

## Combining different genetic data types and additional ecological covariates in population genetic analyses

Biologists use both nuclear DNA (inherited from both parents) and mitochondrial DNA (inherited maternally) to determine which individuals of a given species cluster into distinct populations with separate gene pools. They also collect increasing amounts of auxiliary data such as dietary isotopes or behavioural and environmental covariates. The information is used to better manage those populations. Although many genetic clustering methods exist, none of them can combine nuclear and mitochondrial DNA data, and few can incorporate non-genetic data so biologists are currently forced to analyse the various data types separately and manually piece together the conclusions.

We are working on a method to cluster combinations of nuclear and mitochondrial DNA data, along with additional covariates that may be categorical or continuous. The first step is to modify the widely-used STRUCTURE software to accommodate both nuclear and mitochondrial data in the same analysis. The second step, incorporating non-genetic covariates, will produce a novel clustering algorithm using a technique known as 'surrogate residuals', which are akin to randomised quantile residuals (or PIT residuals). Essentially, we will convert any categorical variables into continuous values that adhere to the ordinal structure of the original data, and then will cluster those converted categorical variables along with any data that are already continuous. This approach will be fast, and scalable to thousands of SNPs.

The new method will be particularly valuable for migratory species such as whales, for which mitochondrial DNA patterns can indicate maternally-mediated elements of migratory behaviours. We will focus on data from southern right whales and humpback whales, but these methods will be applicable to hundreds of other animal species.

In the presentation we will discuss the workings of the new algorithm, and illustrate it with an ecological case study.