An Efficient Method for Inferring Tandem Duplication History

Louxin Zhang
National University of Singapore, S'pore
matzlx@nus.edu.sg

Abstract

Genome analysis suggests that tandem duplication is a primary mechanism for generating gene family clusters. Inferring the tandem duplication history of a gene family could yield valuable insight into functions of the family members. In this paper, we study this inferring problem using the duplication model proposed by Tang *et al.* We provide an efficient algorithm for inferring a duplication model for a given set of (gene) sequences by combining a linear-time algorithm, which is for determining whether a rooted tree is associated with a duplication model, with the nearest neighbor interchange operation. Finally, we derive duplication hypotheses for an exon of a mucin gene MUC5B, a ZNF gene family, and a OR gene family using our proposed method.

This is a joint work with B. Ma and L. Wang.