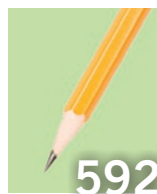**A bone to pick:** A wide range of drugs can inadvertently influence bone health
588

Advice granted: Two experts discuss tips on writing grant applications
592

Gut reaction: Probiotics are pitched to help protect the intestine from HIV
594

Computational tools evolve to reveal patterns in flu data

When a new and suspicious strain of influenza emerges, authorities meticulously plan their public health response on the basis of how dangerous scientists think the strain is or could become. With the emergence of swine flu, or 'H1N1 influenza', this spring, scientists turned to a diverse range of computer-based tools to help predict the virus's virulence.

Rather than focusing on the statistical significance of individual mutations in flu sequences, the latest generation of analytical tools compare multiple genetic signatures among multiple flu viruses at once. These more sophisticated algorithms will prove key in helping to predict the course of future pandemics, some experts say.

Although scientists have started crunching swine flu data with computers, some have also attempted to get an early glimpse into the virus's traits on the basis of manual comparisons to previous flu sequences. Wendy Barclay, chair of influenza virology at the Imperial College, London, says that a comparison of swine flu sequences made available on public databases suggest that two key amino acid residues in the hemagglutinin, or H1, protein of the newly emerged virus are the same as those found on other H1N1 viruses such as the one responsible for the 1918 flu pandemic. Previous studies of the 1918 flu virus have found that these amino acid residues promote virus attachment to cells in the upper respiratory tract, rather than deep in the lungs (*Science* **315**, 655–659; 2007). This suggests that swine flu might spread more easily but perhaps be less deadly than H5N1 avian influenza, which earlier work found binds deep in the lungs (*Science* **312**, 399; 2006; *Nature* **440**, 435–436; 2006).

Influenza data is accumulating thanks to efforts such as the US-sponsored Influenza Genome Sequencing Project and the Influenza Virus Database at the Beijing Institute of Genomics. Today, online databases, such as GenBank (run by the US National Center for Biotechnology Information), hold thousands of genome sequences and allow researchers to compare a swine flu samples with earlier influenza viruses.

But as the amount of available flu data increases, computer models will increasingly help identify new patterns that relate to flu virulence, according to influenza expert Jackie Katz of the US Centers for Disease Control and Prevention. For example, computers can identify mutations that coevolve within viruses and act together to make a flu virus more deadly, says Katz.

Earlier algorithms have focused on single mutations in flu viruses that were individually associated with a statistically significant influence on virulence. But the latest generation of algorithms assesses combinations of mutations that on their own might not have a statistically significant effect but do seem to influence virulence when they appear together. For example, Jonathan Allen and Tom Slezak, both computer scientists at the US Department of Energy's Lawrence Livermore National Laboratory in Livermore, California, have developed an algorithm that looks for combinations of mutations that viruses share.

Just this past April, Allen and Slezak reported the findings of their analysis of samples from 2,100 influenza cases representing numerous flu viruses that occurred within the past century—including the past three major pandemic outbreaks, which occurred in 1918, 1957 and 1968. The team found 34 amino acids conserved at specific points in the genomes of all three of these pandemic viruses (*BMC Microbiol.* **9**, doi:10.1186/1471-2180-9-77; 2009). On the basis of their examination of a 2009 swine flu sequence, they determined the strain had only half of the 34 amino acid markers correlated with the deadly pandemics.

Allen explains their computational tool could help shed some light on the virulence of swine flu and "point[s] out [genome] locations of interest that can be subjected to further experimental studies." But although the computer model can sift through piles of genomic data, it does not take into account certain factors, such as interactions between flu virus genes and host immunity, that also influence how sick a person will become.

"From a computational stand point, developing new ways to selectively prune the expansive space of candidate marker combinations remains an area for continued development," says Allen.

Preparing for the worst: Swine flu has caused fear worldwide

Beyond sequencing

In addition to sequence data, researchers are plugging epidemiological evidence into a mathematical model to approximate how virulent the 2009 H1N1 strain is and how quickly it spreads. An early use of this approach has suggested that in Mexico swine flu is fatal in around 4 in 1,000 cases (*Science* doi:10.1126/science.1176062; 2009). But until experts can obtain data from people who had silent or very mild illnesses, much uncertainty remains.

In addition to computational analyses, researchers are testing how the virus behaves in tissue and animal models. Scientists are using a technique called reverse genetics in which they introduce mutations into the swine flu sequence and test the mutated version in these models. Barclay emphasizes the importance of these experiments: "you can't always be sure that what you have predicted from the [genetic] sequence will be reflected" in the biological action of the virus.

William Gallaher, a professor emeritus at Louisiana State University in New Orleans, cautions against blind faith in computers when it comes to flu prediction; he says, "if you project out a model by computer or any other way, it's dependent on the assumptions that you put into the model. Computers are just going to take those assumptions and project them."

"The assumptions have to be correct in order for the outcome to be correct," he adds.

Kirsten Dorans, New York