Project Details	
Project Code	MRCIIAR25Ba Iqbal
Title	Using AI to understand defence and counter-defence systems between bacteria and phage
Research Theme	Infection, Immunity, Antimicrobial Resistance & Repair
Summary	Bacteria have been engaged in a war with the viruses that predate on them ("phage") for millions of years, both sides developing defence/counter-defence systems. We humans would like to use phage as an alternative to antibiotic therapy, but it is hard to decide which phage will successfully kill a specific bacterial strain, in the light of the huge diversity of defence/counter-defence systems. This project set out to apply statistical and modern AI methods to rich new experimental datasets, to build a deeper understanding of the interactions between bacteria and phage and discover how their defence systems work and interact.
Description	This project aims to apply statistical and modern AI methods to study one of the most rapidly developing fields in modern biology - the study of how bacteria defend themselves against the viruses that infect them ("phage"), and how phage in turn counteract those defences. The co-supervisors combine expertise in bacterial genomics, bioinformatics, big data, mathematics (Iqbal, Bath)[refs 1,2,3] and bacteria/phage interactions, high throughput phenotyping, defence and anti-defence systems (van Houte, Westra, Exeter) [refs 4,5,6] and are already equipped with extremely rich datasets for analysis. This is an opportunity, thanks to the unique datasets, rapid innovation in modern AI methods, and great expertise within the teams within which the student would be embedded, to make real discoveries impacting both our understanding of fundamental biology, and the development of phage therapy. **Background:** The simplest way for bacteria to become resistant to phage is by mutation of the cell surface receptor that phage use to attach to the cell. However, mutation of these receptors is often associated with significant fitness costs, as receptors carry out important cellular functions. An alternative strategy to resist phage is through cellular "immune systems" that act inside bacterial cells to clear infections. Probably the best-known bacterial immune systems are Restriction-Modification (RM) and CRISPR-Cas, which revolutionised our ability to manipulate DNA in vitro and in vivo. Bacteria can also trigger dormancy/cell death upon infection to prevent mobile genetic element (MGE) replication. Literally dozens of previously unknown defence systems have been discovered in recent years, often aided by their clustering in defence-islands. The first systematic, large-scale bioinformatics prediction and experimental validation study led to the discovery of 9 new families of antiviral defence and a novel family of anti-plasmid defences. A subsequent study identified a further 29 antiviral cassettes that are present in approximately a third of all bacterial genomes. Using similar approaches, a vast diversity of bacterial defences has been revealed. These different defences act at different stages of the MGE lifecycle: some cleave MGE genomes immediately following infection, others interfere with MGE transcription or replication or induce cell death or dormancy responses.

	These diverse defence systems frequently coexist in the same genome and it has been hypothesised that bacterial defence systems consist of multiple integrated layers that act in concert to constrain MGE infections, by providing broader spectrum or higher levels of defence than single systems. In response to the evolution of bacterial immune systems, phages have evolved an equally diverse set of anti-defence systems (ADS). ADSs can counteract bacterial defense systems by modifying the target of the bacterial defence system, interference with the activation of bacterial defenses, inhibition of defence surveillance and effector proteins or by alleviating the impact of programmed cell death or dormancy induction. Crucially, while DS and ADS are very common in bacterial and phage genomes, we lack a fundamental understanding how important they are relative to receptor mutation and relative to one another in defining phage infectivity/resistance patterns. **Aims:** This project will focus on the opportunistic pathogen Pseudomonas aeruginosa, which causes lung infections in cystic fibrosis patients and immunocompromised people, as well as a range of other infections in healthy people. The Van Houte and Westra labs at Exeter have a collection of >2000 P. aeruginosa isolates and >150 P. aeruginosa-specific bacteriophages. The team has already set up a workflow for high-throughput phage infection experiments and the entire isolate collection is being used to generate large infection datasets that will feed into this PhD project. **Key aims:** 1. Identifying host receptors for specific phage via classical statistical approaches (Genome-Wide Association Studies, GWAS) applied to infection data (few phage, many bacteria) 2. Identifying anti-defence systems via GWAS (many phage, few bacteria) 3. Developing AI systems which incorporate both host and virus genomic data to predict susceptibility and use them to explore the biology of host-viral interactions. These aims combine well-understood classical statistical approaches (GWAS) which leverage the unique dataset, with cutting edge new artificial intelligence architectures, which provide a completely orthogonal approach to try to learn the properties of the data. This is a very interdisciplinary project, which would suit someone with data science strengths interested in using it to decipher new biology, or someone with evolutionary and microbiological experience who is keen to engage deeply with new statistical/AI methods.

Supervisory Team	
Lead Supervisor	
Name	Professor Zamin Iqbal
Affiliation	Bath
College/Faculty	Milner Centre for Evolution
Department/School	Department of Life Sciences
Email Address	zi245@bath.ac.uk
Co-Supervisor 1	
Name	Professor Stineke Van Houte

Affiliation	Exeter
College/Faculty	ESE
Department/School	CEC, ESI
Co-Supervisor 2	
Name	Professor Edze Westra
Affiliation	Exeter
College/Faculty	ESE
Department/School	CEC, ESI
Co-Supervisor 3	
Name	
Affiliation	
College/Faculty	
Department/School	