

**University of California, Irvine
Statistics Seminar**

***Towards Bayesian Phylogenetics via Systematic
Search and Gradient Ascent***

**Frederick "Erick" Matsen IV
Professor
Fred Hutchinson Cancer Research Center**

4 p.m., Thursday, January 20, 2022

Phylogenetic (evolutionary tree) inference is a key tool for understanding evolutionary systems. This includes viral adaptation and genomic epidemiology, as well as the antibody response to infection and vaccination. Bayesian phylogenetic analysis allows us to assess and integrate out tree uncertainty to obtain more reliable estimates of other model variables of interest (e.g. transmission rates). However, Bayesian posterior distributions on phylogenetic trees remain difficult to sample despite decades of effort. The complex discrete and continuous model structure of trees means that recent inferential methods developed for Euclidean space are not easily applicable to the phylogenetic case. Thus, we are left with random-walk Markov Chain Monte Carlo (MCMC) with uninformed tree modification proposals; these traverse tree space slowly because phylogenetic posteriors are concentrated on a small fraction of the very many possible trees. In this talk, I will describe our work to design new scalable approaches to inferring the Bayesian posterior on phylogenetic trees. This includes establishing a new discrete inferential target, which we call the "subsplit directed acyclic graph," and a new algorithm that will allow us to infer this structure using methods analogous to much faster maximum-likelihood (point-estimate) methods for phylogenetics. I will also describe how, once this structure is in hand, we can perform variational inference for continuous parameters via stochastic gradient descent.