Bayesian Inference of Evolutionary Histories from Genomic Data.

The evolutionary history of a group of organisms explains the process of their genetic variation over time. Due to recent sequencing and computing advances, statistical inference has become an essential discipline in the study of evolutionary history from genomic data. However, typical analyses are either limited to a small amount of data or fail to estimate complex and diverse evolutionary models. In this talk, I will present a new Bayesian method for estimating population/species-level history including population sizes, splitting time of two populations, and migration rates. The new method resolves statistical limitations and overcome major roadblocks to analyze genome-scale data. Using importance sampling and a Markov chain representation of genealogy, the new method scales to genomic data without mixing difficulty in a Markov chain Monte Carlo simulation. The method also provides for the calculation of the joint posterior density for all model parameters, thus resolving the problem of high false positive rates that arises for the likelihood ratio tests for migration rates using other existing Bayesian approaches. I will demonstrate the new method with simulated data and real DNA sequences.