Maximum Likelihood Models and Algorithms for Gene Duplications

Pawel Gorecki
Faculty of Mathematics, Informatics & Mechanics
University of Warsaw, Poland

Thursday September 26, 2013
3:40pm, B29 Atanasoff Hall

The abundance of new genomic data provides the opportunity to map the location of gene duplication and loss events on a species phylogeny. The first methods for mapping gene duplications and losses were based on a parsimony criterion, finding the mapping that minimizes the number of duplication and loss events. Probabilistic modeling of gene duplication and loss is relatively new and has largely focused on birth-death processes. During this presentation I will present a maximum likelihood model that estimates the speciation and gene duplication and loss events in a gene tree within a species tree with branch lengths. Several examples will be shown how to use this model to find optimal gene duplication and loss scenarios, even when the gene trees contain sequences from several hundred species. In many cases, these optimal scenarios differ from the lca-mapping that results from a parsimony gene tree reconciliation.

Pawel Gorecki is an Assistant Professor at the Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland. He received his scientific degrees from the University of Warsaw: MS in computer science (1999), MS in mathematics (2000) and PhD in computer science (2006). He was a postdoctoral fellow at the Max Planck Institute for Molecular Genetics in Berlin. His research is focused on the algorithmic and theoretical properties of problems and models in computational biology.

http://www.cs.iastate.edu/~colloq/new/

ALL SCIENCE IS COMPUTER SCIENCE