New way of predicting dominant seasonal flu strain

November 15 2010

Rice University scientists have found a way to predict rapidly whether a new strain of the influenza virus should be included in the annual seasonal flu vaccine. While it sometimes takes new flu strains up to three years to become dominant worldwide, the new method can predict whether they will become dominant as little as two weeks after the sequence first appears in the GenBank database, the National Institutes of Health's collection of all publicly available DNA sequences.

"We studied a new strain of the virus that evolved in British Columbia in the middle of March 2009," said Michael Deem, co-author of a new study featured on the cover of the Dec. 12 issue of *Protein Engineering Design and Selection*. "By the end of March, just about two weeks after it came out, we could detect that it would become the dominant strain of H3N2 in 2009. By contrast, it wasn't detectable as a novel strain by the standard methods that the World Health Organization uses until July or the middle of August."

It takes several months to produce the millions of doses of [flu vaccine](#) needed each year, and officials at the [World Health Organization](#) (WHO) use a combination of statistical methods and animal tests to choose the following year's formula.

Just a month before the British Columbia strain was first recorded in GenBank, the WHO had made its recommendations for the annual 2009-2010 vaccine. While the biggest flu story of 2009 was the H1N1 pandemic that began in Mexico and spread rapidly worldwide, the

British Columbia strain went on to become the dominant variant of H3N2 the following year. Because it was significantly different from the H3N2 strain that had been included in the seasonal vaccine for that year, the vaccine's efficacy against British Columbia was estimated at about 20 percent.

"It's not that we could have predicted that British Columbia would have emerged out of thin air, but once it had emerged, our method could detect the signature of its eventual dominance with a very limited amount of sequence data," said Deem, the John W. Cox Professor in Biochemical and Genetic Engineering and professor of physics and astronomy at Rice.

Deem and study co-author Jiankui He, a graduate student in physics and astronomy, developed a mathematical method that used freely available genetic profiles of new flu strains to predict whether a strain will become dominant. Using the method, they examined the H3N2 flu strain for the past 14 years and made their own predictions based on the available data in GenBank, where public health officials post all the latest genetic sequences of new flu strains.

Deem and He compared their predictions with the WHO's predictions from 1996 to 2010. They found their new method correctly predicted the dominant strain of H3N2 for most years, including three years -- 2002, 2003 and 2009 -- when the WHO vaccine was formulated with an H3N2 strain that turned out not to be the dominant strain that year.

The new method involves a statistical technique called multidimensional scaling that is used to create graphical plots of complex data in fields as diverse as marketing and physics. In their study, He and Deem used multidimensional scaling to create a graphical plot of amino acid sequence data for all strains of H3N2. They limited their study to a 329-amino-acid region of the virus that mutates regularly to avoid

detection by immune system.

"Using multidimensional scaling, we project from those 329 dimensions to the two dimensions that contain the most information," Deem said. "We just plot all of the points as a function of two variables instead of listing all 329, which is too much information to work with. With the two-dimensional scaling, we have a workable problem and we still have enough information to see clusters of new strains that will eventually become dominant."

Deem said the results of the study suggests that public health officials could benefit by using the new method, which is both fast and inexpensive, in addition to the well-accepted methods that are currently used to formulate vaccine strain recommendations.

Provided by Rice University