Inference of fitness from genealogical trees

MARIJA VUCELJA, Courant Institute of Mathematical Sciences, NYU, ADEL DAYARIAN, BORIS SHRAIMAN, Kavli Institute of Theoretical Physics, UCSB — Natural populations are fitness diverse and can have numerous genes under selection. The genealogical trees, that one obtains by sampling, often bear hallmarks of selection, such multiple mergers, asymmetric tree branches and long terminal branches (the trees are squished towards the root). These are qualitative differences compared to trees in the absence of selection. We propose a theoretical model that links the morphology of a tree with the fitness of the leaves. We obtain multipoint correlation functions of the fitness along the tree. In this way we are able extract some quantitative information about the strength of selection from data-reconstructed trees. The extensions of this approach can potentially be useful for inferring relative fitness of sequenced genomes of tumors and for predicting viral outbreaks.

Marija Vucelja
Courant Institute of Mathematical Sciences, NYU

Date submitted: 05 Nov 2012
Electronic form version 1.4