Communication libre
« Théorie de la coalescence : nouveaux développements et applications »
7–11 octobre 2013

Contributed talk
"Coalescent Theory: New Developments and Applications"
October 7–11, 2013

On the accuracy of deterministic approximations to coalescent formulas

Ethan Jewett*

emjewett@stanford.ed

Many coalescent distributions and expectations can be derived by conditioning on the number, n_t, of alleles at time t in the past that are ancestral to a data set of n_0 alleles sampled in the present. However, summing over the conditional distribution of n_t, given n_0, can be computationally challenging when n_0 is large. Therefore, such formulas can be difficult to evaluate on modern genomic datasets with hundreds or thousands of sampled lineages. One alternative to conditioning on all possible values of n_t is to use an approximation in which n_t is assumed to equal its expected value $\mathrm{E}[n_t]$ with probability one (Slatkin, 2000). This approximation greatly reduces the number of terms in conditional expressions, significantly reducing their computational complexity. However, despite the utility of the approximation, its theoretical accuracy is not known. Instead, the accuracy of any given version of the approximation must be evaluated empirically by comparing it with the true distribution or with simulated values. Here, we show that approximate distributions converge uniformly to the true distributions under certain simple assumptions, and we derive an expression for the asymptotic approximation error. We also obtain approximations of $\mathrm{E}[n_t]$ for the case of multiple populations of time-varying size with migration among them. Our results provide a theoretical basis for understanding the ranges of parameter values over which any given approximation is accurate, facilitating the application of the approximation $n_t = \mathrm{E}[n_t]$ to reduce the complexity of computing coalescent formulas on large genomic data sets.

Joint work with Noah A. Rosenberg

*Department of Biology, Stanford University, 371 Serra Mall, Stanford, CA 94305-5020, USA.