Science and Technology Facilities Council

Scientific Computing

Can my cheeseburger harm me?

Using computers to study the genomics of *E. coli*.

Every year, around 2.8 million cases of food poisoning from *Escherichia coli* (*E. coli*) are reported worldwide. Often, these harmful bugs get into our bodies by eating unwashed, contaminated salad leaves or undercooked and processed meats such as those used in burgers.

E. coli are generally benign inhabitants of our gut but occasionally we are exposed to more toxic variants such as Shigatoxigenic *Escherichia coli,* notorious foodborne pathogens that are considered to be a public health risk.

The most common form of this pathogen is Shigatoxigenic *E. coli* (STEC) serotype O157:H7. In 2019 there were 539 confirmed cases of food poisoning caused by STEC O157 in England and Wales, with nearly one-third of those in England being hospitalised. However, an increasing number of infections are now caused by non-O157 STEC, causing uncertainty about the disease and the methods of controlling it.

Scientists from STFC Scientific Computing and Scotland's Rural College (SRUC) are now finding ways to combat this issue, which could lead to the development of a simple PCR test (similar to those used to identify Covid-19) to test for types of *E. coli.*

E. coli Image Credit: NIAID. Licensed under creativecommons.org/licenses/by/2.0/deed.en

The Challenge

With the rise in infections now being caused by non-O157 STEC, there is uncertainty about the pathogenicity (ability to cause disease) of new strains, which methods of control to use, and the likelihood of disease. Food-borne diseases are caused by the consumption of a poisonous microbe. If scientists can identify what type of microbe is causing the food poisoning, they can determine the methods of control and prevent the spread of disease to other people.

Our Approach

Computational scientist Dr Martyn Winn from STFC Scientific Computing and Dr Nicola Holden, a Professor in Food Safety at SRUC, joined forces to explore this issue. They investigated the genomic signatures (the types of genes in a cell) of non-O157 STEC to build up knowledge of these bacteria.

"This small scoping study enabled a new partnership to be built, allowing new ways of working that would otherwise not be open to me. It has resulted in some terrific opportunities for how to take the science forward for much needed diagnostics and surveillance of a priority pathogen".

Dr Nicola Holden, SRUC

"Our project was to look at the genomic signatures of non-O157 *E. coli* - so what are the characteristics of the bacteria; when do they become pathogenic and how can we identify them?

"*E. coli* have between four and five thousand genes and around two thousand variable genes, creating a wide diversity of strains. Understanding which genes are present in the pathogen makes it possible to determine the likelihood of the pathogen causing disease."

Dr Nicola Holden, SRUC

The scientists used data from 229 isolated samples from food poisoning outbreaks. The samples came from sources such as cheese (thought to be the source of an *E. coli* outbreak in Scotland in 2016), minced meat, wild deer (venison), and clinical samples from affected patients.

The samples had been supplied by the UK Health Protection Agency reference laboratories, and the Scottish *E. coli* reference laboratory, which extracted and sequenced the DNA.

Dr Winn mapped the genomes to indicate their potential to be pathogenic. Using SCARF, one of Scientific Computing's advanced computer platforms, he ran data from the samples through various bioinformatics pipelines. His aim was to determine which genes were present or absent in the samples.

Doctors Winn and Holden then compared the genomes to determine how the disease evolved and the relationship between the sample and the suspected source of the *E. coli* outbreak.

Benefits

This work will help identify different and emerging strains of STEC, predict the likelihood and severity of the disease, and, ultimately, prevent it from spreading.

Further work is needed but there is potential to have a PCR test, customised to one or two genes, for patients with possible food poisoning. This would produce an immediate result, indicating the strain of STEC causing the disease and enabling faster treatment for patients.

"We have a path; we know that there are more bioinformatics we need to do to pin down what genes are useful to identify specific STEC strains that are becoming more common in cases of food poisoning. We already have companies that are interested in producing a test for use in hospitals, in facilities where food is prepared, or even in the environment where wild deer can be found. The test would identify these strains and determine the likelihood of disease. The work will take a couple of years and more funding, but there is a clear path to impact."

Dr Martyn Winn, STFC Scientific Computing

This initial project was funded through a grant from the STFC Food Network +, which brings together STFC researchers and facilities with research and industry in the agri-food sector. **stfcfoodnetwork.org**

Case study produced by Briana Doherty, Industry Placement Student, STFC Scientific Computing