Atelier
« Les mathématiques de l'évolution des séquences : modèles biologiques et applications »
23-27 septembre 2013

Workshop
"Mathematics of Sequence Evolution: Biological Models and Applications"
September 23–27, 2013

Phylogenetic Gaussian process model for the inference of functional protein regions

Yifei Huang*

huangy67@mcmaster.ca

A variety of phylogenetic models have been developed to identify slowly evolved functional sites. Most of the existing software, e.g., Rate4Site, assumes that site-specific substitution rates are independent across sites. However, site-specific substitution rates may be strongly correlated in the protein tertiary structure. We developed a new program, GP4Rate, which incorporates Gaussian processes and phylogenetic model to identify slowly evolved regions in protein tertiary structures. Simulations and real data analyses suggest that GP4Rate outperforms Rate4Site.

*Department of Biology, McMaster University, 1280 Main St. West, Hamilton, ON L8S 4K1, CANADA.