Identifying genes associated with severity and spread of infections

Supervisors: Xavier Didelot and Noel McCarthy

Evolution in bacteria, as in humans and other forms of life, is driven by genetic variation and selection. Different forms of a gene, or the presence or absence of a gene can lead to different function or phenotype. For example bacteria might cause more or less severe disease, or survive better in the environment, or in one host species compared to another according to their genetic make-up.

However since bacteria do not reproduce sexually they have an intense population genetic structure. This is described as clonal with the off-spring a clone of the parent. Populations from a common ancestor thus share a lot of genetic material. This structure complicates analysis. The tree like relatedness must be estimated and then adjusted for in all analyses, while accounting for recombination that can disrupt the patterns of clonal evolution. Additionally, the large amount of information (over 1,000,000 potentially variable sites in each bacterium, plus variations in gene content, and analyses including thousands of bacterial strains) complicates analysis and increases computational requirements.

This project will work on extensive available genetic data to link genotype with phenotype. This can include predicting antibiotic resistance, measuring survival of infections through the food chain, better survival in one animal host (e.g. cattle) than another (e.g. sheep). Analysis will use a suite of available programmes developed by Prof Didelot (TreeWAS) and also compare approaches such as machine learning.

The initial MSc project will focus on a single question and technique. Extending the project to a PhD the student would apply and develop other approaches and compare them as well as addressing a range of substantive questions. This will be in collaboration with Public Health England and the Animal and Plant Health Agency with access to national data.

The student would have the opportunity to spend time in Public Health England and work as part of and the NIHR Health Protection Research Units funded at Warwick from April 2020 in collaboration with Public Health England and partner universities. Students would therefore have access to the NIHR academy and professional training offered by the NIHR (National Institute for Health Research).

Starter references

1. Earle SG, Wu C, Charlesworth J, et al. Identifying lineage effects when controlling for population structure improves power in bacterial association studies. Nat Microbiol 2016; 1:16041
2. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Comput Biol 2018; 14:e1005958
3. Méric G, Mageiros L, Pensar J, et al. Disease-associated genotypes of the commensal skin bacterium Staphylococcus epidermidis. Nat Commun 2018; 9:5034
4. Sheppard SK, Didelot X, Meric G, et al. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. Proc Natl Acad Sci USA 2013; 110:11923–7