

Bayesian inference of past population dynamics from genomic data

Professor Xavier Didelot and Professor Matt Keeling

Background

Variations in the size of a population have an effect on the patterns of relatedness between individuals in the population. For example, if the population size was constant, the genealogy of sampled individuals is expected to follow a simple model called the coalescent. On the other hand, if the population grew exponentially, then the genealogy will have longer terminal branches and shorter internal branches than in the coalescent model. This basic principle can be reverted to learn about past population dynamics from recently sampled genomic data by first reconstructing a genealogical tree and then interpreting the tree in terms of temporal trends in the population size (or more exactly by performing these two steps jointly rather than successively). These trends could be assumed to follow a specific model (for example constant or exponentially growing population size as mentioned above), although a nonparametric approach is often used for more flexibility.

Objectives

The aim of this project is to develop, implement and test a new statistical methodology to perform Bayesian inference of past population dynamics given a dated genealogical tree.

Proposed research plan

1. Get familiar with coalescent theory (no prior knowledge on this is required)
2. Simulate genealogies under a constant population size model
3. Simulate genealogies under a variable population size model
4. Formulate a flexible model under which to perform inference and establish a strategy to calculate the likelihood
5. Implement a method to perform inference (probably Markov Chain Monte Carlo)
6. Test the method thoroughly using simulated data and real data if time allows

Starter References

1. Lan S, Palacios JA, Karcher M, Minin VN, Shahbaba B. An efficient Bayesian inference framework for coalescent-based nonparametric phylodynamics. *Bioinformatics*. 2015;31: 3282–3289.
2. Karcher MD, Palacios JA, Lan S, Minin VN. *phylodyn*: an R package for phylodynamic simulation and inference. *Mol Ecol Resour*. 2017;17: 96–100.
3. Volz EM, Didelot X. Modeling the growth and decline of pathogen effective population size provides insight into epidemic dynamics and drivers of antimicrobial resistance. *Syst Biol*. 2018;67: 719–728.