New maths to predict dangerous hospital epidemics

August 17 2016, by Yngve Vogt

1.5 million people worldwide die of tuberculosis annually. In the Russian prison in Kemerovo in Sibir, more than 1000 prisoners suffer from this heinous disease. Two years ago professor Jukka Corander made a mathematical model that explained why the tuberculosis epidemic spread so quickly. Credit: NTB/Scanpix

Mathematicians are now developing completely new statistical calculations on the world's fastest computers in order to predict how epidemics of dangerous hospital bacteria spread.

'We are going to use mathematical models and statistical methods to understand why and how lethal, drug-resistant hospital bacteria arise and spread,' says Professor Jukka Corander at Oslo University's new innovation centre, Big Insight. The Innovation Centre was started two years ago, by Professor Arnoldo Frigessi, for crunching enormous amounts of data using completely new statistical methods.

Studying the entire genomes of bacteria has now thrown open entirely new possibilities for revealing their secrets. It is this genetic knowledge that scientists use to understand bacterial epidemics.

A single bacterial genome consists typically of several million building blocks. The individual building blocks may be one of four compounds: adenine, thymine, guanine and cytosine, popularly called A, T, G and C. The order in which they are assembled is the actual recipe for the genome.

Checking the genes of a single bacterium is not enough. The scientists have to check the genes of thousands of bacteria in order to understand what the individual genes do and how the population evolves.

The dataset quickly becomes so enormous that it is not possible to interpret by means of ordinary computing operations.

In order to reveal the secrets of the bacteria, statisticians are now developing entirely new mathematical methods for describing the evolution of bacteria, how genomes change over time, and what affects these genes.

Deadly bacteria

There are a number of deadly hospital bacteria. The best known is a type of Staphylococcus aureus called MRSA. This feared bacterium kills 25

000 Europeans each year.

Since the 1990s, certain Enterococci have caused a series of severe epidemics at many hospitals all over the world. The mortality rate can be high. Enterococci constitute a natural part of the intestinal microbiota, but unfortunately also include some highly dangerous species. Like other nasty bacteria, they may circulate from the intestines via the blood and infect the heart valves.

Both MRSA and some enterococci are multi-antibiotic-resistant and therefore very difficult to eradicate.

The bacteria unfortunately have the feared ability to warn others of their kind by transferring genes to them when they are subjected to antibiotics attack. This is called horizontal gene transfer, and makes hospital bacteria even more difficult to gain control of. In other words, the bacteria can borrow genes from one another. This mechanism gives them extra resistance to antibiotics therapy.

In 2013, while he was professor at the University of Helsinki, Jukka Corander discovered that hospital enterococci resembled bacteria in animals more than the bacteria in the human intestinal system. Antibiotics are used extensively in veterinary medicine, and in many countries it is still permissible to feed animals with antibiotics to increase their muscle mass. Farmers may then bring infection from their animals to hospitals.

'This is how some epidemics have started in countries with intensive agriculture, such as the Netherlands and Denmark,' states Jukka Corander.

He has used his mathematical methods to analyse why these bacteria become more dangerous and more infectious by sharing their genes with

other bacteria.

Tuberculosis bacteria

Two years ago, Jukka Corander published a major study on genetic changes in tuberculosis bacteria in *Nature Genetics*. One and a half million people die each year from this unpleasant lung disease. Every third person in the world is a carrier. In the former Soviet countries, the proportion is as much as a half. But it is only when the illness breaks out that the carriers themselves become infectious.

Scientists performed genetic tests on the bacteria from sputum samples from a thousand tuberculosis patients. The statisticians were then able to explain why the tuberculosis epidemic spread so fast in Russia. The bacteria became extra drug-resistant through a mutation process.

'Normally becoming resistant to antibiotics comes at an extra evolutionary cost to bacteria. They grow more slowly and divide more inefficiently. We discovered that certain mutations in the bacteria counteract this effect, with the result that they were able both to be drug-resistant and to grow just as well. This is an enormous advantage for the bacteria, because they can spread much faster,' explains Corander.

The objective of his research is to optimise treatment of patients who are infected by dangerous hospital bacteria. Many types of antibiotics have severe side effects, and must therefore only be used when absolutely necessary. Sometimes no antibiotics are effective.

'There are at least seventeen different antibiotics against tuberculosis. The treatment is normally a combination of two or three different drugs, and may take two years. We know which bacterial genes cause which drug-resistance. The gene test enables hospitals to drop antibiotics that are not efficacious.'

Dangerous bacteria in refugee camps

Jukka Corander has primarily investigated how fast bacteria become drug-resistant, and where the different types of tuberculosis bacteria are to be found.

One of the studies is from a Karen refugee camp located between Burma and Thailand. There he discovered enormous variations in the antibiotic-resistant bacteria. Corander compared the genomes of 3000 bacteria, and found 20 per cent variation in the genomes of individual bacterial species. This is a huge difference. The genetic difference between humans and chimpanzees, by way of comparison, is only one per cent.

'This shows how fast bacterial evolution takes place. The bulk of the evolutionary changes are due to the fact that bacteria borrow genes from one another. They have different strategies for this borrowing. Some bacteria have an enormous capacity for borrowing genes from one another and passing the genes on to others when they are subjected to a selective pressure, such as antibiotics.'

By combining knowledge of bacterial genes with which types of antibiotics were prescribed when, Jukka Corander was able to use mathematical models to calculate back in time to a year or twenty years ago. This enabled him to find out which genes the bacteria lent one another when, and to calculate which types of antibiotics led to which type of drug-resistance.

The amount of data was so enormous that the statisticians used a supercomputer running two thousand parallel computations for two full months before they got the answers. On an ordinary PC, these computations would have taken 330 years.

The massive need for computations does not frighten Jukka Corander.

He is now busy studying a dataset ten times larger on pneumococci bacteria. This is a collaboration with the Sanger Institute, near Cambridge, England, which is a world leader in bacterial research. Their speciality is how bacteria spread, and how they develop so as to become even more dangerous.

'In 2015 there were 100 000 bacterial genome sequences. This year the figure is one million. The amount of data is exploding. In three years, it will be a hundred times larger again, and in five to ten years perhaps a thousand times larger. We cannot rest. We do not have good enough algorithms for the future. Today's algorithms are far too slow. In order to analyse this vast quantity of data, we must develop far faster algorithms.'

If not, the scientists risk computations taking years, or even ten thousand years, before we get the answers.

'We hope that in a few years we will have made algorithms that are so fast that we can issue a warning when a new epidemic is starting, rather than being able to say how it developed, as is the case today,' states Jukka Corander.

If we are to be able to make even better analyses in the future, the underlying data must also be better.

'Every time a patient is admitted to a hospital in England, genetic tests for various hospital bacteria are carried out. If the patient is infected, the entire genome of the hospital bacterium is sequenced. This will make it possible at a later stage to determine the most probable cause of the infection and where it originated. We will also be able to immediately apply a simple gene check to determine what sort of drug-resistant bacteria the patient is infected with. Today it takes several days to get answers. Faster answers will enable patients to get the right treatment much more quickly.

The new mathematical and statistical methods are available on the internet.

'Bacteria cause several billion infections each year. It is often the same types of bacteria that kill and disable people. Our aim is to reduce the burden of disease on ordinary people, the spreading of infection and the number of illnesses in the world', affirms Professor Jukka Corander at the new Big Insight research centre, which is physically located at Oslo University's Faculty of Medicine.

Provided by University of Oslo