PI profile

## Azim Ansari

|  |  |
| --- | --- |
|  | **Dr. Azim Ansari**  **Titles**: Sir Henry Dale Fellow, Group Leader in the Experimental Medicine Division, Associate Group Leader at the Wellcome Centre for Human Genetics  **Location**: Peter Medawar Building for Pathogen Research and Wellcome Centre for Human Genetics  **Department**: Nuffield Department of Medicine  **Group**: Ansari: Host Pathogen Genomics  **Webpage**: https://www.ndm.ox.ac.uk/team/azim-ansari  **Email**: azim.ansari@ndm.ox.ac.uk |

GMS themes:

* Genomic and –omic technologies
* Genome biology (genomes and genetic variation)
* Genomics of infectious disease
* Genomic analysis (bioinformatics and statistical genetics)
* Application of genomics in the clinic (diagnostics and therapeutics)

### Research Overview

We investigate infectious diseases and the role of host and pathogen genetic variation on the outcome of infection. Genome-wide association studies (GWAS) aim to identify the genetic basis of phenotypic traits using the variation that exists within natural populations. Uniquely for infectious diseases, the inter-individual heterogeneity in disease phenotype is linked to both host and pathogen genetic variation. Traditionally, genetic studies of infectious diseases have sought to explain between-individual variation in disease phenotypes by assessing genetic factors separately in humans or pathogens, under the assumption that these factors are independent. Although reasonable for some variants, there is strong theoretical and empirical evidence that genetic interactions between host and pathogen play a major role in infectious disease aetiology. We integrate host and pathogen genomic data from the same patients to better understand pathogenesis and between-individual heterogeneity in disease outcomes. Using paired host-pathogen data we aim to identify (a) host polymorphisms linked with pathogen sequence variation, (b) pathogen sites under strong host genetic selective pressures, (c) host and virus genetic factors independently contributing to disease phenotypes and (d) host-virus genetic interactions contributing to disease phenotypes. The findings have the potential to: (I) revolutionize our understanding of host-pathogen interactions and human biology; (II) aid in development of more effective vaccines, drug targets and immunotherapies; and (III) permit better use of therapies through patient stratification.

With the development of mobile and rapid sequencing platforms and exponential decrease in the cost of sequencing, the ability to generate pathogen and host data on the same subjects is becoming widely available. However, there is a lack of analysis tools and pipelines for pathogen GWAS, host-pathogen genome-to-genome analyses and to uncover interactions that modify disease phenotypes. We develop methodologies to integrate host and pathogen data with a focus on chronic infections of hepatitis B and C viruses (HBV and HCV) and human immunodeficiency virus (HIV) where we analyse data from large well characterised cohorts where we have generated paired host pathogen genomic data. In the age of “Big Data” and “Personalised Medicine”, analysis of paired host-pathogen genomic data will become increasingly important to uncover the mechanisms driving pathogen adaptations and heterogeneity of infection outcomes.

We are a cross‐disciplinary group, based both at the Peter Medawar Building for Pathogen Research and the Wellcome Centre for Human Genetics (WCHG). Our focus is data analysis and developing statistical and computational methodologies, although we do wet-lab work too. We have strong collaborations with clinicians, immunologists and wet-lab scientists nationally and internationally (Pakistan, Vietnam, China, South Africa, Uganda) and the students will have the opportunity to work with them. We are flexible about the projects, and they will be tailored around the interests of the students.

Project areas: Infectious diseases, Evolution, Bioinformatics, Host Pathogen interactions, Population genetics, Computational Statistics, Machine Learning

### Specific project proposals:

* Developing Statistical Methodology for integrated analysis of host-pathogen genomic data
* Developing Statistical Methodology to model mixed infection and its contribution to maintenance of pathogen genetic diversity
* Understanding mechanisms of sex disparities in infectious diseases
* The impact of host genetic variation on the within-individual pathogen diversity

We have many other projects in the group, please contact directly for further information on other possible projects on both data analysis and model development. We also will consider projects proposals from students.

*These pages were reviewed/updated:* ***27/September/2022***

Project proposal

# **Title**: Developing Statistical Methodology for integrated analysis of host-pathogen genomic data

Supervisors: Dr. Azim Ansari and Dr. Gavin Band

Wet/dry lab mix (approx): 0% wet lab, 100% dry lab

### Description:

To date, most human pathogen genetic interaction studies have focused on HLA-associated escape variants and the lack of complete host and pathogen genetic variation data has meant study of other selective effects has been neglected. Host-to-pathogen genome-to-genome analysis will detect host genetic “footprinting” on the pathogen in a non-biased fashion. At present, there is a paucity of such studies. Some of key challenges are as follows:

* Pathogen and host populations are co-structured. If not accounted for, this co-structuring will result in many false positive associations. There is a need to develop statistical methodologies and associated software to account for the strong co-structuring between the pathogen and the host.
* The current methods have low Statistical power as typically one would test for association between millions of host SNPs and thousands of pathogen SNPs. There is a need to develop methods that will increase the power of such studies. Two possible strategies are either marginalising over one set of variables or using omnibus tests that combine evidence across many hypotheses.
* A third area for methodological development relates to within-infection diversity. For instance, in the case of malaria this occurs when multiple mosquito bites generate overlapping infections (superinfection) or when mosquitos themselves carry several strains (coinfection). In the context of chronic infections (such as HCV) within-patient diversity occurs because mutations arise during the long infection periods, with the level of pathogen diversity typically varying between patients. These features are not well handled by current association testing methods, and we would like to develop new methods that account for this.

The methods will be applied to large HCV, HBV, HIV and malaria cohorts where we have generated both host and pathogen genomic data.

We can shape the short project around the student’s interest.

### Training Opportunities:

The student will develop expertise in Statistical genomics, Statistical Modelling, Machine Learning, bioinformatics, infectious diseases, evolution and population genetics. This studentship will be based at the Peter Medawar Building for Pathogen Research (PMB) and the Wellcome Centre for Human Genetics (WHG) at Oxford. The PMB houses around 150 scientists working on HIV, HCV, influenza, TB, malaria, SARS-CoV-2 and dengue and many of the PIs are global leaders in the study of infections.

### Background reading / references:

* Band, G, et al. “Malaria protection due to sickle haemoglobin depends on parasite genotype” *Nature* **602**, 106–111 (2022). https://doi.org/10.1038/s41586-021-04288-3
* Ansari, M. Azim, et al. "Genome-to-genome analysis highlights the effect of the human innate and adaptive immune systems on the hepatitis C virus." *Nature genetics* 49.5 (2017): 666-673.
* Behr, Merle, et al. "Testing for dependence on tree structures." *Proceedings of the National Academy of Sciences* 117.18 (2020): 9787-9792.
* Crawford, Lorin, et al. "Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits." *PLoS genetics* 13.7 (2017): e1006869.

Project proposal

# **Title**: Understanding mechanisms of sex disparities in infectious diseases

Supervisors: Dr. Azim Ansari and Dr. Gavin Band

Wet/dry lab mix (approx): 0% wet lab, 100% dry lab

### Description:

The mortality rate for COVID-19 pandemic has been two to three times higher in men than women. Similar observation extends to susceptibility and outcome of most other infectious diseases. For instance, after initial Hepatitis C Virus infection women are more likely to spontaneously clear the virus without any interventions. The rate of progression to severe liver disease is also slower in women than men and there is some evidence that women respond better to direct-acting antiviral therapies. On the other hand, many auto-immune diseases (such as lupus) have a much higher rate of occurrence in women than men. Despite large evidence for sex differences in autoimmune diseases and susceptibility and outcome of infectious diseases, data addressing the biological mechanism are remarkably scarce.

In this short project you will use computational and (and potentially experimental) methods to probe differences in immune system that lead to sex differences in infectious diseases. One hypothesis for the sex differences in immune responses are that in females, due to the random nature of X-chromosome inactivation process, organs are mosaic and consist of two populations of cells with preferential expression of either paternal or maternal X-chromosome. This increased level of genetic heterogeneity at organ level in females relative to males could explain, better immune responses against infections.

To test this hypothesis, you will use liver RNA-seq and genomic data from a cohort of 200 patients infected with HCV to estimate what proportion of cells express paternal and maternal X-chromosomes in each patient. We will then test for association between clinical phenotypes (viral load and treatment outcome) and the level of X-chromosome expression heterogeneity.

### Training Opportunities:

The student will develop expertise in Statistical genomics, Statistical Modelling, Machine Learning, bioinformatics, infectious diseases, evolution and population genetics. This studentship will be based at the Peter Medawar Building for Pathogen Research (PMB) and the Wellcome Centre for Human Genetics (WHG) at Oxford. The PMB houses around 150 scientists working on HIV, HCV, influenza, TB, malaria, SARS-CoV-2 and dengue and many of the PIs are global leaders in the study of infections.

### Background reading / references:

* Vieira VA, Zuidewind P, Muenchhoff M, Roider J, Millar J, Clapson M, Van Zyl A, Shingadia D, Adland E, Athavale R, Grayson N, Ansari MA, et al. Strong sex bias in elite control of paediatric HIV infection. AIDS (London, England). 2019 Jan 2;33(1):67.
* Shvetsova, E, et al. Skewed X-inactivation is common in the general female population, European Journal of Human Genetics, 2019.
* Tukiainen, T, et al. Landscape of X chromosome inactivation across human tissues, Nature, 2017.
* Oliva, Meritxell, et al. "The impact of sex on gene expression across human tissues." Science 369.6509 (2020).