

Applying Deep Learning to Salmonella Genomics: A New Horizon for Source Tracing of Pathogens in the Food Chain

Supervisory team:

Main supervisor: Prof Kristen K Reyher (University of Bristol)

Second supervisor: Prof Andrew Dowsey (University of Bristol)

Dr Sion Bayliss (University of Bristol), Dr Lauren Cowley (University of Bath)

Collaborators: Dr Claire Jenkins (Gastro and Food Safety One Health Division, UK Health Security Agency), Dr Marc Allard (Center for Food Safety and Applied Nutrition, US Food and Drug Administration), Dr Erica Rose (Centers for Disease Control and Prevention)

Host institution: University of Bristol

Project description:

Infections caused by members of the bacterial genus *Salmonella* represent a major threat to both animal health and public health; requiring significant resources to identify sources of infection. Over the last decade, food safety and public health agencies have begun to routinely utilise whole genome sequencing (WGS) for pathogen surveillance. WGS is a powerful tool which provides information on genetic relatedness between disease isolates, allowing for confident identification of clusters of infections arising from common sources. Traditional investigative research approaches, however, are unsuitable for application to the increasingly vast amount of WGS data being generated by surveillance of the global food supply chain. Novel tools are required to translate the large amounts of WGS data gathered by surveillance programmes into actionable information for disease control.

The successful candidate will be supervised by world-leading researchers at the Universities of Bristol and Bath, and will liaise and intern with epidemiologists at the US Food and Drug Administration, UK Health Standards Agency and the US Centers for Disease Control to develop cutting-edge machine learning (ML) models to predict the source of *Salmonella* disease outbreaks. The project will innovate and refine upon recent work (DOI: <https://doi.org/10.7554/eLife.84167>) to identify optimal state-of-the-art machine learning frameworks for application to bacterial genomic data, optimise a hierarchical models to predict the source of outbreak isolates and generate human-readable outputs for epidemiologists. The resulting tool will allow for rapid and effective predictions of the source of *Salmonella* infections which will be used to support public health decision making and facilitate rapid responses to future outbreaks. The studentship will be hosted in the University of Bristol, which is ranked 5th in the UK for research. The student will be supported by a comprehensive multi-disciplinary supervisory team who are embedded in a number of rich active research communities in the areas of genome bioinformatics, epidemiology, machine learning, antimicrobial resistance and public health in the Universities of Bristol (Prof Kristen Reyher, Prof Andrew Dowsey, Dr Sion Bayliss) and Bath (Dr Lauren Cowley).

This project would suit an applicant with a strong first degree or masters involving bioinformatics or computational biology. Students with a mathematical background open to learning skills in bioinformatics/ML or those with a biological/biomedical background and experience or interest in the areas of basic programming, data science or ML are encouraged to apply.

Our aim as the SWBio DTP is to support students from a range of backgrounds and circumstances. Where needed, we will work with you to take into consideration reasonable project adaptations (for example to support caring responsibilities, disabilities, other significant personal circumstances) as well as flexible working and part-time study requests, to enable greater access to a PhD. All our supervisors support us with this aim, so please feel comfortable in discussing further with the listed PhD project supervisor to see what is feasible.