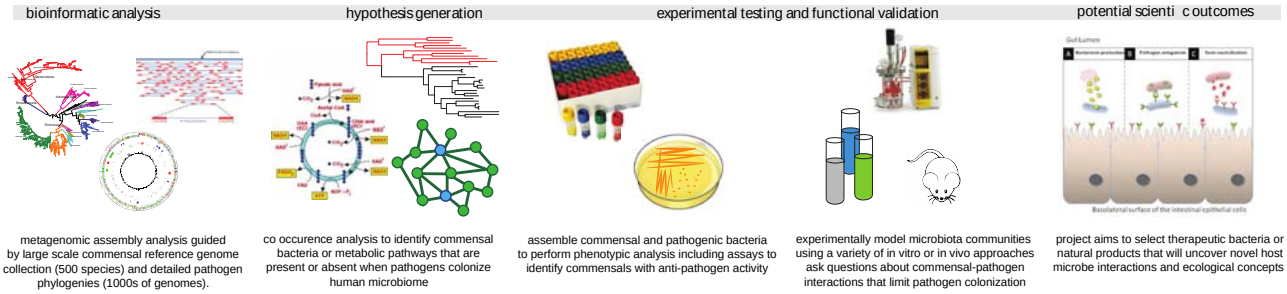
Rational Mining of the Human Microbiome for Novel Antibiotic Candidates that Target Antimicrobial Resistant Pathogens

Widespread antimicrobial resistance (AMR) is now rendering some bacterial infections untreatable with current antibiotics. The problem is dire, especially considering there are very few new or novel classes of antibiotics in the discovery pipeline. Many believe we are headed for a post-antibiotic era where common bacterial infections and routine clinical procedures become life or death scenarios. For example, colistin is a last resort antibiotic and there are now reports of the emergence of colistin resistant *Klebsiella pneumonia* meaning that lives are now at risk of infections once thought to be easily and routinely treatable. Novel approaches and technologies are required to find candidate antibiotics to replenish the discovery pipeline with products that spare the co-resident microbiota.

Over the past few years, extensive research into the human microbiota has led to insights on interactions between health-associated bacteria and enteric pathogens that could be exploited to discover novel metabolites or other bacterial by-products that could be novel classes of antibiotic candidates. We believe that during intestinal health, some commensal bacteria produce molecules that can interfere with or kill pathogens or hold a pathogen colonization in stasis, thereby preventing overgrowth. How do we accurately identify these pathogen restraining commensals?

Using metagenomic mining of the human microbiome we find that about 5-10% of healthy individuals carry low levels of known pathogens (*Klebsiella*, *E. coli*, *Enterococci*, *Camploybacter*, *Salmonella*, *C. perfringens* - unpublished data). However, nobody has yet systematically and thoroughly data mined the human microbiome for beneficial bacteria with anti-pathogen activity. Such surveys are hampered by the variability of the microbiota, the complexity of the system, heterogeneity of informatics analysis, and the broadly assumed viewed that the microbiota is "unculturable". The Lawley group recently overturned the latter dogma by developing novel methods to culture the majority of bacteria from the human intestinal tract and identifying many novel bacterial taxa (family, genera and species) that are ubiquitous to humans (Browne et al. 2016. Nature). Importantly, we now have over 700 reference genomes (500 species) generated from commensal bacteria providing a basis for fine grain taxonomic analysis to the subspecies/strain level (using multiple loci) and a framework for high throughput, precision assembly of genomes found in microbiota metagenomes (deep, short read datasets). We believe this approach opens the opportunity for functional studies of the human microbiome, particularly for identifying commensals with anti-pathogen activity.



Scienti c work ow of ESPOD project. Metagenomic analysis will be used to identify commensal bacteria or metabolic pathways in the human microbiome that can exclude pathogen colonization and then functional validated and mechansisms studied using in vitr o and in vivo methods.

The project will leverage the informatics expertise in the Finn group to initially perform large-scale metagenomic analysis using over 7,000 datasets with high quality and rich sample metadata. The data contains sufficient read depth to perform reference genome based mapping (using data from the Lawley group) and *de novo* assembly to accurately identify bacterial species and subspecies and associated abundances with a high degree of confidence. With this detailed listing, we will use a co-occurrence network analysis to provide insights into which commensal bacteria could exclude pathogens from the community. Detailed phylogenies for pathogens carried at low level within the

healthy human population will be analysed in the context of metadata including known pathogenic outcomes to differentiate the genetic vs environmental factors underpinning the pathogen inhibition factors of interest. This analysis will be based on 1000s of pathogen genomes to allow for precise identification. Based on the co-occurrence analysis we will identify species and/or specific metabolic pathways/gene clusters that could lead to the identification of novel commensal or antibiotic candidates.

SPECIFIC AIMS

One key strength of this project is the complement of a high-end bioinformatics group with a well-supported, state-of-the-art CL2 experimental laboratory that will allow the post-doctoral fellow to achieve the following specific aims:

1. To develop pipelines and perform comprehensive genomic and phylogenetic analysis of extensive pathogen genome collections to identify key marker genes capable of detecting and determining the lineage of AMR pathogens within metagenomics datasets.

2. To use commensal and pathogenic genome sequences and associated phylogenies for reference based analysis of WGS metagenomes to establish positive and negative co-occurrence correlations for AMR pathogens. Unmapped/unassembled sequences will undergo *de novo* assembly and binning to establish potential commensal bacteria absent from the reference genome collections that could be cultured. Metagenomes containing commensal organisms that have the strongest co-occurrence correlations (both positive and negative), hence likely to be harbouring antimicrobial resistance genes or antibiotic synthesis operons, will have metadata verified and enriched from associated publications. The Finn group will help use the reference genomes, genomes assembled from metagenomes (both referenced based and *de novo* assemblies) and sample metadata to mine for potential pathways (e.g. antibiotics, immunity proteins, metabolic pathways) to provide a rationale for the observed relationships.

3. To assemble a collection of pathogens and commensal organisms (based on the findings from Aims 1 and 2) to initially test commensals for anti-pathogen activity. The Lawley Lab maintains a large diverse culture collection of over 400 species with whole genomes and the capacity to target culture desired bacteria from microbiota samples. The lab has a variety of experimental platforms to investigate the molecular and genetic basis of host-microbiota interactions (Figure 1). Candidate bacteria will be taken forward for validation and mechanistic studies in the Lawley Lab with the options to use a variety of *in vitro* (e.g. microbiology, intestinal organoids) and *in vivo* methods (e.g. germ free mouse colonization) in combination with metatranscriptomic and metabolic analysis. The lab maintains a fermentation system to grow pure cultures or model communities, if warranted. Thus, the fellow will have a variety of experimental platforms to interrogate and validate hypotheses generated from Aims 1 and 2 – fostering a cycle of experimental validation follow by refinement of informatics analysis. The overarching goal is to determine the unpinning molecular mechanisms of candidate commensals/natural products, both in terms of the target in the pathogen and the spectrum of action against other pathogens and the microbiota.

EXPECTED OUTCOMES

This project offers a unique opportunity for a post-doctoral fellow in the burgeoning field of human microbiome research and asks deep questions about the co-evolution of human commensals and pathogens. The resulting human microbiome data generated will have broad applications for deciphering the mutualistic co-evolution of the human microbiome and understanding health and diseases of the intestine that are associated with the microbiota. Once developed, the pipelines and experimental work could be readily applied to other biomes (e.g. oral cavity or cow rumen) to deepen our understanding of the interplay of micro-organisms in different biomes. Finally, this collaboration between the EBI and WTSI will form the basis of a unique, world-class research programme, integrating state-of-the-art bioinformatic and experimental approaches.