Virus hunters tracking down the next pandemic

05 October 2016

by Gary Finnegan

Genetic detectives have used big data to map early HIV epidemics – now they hope that new technologies can help doctors respond to outbreaks of Ebola, Zika and flu in real time.

Deadly viruses are a major threat to public health. Not only can they kill, they are often unpredictable and are capable of changing over time.

Just as the sudden surge in Ebola cases took the world by surprise in 2014, the little-known Zika virus has dominated the health headlines in 2016 after being declared a public health emergency by the World Health Organization in February.

However, the spontaneous genetic changes – or mutations – that can, for example, help influenza viruses to evade the annual flu vaccine can also be used to track their spread. Scientists have developed sophisticated algorithms that allow them to trace the origins of viral outbreaks, offering a window on history, and helping to predict the future.
Professor Philippe Lemey of the University of Leuven in Belgium examines telltale mutations in the genomes of viruses to learn about their past. Like a cold-case detective using new tools to reopen an investigation, he has studied the spread of rabies, influenza and HIV.

‘Studying HIV is what motivated me to work in this field,’ he said. ‘There are many viruses that have made the jump from primates to humans but only HIV became pandemic. We wanted to know why.’

Prof. Lemey, who ran the VIRALPHYLOGEOGRAPHY project, funded by the EU’s European Research Council (ERC), says the answer lies in a ‘perfect storm of factors’, including the development of railways in Kinshasa, in the Democratic Republic of Congo (DRC), in the early 20th century. By poring over data on the genetics of a major strain of HIV, the team was able to build a kind of family tree for the virus.

From there they mapped the evolution of HIV and its explosive spread in the DRC. Once they had figured out the geographical spread of the disease, they saw a pattern: HIV was spread through the fledgling Congolese rail network.

‘The virus seems to have followed transport links, particularly along the railway network which was developing rapidly at that time,’ Prof. Lemey explained. ‘Alongside transport, there were cultural changes too – the behaviour and location of sex workers was shifting and this is likely to have accelerated the spread.’

The power of these genetic tools is not limited to unravelling the mysteries of historical epidemics. Just as the Kinshasa railway network maps onto the spread of HIV, today’s aviation networks help to predict the spread of flu viruses.

Predicting flu

Knowledge of transport links and genetic analysis can be a useful guide to how outbreaks will progress once they begin. Indeed, experts already apply this logic when predicting which strains of flu should be included in the annual flu vaccine. If they know which strains are spreading in the southern hemisphere they can make an educated guess about which ones are likely to dominate the winter flu season north of the equator.

‘We don’t predict the location of the next big outbreak, but once a virus has entered the human population it may be feasible to predict how it will spread spatially,’ Prof. Lemey said. ‘If you can characterise the patterns of spread, you can estimate what will happen next.’
Researchers were able to map the spread of a swine flu outbreak between 30 March 2009 and 12 July 2009 by analysing the genetic sequences of 242 virus samples. Video courtesy of Prof. Philippe Lemey

The power of this approach to disease mapping will be dramatically improved thanks to advances in genetic sequencing. The availability of portable gene-sequencing machines during the Ebola outbreak generated a wealth of data.

Genomic data has been collected from around 5% of all Ebola cases during 2014 and 2015. Despite pressure on health systems in West Africa, mobile diagnostic kits meant that results were available within hours of taking a blood sample from an infected patient. Now, experts are working on how this data can be used to inform on-the-spot policy decisions, such as targeting resources towards preventing the spread of the disease.

Scientists want to build an automatic system that would incorporate each new piece of genetic information into an algorithm to calculate what is likely to happen next. ‘You can identify transmission hotspots,’ said Prof. Lemey. ‘For example, we know Ebola outbreaks were seeded in capital cities – this helps you to target public health interventions at high-risk areas.’

Real-time sequencing

Professor Oliver Pybus of Oxford University, UK, who worked on the original HIV investigation described by Prof. Lemey, believes that this kind of real-time gene sequencing and disease surveillance could revolutionise healthcare the world over.

‘In the not-too-distant future, every doctor could have a real-time genetic sequencing device on their desk,’ he said. This would help doctors choose the right treatment for their patient while also generating a huge volume of data that could be analysed by researchers.

Prof. Pybus, with the support of the ERC, is examining how to get the most out of this data. ‘We will have more data than we currently know what to do with,’ he said. ‘So instead of playing catch-up, we are looking forward five-to-10 years and developing methods for interpreting large volumes of genomic information.’

Provided all the data is shared, outbreaks of diseases would be picked up before they could get out of control. New treatments and vaccines could be developed or, where these already exist, deployed to where they are needed most.

‘Genetic data can tell you things that no other data can,’ he said. ‘HIV and hepatitis C were only discovered in the 1980s, but had been infecting millions of people long before that. There are no useful medical records on patients before that time but mathematical modelling, using genetic data, can take us back much further.’
In future, real-time data on HIV and Hepatitis C outbreaks – paired with knowledge of historical epidemic patterns – would allow rapid intervention among communities where the disease might otherwise find its next victims.

More info

VIRALPHYLOGEOGRAPHY
PATHPHYLODYN