New targets for the control of HIV predicted using a novel computational analysis

September 22 2011

A new computational approach has predicted numerous human proteins that the human immunodeficiency virus (HIV) requires to replicate itself. These discoveries "constitute a powerful resource for experimentalists who desire to discover new targets for human proteins that can control the spread of HIV," according to the authors of this study that appears in the Sept. 22, 2011 issue of *PLoS Computational Biology*, a journal published by the Public Library of Science.

The authors of the article are: T. M. Murali, a computer scientist, and Brett Tyler of the Virginia Bioinformatics Institute www.vbi.vt.edu/faculty/personal/Brett_Tyler, both located at Virginia Tech, and Michael G. Katze, a microbiologist and associate director of the Washington National Primate Research Center at the University of Washington.

David Badger of Blacksburg, Va., one of Murali's graduate students, and Matthew D. Dyer of Applied Biosystems of Foster City, Cal., also contributed to the study, which was funded by grants from the National Institutes of Health to Katze, the Virginia Bioinformatics Institute Fellows Program to Murali and Tyler, and the Virginia Tech's Support Program for Innovative Research Strategies to Murali.

When a person contracts HIV, it causes the progressive failure of the body's immune system, with the onset of life threatening infections and diseases such as cancer. Over 25 years of intensive research have failed to create a vaccine for preventing HIV. Moreover, drugs used to cure

HIV become rapidly ineffective because HIV is able to develop mutations against drugs, Murali said.

A recent line of research is examining whether human proteins can be targeted to cure HIV. Since viruses such as HIV have very small genomes, they must exploit the cellular machinery of the host to spread. Therefore, disrupting the activity of selected host proteins may impede viruses. Moreover, since human proteins evolve at a much slower rate than HIV proteins, human proteins that are targeted by drugs are very unlikely to develop mutations that render the drugs ineffective.

In fact, three studies published in 2008 systematically silenced virtually every human gene in order to discover HIV Dependency Factors (HDFs), i.e., those genes that are necessary for HIV to survive and replicate. Each of these three studies discovered hundreds of HDFs. However, a puzzling aspect was that only a handful of HDFs were common to two or more experiments.

"We set out to untangle this mystery," Murali said. "We hypothesized that many HDFs have not yet been discovered. Other papers had suggested that HDFs may themselves interact with each other. Inspired by these observations, we hypothesized that we could predict new HDFs by exploiting the proximity between HDFs within networks of interactions between human proteins."

To this end, they used an algorithm called SinkSource developed by Murali and Tyler. Tyler explained the algorithm using this analogy: "We treated the human protein network as if it were a system of tanks connected by pipes carrying water. This arrangement allowed us to study the flow of predictive information (water) from proteins we are certain about (full tanks) to those we are uncertain about (empty tanks). The further you get from the full tanks, the weaker the trickle, and the less water accumulates in the bottom of the tank. Mathematically you can

show that, over time, every empty tank accumulates some stable level of water. At the end of the analysis, tanks accumulating lots of water were judged to be good predictions."

"We found that SinkSource and one of its variants made predictions of very high quality," Murali added. "We evaluated predicted HDFS using a number of additional datasets that we did not use during the prediction step."

Their most exciting results used an analysis of HDF activities in two non-human primate species that respond differently to Simian Immunodeficiency Virus (SIV). One species, the African green monkey, does not develop disease when infected by SIV, in contrast to the other species, pig-tailed macaque. Using data already published by Katze, the authors showed that predicted HDFs had very different patterns of expression in the two species, especially in lymph nodes and within 10 days after infection with the virus. They also showed that predicted HDFs participated in human cellular processes that are known to be subverted by the virus, including gene transcription and translation, energy production, protein degradation, and transport across the nuclear membrane. Moreover, many predicted HDFs themselves directly interacted with proteins in HIV.

From these results, Murali, Tyler, and Katze concluded that existing genomic screens are "incomplete and many HDFs are yet to be discovered experimentally. Our results suggest that many HDFs are yet to be discovered and that they have potential value as prognostic markers to determine pathological outcome and the likelihood of Acquired Immune Deficiency Syndrome (AIDS) development."

Provided by Virginia Tech