Inferring tipdated phylogenies from distances

Since Zuckerkandl and Pauling proposed the molecular clock in the 1960s it became common to trace mutations in genomes to infer their evolutionary history. For fast evolving organisms or viruses we can estimate mutation rates and divergence times without the need of fossil calibration. We only need molecular sequences and their associated sampling times. Thanks to steady improvements in sequencing technologies thousands and in the case of SARS-CoV-2 millions of sequences are available. Bayesian MCMC algorithms are the gold standard for inferring divergence times and mutation rates, but are computationally costly. Here we present a fast algorithm to estimate tip-dated phylogenies using a strict molecular clock with (non-negative) least squares. We will further discuss the relationship to other distance based algorithms and possible extensions like relaxed clock models. An implementation of the algorithm will be part of the R package phangorn.