



Title of PhD project	Developing machine learning tools to investigate evolutionary trajectories in emerging viral infectious diseases	
Supervisor	Dr Nick Furnham	LSHTM
Co-Supervisor	Dr David Allen	LSHTM
Brief description of project	<p>Understanding the evolutionary trajectory of an emerging viral pathogen is crucial to both surveillance and the development of therapeutic interventions. This project aims to advance our understanding of the mutational pathways followed as a virus adapts to its host. The unprecedented and large quantity of genomic and molecular data linked to phenotypic health data that has been generated in response to the current Sars-CoV-2 pandemic, as well as existing data from influenza and other viruses, allows us to address this question at the molecular level. Using a combination of computational techniques, including structural bioinformatics and machine learning / deep learning, the effects of successive combinations of mutations in proteins involved in viral/host interactions will be evaluated and predicted. To test and appraise the computation tools a range of lab based molecular virology techniques will be employed. Including in vitro virus culture, cloning, site-directed mutagenesis, reverse genetics, recombinant protein expression and purification, pseudovirus construction, classical and pseudoneutralisation assays. The computational tools and insights from this project can be applied to future emerging viral pathogens. The range of highly sought after skills developed within the project will give a firm bases for a future career in academia and/or industry.</p>	
Skills we expect a student to develop/acquire whilst pursuing this project	<p>Advanced skills in structural bioinformatics and computational biology including genome analysis and machine learning / deep learning approaches. Experimental molecular virology techniques including <i>in vitro</i> virus culture, cloning, site-directed mutagenesis, reverse genetics, recombinant protein expression and purification, pseudovirus construction, classical and pseudoneutralisation assays.</p>	
Particular <u>prior</u> educational requirements for a student undertaking this project	<p>As a significant proportion of the project is computational, a strong interest in bioinformatics techniques and some programming experience would be helpful. A background in</p>	

	the lab based molecular biosciences would be beneficial, especially if that included molecular or cellular virology, though not essential. Appropriate training will be made available to fill gaps in knowledge/experience.
Project key words	Molecular virology Sars-CoV-2 Structural bioinformatics Machine learning Host-pathogen interactions Emerging infections
Possible under 1+4 route? Master's options identified.	Yes LSHTM – MSc Medical Microbiology
MRC Core Skills developed through this project	Quantitative skills
MRC LID themes	Health Data Science Infectious Disease
Further reading	<p>Epistatic interactions can moderate the antigenic effect of substitutions in haemagglutinin of influenza H3N2 virus</p> <p>Effects of common mutations in the SARS-CoV-2 Spike RBD and its ligand, the human ACE2 receptor on binding affinity and kinetics</p> <p>SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals</p> <p>Antigenic cartography reveals complexities of genetic determinants that lead to antigenic differences among pandemic GII.4 noroviruses</p> <p>Positive epistasis between viral polymerase and the 3' untranslated region of its genome reveals the epidemiologic fitness of dengue virus</p>