Project title: Clustering of admixed genomes using closely linked markers
Supervisor: Dr. Michael Salter-Townshend

Summary: Population genetics is concerned with modelling the shared variation in heterogeneous populations and can shed light on human demographic history, natural selection, and disease association. Admixture is of particular interest and is when individuals from previously separated groups come together and produce offspring. The mechanism of recombination leaves tell-tale patterns in descendants' genomes that allow us to date and otherwise quantify the admixture event. Popular models for clustering such individuals to sub-populations ignore the correlation of mutations along the genome, or rely on having good surrogates for the mixing groups that are not themselves admixed, or do not yield accurate estimates of ancestry along the genome. You will build a model and inferential method that addresses these concerns, performing accurate clustering of chunks of DNA of admixed individuals, using only these same individuals as input. You will need a background in mathematics or statistics, an interest in genetics, and knowledge of Bayesian inference. An ability to code in R or C is also required.