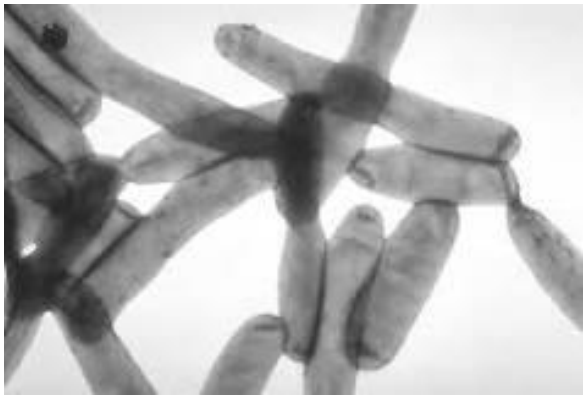
PROJECT TITLE: **Improving genomewide association study strategies for application on environmental bacteria**

DTP Research Theme: **Living World**

Lead Institution: **University of Bath**

Lead Supervisor: **Dr Lauren A. Cowley, University of Bath, Milner Centre for Evolution**
Co-Supervisor: **Prof. Samuel K. Sheppard, University of Bath, Milner Centre for Evolution**
Co-Supervisor: **Prof. Ed Feil, University of Bath, Milner Centre for Evolution**
Co-Supervisor: **Dr Ruth C. Massey, University of Bristol, School of Cellular and Molecular Medicine**
Project Enquiries: **drlacowley@gmail.com**



Project Background

Genomewide association studies (GWAS) are a powerful tool in discovering the genetics underlying observed important phenotypes and have been used very successfully in human genetics for several heritable diseases. In bacteria that reproduce asexually, the challenge with GWAS is the high linkage between polymorphisms due to low levels of recombination. A lot of the mutations or genes in bacterial genomes are not acquired independently and it is therefore difficult to differentiate the mutation or gene that is causing the phenotype change from those that are simply hitch hiking with it (linkage disequilibrium).

Methods developed to achieve bacterial GWAS results of high confidence could be extremely useful in determining the genetic elements associated with phenotypes of environmental importance, such as toxin production, environment survivability, antibiotic resistance or host niche. Previous studies by co-supervisors Professor Sam Sheppard and Dr Ruth Massey (Sheppard et al., 2013, Recker et al., 2017) have shown the capabilities of GWAS in bacteria.

Project Aims and Methods

NERC funds world-leading basic, strategic and applied research. This project comes under the Living World NERC research theme and more specifically under the genomes and evolution section within that. This studentship will use state of the art genomics and cutting edge computational tools to wrangle big data to understand phenotypes of environmental bacteria.

The aims of the studentship will be:

- To develop a novel approach of longitudinal and spatial sampling in big data that offers an opportunity to resolve genuine associations from spurious one's due to linkage.
- This will be tested by using datasets that are deeply sampled longitudinally and testing time periods individually to see if association gene lists change over time and if certain genes remain constant in their assigned significance. In recombining bacteria, sampling over time causes linkage between loci to decline (Arnold and Hanage 2017).
- Simulations will be run using the state of the art computational simulation tool Fwdpp (Thornton 2017) to gauge the optimum time period needed to allow enough recombination to happen to decrease spurious linkage enough to make GWAS accurate. The student will aim to estimate the

minimum number of required genomes and frequency of sampling when given the recombination rate of the species.

- Systematic testing will be performed in a variety of large genomic datasets to develop a robust protocol to provide high confidence results.

To limit the complexity of linkage disequilibrium, highly recombinogenic environmental bacteria such as *Legionella pneumophila* datasets will be used in initial strategies.

Candidate Requirements

Essential:

- A high scoring degree in microbiology, biology or computer science (2.1 or higher)

Preferred:

- A master's degree in Bioinformatics or computer science
- Experience of at least one computer programming language

Training

The NERC Great Western Four+ Doctoral Training Partnership is designed to train tomorrow's leaders in earth and environmental sciences. A wide depth of expertise in Bacterial genomics and GWAS experience is held by the supervisors. Lauren Cowley is an expert in microbial bioinformatics and dealing with very large datasets. Ed Feil is a very experienced expert in population genetics in bacteria. Sam Sheppard and Ruth Massey will provide experience as early innovaters in bacterial GWAS.

Key skills the student will acquire:

- The ability to apply cutting-edge computational tools to Big data genomic datasets
- Excellent programming skills transferable to a wide variety of modern careers
- Generalizable data science skills

References / Background reading list

Genome-wide association study in *Campylobacter*
Samuel K. Sheppard, Xavier Didelot, Guillaume Meric, Alicia Torralbo, Keith A. Jolley, David J. Kelly, Stephen D. Bentley, Martin C. J. Maiden, JulianParkhill, Daniel Falush
Proceedings of the National Academy of Sciences Jul 2013, 110 (29) 11923-11927; DOI:10.1073/pnas.1305559110
Clonal differences in Staphylococcus aureus bacteraemia-associated mortality
Mario Recker, Maisem Laabei, Michelle S. Toleman, Sandra Reuter, Rebecca B. Saunderson, Beth Blane, M. Estee Török, Khadija Ouadi, Emily Stevens, Maho Yokoyama, Joseph Steventon, Luke Thompson, Gregory Milne, Sion Bayliss, Leann Bacon, Sharon J. Peacock & Ruth C. Massey
Nature Microbiologyvolume 2, pages1381–1388 (2017)
Longitudinal samples of bacterial genomes potentially bias evolutionary analyses
Brian John Arnold, William P Hanage
bioRxiv 103465; doi: https://doi.org/10.1101/103465
A C++ Template Library for Efficient Forward-Time Population Genetic Simulation of Large Populations
Kevin R. Thornton
GENETICS September 1, 2014 vol. 198 no. 1 157-166; https://doi.org/10.1534/genetics.114.165019

Useful links

Enquiries relating to the project should be directed to the lead supervisor (see email address above for Project Enquiries). Enquiries relating to the application process should be directed to doctoraladmissions@bath.ac.uk

In order to apply, you should select the relevant University of Bath PhD online application form found here: https://www.bath.ac.uk/study/pg/applications.pl. When completing the form, please state in the 'Finance' section that you wish to be considered for GW4+ DTP funding and quote the project title and lead supervisor's name in the 'Your research interests' section.

Further information about the application process may be found here: http://www.bath.ac.uk/topics/postgraduate-research/

The application deadline is 1600 hours GMT Monday 7 January 2019 and interviews will take place between 4 and 15 February 2019. For more information about the NERC GW4+ DTP, please visit https://nercgw4plus.ac.uk.