Project title: Identifying drivers for antibiotic resistance in river systems

Theme: Organisms and Omics

Key words: metagenomics, bioinformatics, antibiotic resistance, model development

Supervisory team:
Prof. Elizabeth Wellington (University of Warwick: E.M.H.Wellington@warwick.ac.uk);
Dr Christopher Quince (University of Warwick: c.quince@warwick.ac.uk);
Dr Andrew Singer (NERC CEH Wallingford: acsi@ceh.ac.uk)

Project Highlights:

- Analyse prevalence of antibiotic resistance from the Thames watershed
- Study evidence for selection via antibiotic exposure in the river
- Resolve environmental fate of resistance genes

Overview: The development of resistance to antibiotics by human pathogens is one of the greatest threats currently faced by mankind. It has been estimated that if antibiotics become ineffective then routine operations such as hip-replacements could lead to death for 1 in 6 patients. Until recently this was studied purely as a medical problem but it is now realised that the environment has a key role to play as a reservoir of antimicrobial resistance (AMR) genes. AMR genes possessed often as plasmids or other horizontally transmitted elements (HTEs) can be transferred from environmental organisms to pathogens providing them with resistance. The challenge is how to study AMR in environmental organisms that consist of a complex community of possibly unculturable organisms. Metagenomics the sequencing of DNA from the entire community in situ provides a solution to this problem.

The TMD will provide an invaluable resource on the prevalence of AMR genes in the environment but in its raw form it will consist of billions of short reads. To unlock its full potential we will assemble these reads together and then use algorithms developed by Dr Quince (e.g. CONCOCT) to link the assemblies into genomes. Only then will it be possible to determine which environmental organisms possess which AMR genes that are potentially being transmitted to pathogens. This will be challenging and may require the use of some additional novel algorithms that are being developed in the Quince group currently. The analysis will allow us to go on and explore AMR in addition publicly available metagenomes from environments such as the human gut (HMP) and the sea (Tara Oceans) have greatly increased our understanding of the role of the environment in AMR. In addition further sampling will be done to establish selection effects in relation to antibiotic exposure in fish farms where various antibacterial agents have been used. Experimental selection following on from field data will allow a study of gene dissemination and the selective effects of tetracyclines and sulfonamides in model river systems. We hypothesize that plasmids bearing resistance genes improve fitness of the host even in the absence of selection and this effect will occur when the plasmids spread to other hosts in the river, thus spreading the antibiotic resistance genes (ARGs) without the need for antibiotic selection.
Fig 1. Distributions of ARG proxy gene IntI1 across three seasons and 69 sample points on the Thames catchment 2015-2017.

Methodology: Phase 1. The TMD will provide us with metagenome sequences from microbial communities at multiple times from the Thames catchment (TMD – NE/M011259/1). We will pool all these samples together and co-assemble the short reads into longer fragments or contigs. This will generate many hundreds of thousands of contigs deriving from thousands of microbial genomes. We will use statistical techniques that exploit the fact that contigs from the same genome will co-occur across samples to group contigs into the genome they derived from. This will enable us to reconstruct organisms genomes. We will then identify genes within these genomes, both taxonomic markers and AMR genes. In this way we will determine exactly which organisms in the environment are acting as reservoirs for AMR. We will place these results into a wider global context by analiaysing both further metagenome data sets for AMR and by cross-referencing to isolate genome databases for specific organisms. Phase 2. The analysis of field data will seek to resolve evidence for selection in the presence of antibiotic exposure in the river- for example proximity to fish farms or run off from dairy farms. We will seek to establish if the reservoirs for these resistance genes are different from those observed in teh metagenome analysis across the entire river system. Phase 3. We aim to investigate the mobility of plasmdis in model river systems using sampoles taken from the Thames to establish if selection impacts on plasmid spread and if plasmid move within teh reservoirs defines in Phase 1 or outside the reservoirs into other hosts. The use of antibiotic additions will establish the effects of variable selection on plasmid stability and spread.

Training and skills: CENTA students are required to complete 45 days training throughout their PhD including a 10 day placement. In the first year, students will be trained as a single cohort on environmental science, research methods and core skills. Throughout the PhD, training will progress from core skills sets to master classes specific to the student's projects and themes.

The student will receive extensive training in microbial bioinformatics both metagenomics and genomics. They will learn how to run bioinformatics software for assembly, mapping, gene annotation, phylogenetic tree construction etc. They will gain familiarity with sequence databases such as the ABR and NCBI. They will be taught programming in Python, and C, basic statistics using R and more advanced Bayesian statistical modelling. They will also be given an overview of the molecular biology behind the generation of the data sets, sample preparation and next generation sequencing although the focus of the studentship will be on the bioinformatics and statistics.

Partners and collaboration (including CASE): The NERC Centre for Ecology and Hydrology will be a project partner (Dr Andrew Singer). The student will have the opportunity to spend some time visiting Dr Singer to work on integrating the metagenome data into a landscape setting. The bioinformatics training will in part take place through the MRC funded Cloud Infrastructure for Microbial Bioinformatics (CLIMB) consortium (http://www.climb.ac.uk). This project is a collaboration between Warwick and CEH.

Possible timeline:
Year 1: Initial quality control of TMD, assembly, mapping, annotation and contig clustering. Determine the organisms present, their abundances across samples and which possess AMR genes. Study of TMD metadata to determine if evidence for selective effects of antibiotics.
Year 2: Resolve distributions of ARG in host reservoirs and establish evidence for spread under selection. Resolve in detail where ARG are most prevalent and in which hosts.
Year 3: Test hypothesises for plasmid spread and ARG persistence in model river systems in presence and absence of selection. Establish if plasmid impacts on host fitness can override effects of carriage and gene expression in absence of antibiotic selection.

Further reading:

Further details:
Prof. Elizabeth Wellington (University of Warwick: E.M.H.Wellington@warwick.ac.uk)
Dr Christopher Quince (University of Warwick: c.quince@warwick.ac.uk)