Combinatorial Methods for Genome Comparison

Anne Bergeron

Département d'informatique
Université du Québec à Montréal
C.P. 8888, succ. Centre-ville
Montréal, Québec
CANADA H3C 3P8

Abstract

Mechanisms of genome evolution are numerous. Most mutations are local, involving only a small portion of the DNA molecules. However, other types of mutations, much less frequent, are necessary to explain the diversity of existing species. These large scale mutations, called gene rearrangements, include reversals, transpositions, duplications, translocations, fusions and fissions. Each of these rearrangement admits a simple definition in terms of sequences of integers that represent the sequence of genes on a DNA molecule.

Rearrangements are studied with two main approaches. The operational approach explore the sequences of rearrangement operations that transform one genome into another. For a given set of operations, the minimal length of these sequence is a distance. Distances obtained in comparing several pairs of genomes can then be used to reconstruct phylogenies, or ancestral genomes.

A second approach tries to identify invariants in a set of genomes. These can be groups of genes that occur simultaneously in several genomes, with possible variations in gene order, or even gene content. In this case, comparing genomes focuses on similarity rather that distance.