Fourth year projects in Bioinformatics offered by Cathal Seoighe

Prediction of phenotype from genotype data

The prediction of diseases and physical characteristics from whole genome genotype data is an area of great interest in the context of recent advances towards personalized genomics and personalized medicine, with medical as well as forensic applications. Some characteristics can now be predicted with high accuracy, including pigmentation (hair, eye and skin colour) and these predictions can be applied to DNA obtained, for example, from crime scenes to infer physical characteristics of suspects. Pigmentation characteristics are relatively easy to predict because they involve a relatively small number of different genes. Other characteristics, such as height, are polygenic and extremely difficult, if not impossible (at least currently) to predict. The objective of this project will be to investigate methods that have been developed to predict phenotype from genotype data and to apply these to any of a number of genome-wide genotype datasets in the public domain.

Identification of epitopes in viruses using Markov Chain Monte Carlo methods

This project is part of a large international collaboration to identify parts of the HIV virus that are targeted by the immune system by applying models of how HIV evolves over time. The pattern of evolution can give clues to where these epitopes occur and this can provide important information for efforts to design vaccines. The evolutionary models have already been developed in the research group. The role of this project will be to reimplement part of the computer code that applies the Markov Chain Monte Carlo resampling method to come up with a statistical inference about the location of epitopes on the surface of three-dimensional HIV proteins and to investigate whether a similar approach can be used to find immunogenic regions on the surface of influenza viruses.