

Abstract Submitted
for the MAR10 Meeting of
The American Physical Society

A Bayesian Approach to Detecting Amino Acid Covariance in Multiple Sequence Alignments LUCY COLWELL, MICHAEL BRENNER, ANDREW MURRAY, Harvard University — Determining which residues of a protein control its biological function is a classical question in molecular biology. In particular, proteins can change their structure or function by mutating just a small set of residues. An attractive idea is that distinct sets of residues are responsible for different phenotypic properties, so that one property can be changed while another is not. Members of such a set mutate at similar points in a multiple sequence alignment and so are correlated. It has long been proposed that analysis of correlations in the mutation patterns of protein sequences may provide an important means of extracting functional information about proteins from sequence alignments. Here, we propose a methodology for incorporating functional and structural annotations of the sequences analyzed to improve the efficacy of algorithms at detecting such residue sets. We provide a Bayesian framework in which known biological properties of the sequences are used to define a prior probability that quantifies our belief that sequence positions with different conservation levels are associated with the phenotype of interest. Recent experimental data is used to demonstrate that applying these principles results in improved detection ability, allowing us to distinguish between pairs that demonstrate similar levels of correlation but are not of equal relevance to the phenotypic purpose being addressed.

Michael Brenner
Harvard University

Date submitted: 19 Nov 2009

Electronic form version 1.4