Do Neighbor-joining and Maximum Likelihood Methods Produce similar Bootstrap Consensus Tree?

ABSTRACT

Neighbor Joining (NJ) and Maximum Likelihood (ML) are two major phylogenetic tree reconstruction methods. NJ is very computing efficient and simulation studies show high accuracy for NJ. ML has nice statistic properties but is very time consuming. Bootstrapping is a statistic procedure that is applied in the phylogenetic context to test the reliabilities (how much confidence we should put in some monophyletic group defined by the estimated tree) of the tree. Bootstrapping for NJ can be done within several minutes even the number of taxa is large. But bootstrapping for ML is not feasible due to large amount of computing time. So it is interesting to know if NJ and ML produce different consensus trees or not. This project is proposed and designed to address this problem by Dr. Kumar. I'm helping on the data collection, programming and analyzing the results. My focus was on understanding the theoretical, computational, and practical aspects of NJ, ML and Bootstrapping method in reconstructing and testing phylogenetic trees using molecular data. The study of theoretical aspects involved a primary literature survey and the aspect of experimental design, computational aspects dealt with understanding the inner-workings of the PAUP software, and practical work involved writing PAUP scripts to run ML, NJ and Bootstrapping, PERL scripts to extract patterns from the bootstrap bipartition table, and statistical approaches to do data analyses.