

TECHONOMY

Date: 30 April 2013

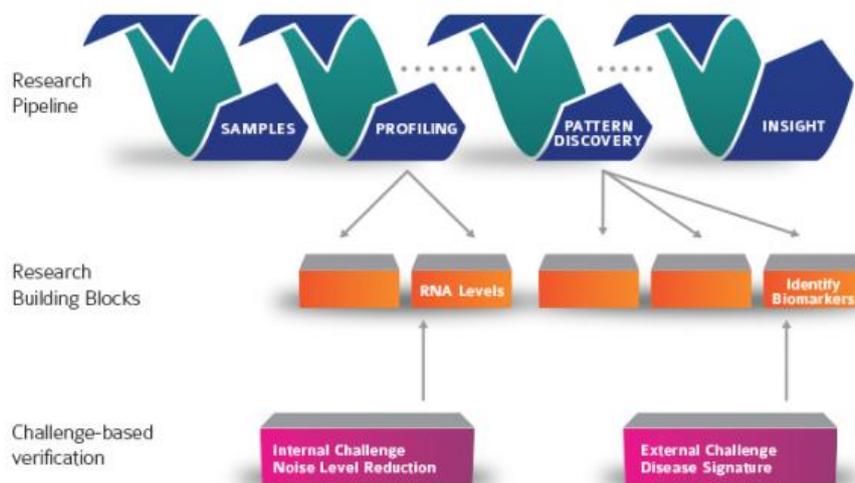
Headline: Do We Get Sick Like Rats? A New Philip Morris Prize Asks the Crowd

Byline: Adrienne Burke

It might be surprising to hear a tobacco giant described as a tech innovator. But Philip Morris researchers are pioneering new territory with a crowdsourced approach to checking the accuracy of life sciences data.

In partnership with computational biologists at IBM's Watson Research Center, Philip Morris's so-called sbv IMPROVER project creates open challenges to encourage scientists to augment traditional peer reviews of research data. On Monday, Philip Morris launched its Species Translation Challenge, which will award three \$20,000 prizes to teams whose results best define how well rodent tests can predict human outcomes.

Similar competitions have emerged in the academic world, but sbv IMPROVER (short for "systems biology verification of industrial methodology for process verification in research" in case you were wondering) is the first that taps the crowd to verify industrial research. An initial challenge last year awarded \$50,000 to two Wayne State University researchers who proved best at confirming genetic features that could be considered "diagnostic signatures" for particular diseases.



sbv IMPROVER at a Glance (via www.sbvimprover.com)

Why is a cigarette manufacturer sponsoring such competitions? "Our number one objective is to do something about our dangerous products," says Philip Morris scientific communications director,

Hugh Browne. (The company is known for its periodic candor about such matters, even as it continues to dominate the industry.) From heart disease to cancer to emphysema, the potential consequences of smoking are well known. But not every smoker suffers all or any of those health effects, suggesting that a combination of environmental and genetic factors lead to disease.

To understand precisely how smoking and chewing tobacco leads to complex interactions in a user's biological systems, "Philip Morris is increasing its investments into systems biology," Browne says. The company is looking at networks of genes, proteins, and biochemical reactions to identify the exact biological mechanisms perturbed by smoking.

But such biological data is notoriously complex to analyze. The profession as yet lacks any standard methodology for verifying results, and traditional peer-review methods have "struggled with the volume and complexity of the data," according to Philip Morris.

IMPROVER breaks research workflows into components and asks the crowd to apply its own computational methods to verify results. IBM computational biologist Gustavo Stolovitsky says the project "provides an excellent platform on which to test and develop some of the most cutting-edge approaches to the analysis of high-throughput biological data."

The 2012 IMPROVER challenge asked participants to identify signs in a patient's set of transcribed genetic material that could be relied on to diagnose any of four diseases associated with smoking: psoriasis, multiple sclerosis, chronic obstructive pulmonary disease, and lung cancer. Competitors looked at clinical data from patients—some of it licensed from third parties and provided by Philip Morris; some from the public domain.

More than 50 teams worldwide competed in the challenge that the Wayne State researchers won. Says Ajay Royyuru, director of IBM's Computational Biology Center in Yorktown Heights, NY: "There was a refreshing variety of competitors." The most successful applied fundamental understanding of biology "rather than brute force machine learning," or automated big-data analysis methods. "Some came at it from a mathematical modeling approach, others came from biology, and others combined those," he continues. (The Wayne State team comprised a bioinformaticist and a perinatal researcher.) Royyuru adds that the challenges can provide young scientists without scientific publications under their belt with a way to get recognition, and computational biology startup companies with a way to showcase what they can do.

A team of IBM computational biology experts scored entries, and a five-man outside panel reviewed the scores. While no single team identified the data perfectly, the leading methods, considered in the aggregate, performed exceptionally well, Royyuru says.

The new challenge launched this Monday seeks to determine if gene expression pathways identified in rodents will predict the same in humans. Scientists typically rely on them to study the impact of products on consumers, even though it remains unclear how well rodent results translate to humans.

Four sub-challenges ask participants to determine 1) if the way signaling pathways in one species react to a given stimulus really predicts similar response in another species, 2) which biological pathway functions and gene expression profiles are most parallel in rodents and humans, 3) how much that depends on the nature of the stimulus or data type collected, and 4) which computational methods are most effective for inferring responses between species.

Competitors will get access to about 5,000 human and rat samples Philip Morris generated for the challenge, and will look at 57 stressors to a single cell line exposed at different time points.

Browne suggests that the IMPROVER approaches for verifying results could be useful as well in the pharmaceutical, biotechnology, nutrition, and environmental safety industries. And Royyuru sees the project as a step toward creating “a verification methodology that will become routine industry practice.” Who knows how Philip Morris might utilize the outcomes? For better or worse, they may seek to create safer tobacco products.

-- ENDS --