcriptome.

Genometry provides L1000 as a service. Clients send lysate—a fluid of broken-apart cells—in 384-well plates, usually dozens at a time, to Genometry's headquarters and lab in Kendall Square. After some initial processing, the samples are mixed with microscopic beads of different

colors coated with DNA from landmark genes, before being run through a version of a flow cytometer. Varying fluorescence intensity of each bead denotes varying degrees of each gene's expression. Based on those measurements, a computational inference model infers how all other genes will behave. With fewer measurements, throughput increases and the price drops to a few dollars per sample, compared with $500 by other methods.

Pharmaceutical firms will use L1000 data from potentially hundred of thousands of compounds as an initial screen, or to characterize their entire library of chemicals, as is the case with Janssen. Firms could also use the data for "hit selection" to narrow down the number of viable candidate drugs, Lamb says. A firm may have, say, 1,000 compounds to potentially treat a certain disease. Running those through L1000 will quickly reduce that number to perhaps 50 of most selective compounds and the ones with the fewest off-target effects. "It's a convenient way to prioritize compounds for drug development," Lamb says.

"Google" for drug discovery

In 2012, Lamb and Subramanian formed Genometry to commercialize L1000, only after it had proven its mettle at Broad Institute. "The purpose of [Genometry] was not so much to take a hand-wavy academic idea and make it into a product—it was to take a tried, tested, and proven technology and come up with a mechanism by which the outside world can get easy and effective access to that innovation," Lamb says.

L1000 was developed for—and has become an integral component of—an ongoing Broad Institute project called the Connectivity Map, which aims to become, essentially, "Google for drug discovery," according to the project's website.

The Connectivity Map, launched in 2005, is a collection of gene-

expression responses to perturbagens, including chemical and genetic manipulations. "If you describe the actions of drugs, the effects of genetic manipulation, and disease states in a common language, connecting a disease with a potential remedy, for example, is a relatively easy pattern-matching problem. Gene expression provides a particularly good vocabulary for this," Lamb explains.

Users of the Connectivity Map, for instance, can upload a list of genes whose expression pattern characterizes a disease and, with one click, receive a list of compounds ranked by their effect on those genes. It's similar to using keywords in Google to call up pages ranked by their relevance, Subramanian says.

At the Broad Institute, L1000 has now helped researchers archive profiles for around 5,000 drugs—about half the drugs ever tested in humans—and perturbation of 3,000 genes using CRISPR genome editing and other tools. The Broad Institute hopes to amass data on the connections among all drugs, genes, and diseases in the near future, Subramanian says.

Lamb says Genometry aims to help pharmaceutical firms use gene-expression data to effectively digitize their compound collections. In doing so, firms can use various tools to "search for compounds with desirable characteristics, determine if new chemical matter will be effective against various diseases, or if they are too similar to existing drugs to be commercially viable," he says.

"It's important to not just digitize libraries available to academics," Subramanian adds, "but to also make sure companies are achieving the impact they want."

  More information: genometry.com/

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology