

Abstracts

James Cai, Genome Research Center, HKU

Probabilistic Models of DNA Sequence Evolution with Context-dependent Mutation

Context-dependent mutation processes manifests specific patterns of dinucleotide frequencies in the genomes of most organisms. The CpG-methylationdeamination process, for example, is one of the prominent mutation processes in vertebrates (CpG effect). Context-dependent mutation processes need to be incorporated in order to make more realistic models of DNA substitutions. Based on a general framework of nucleotide substitutions, we developed a method to identify the most relevant dinucleotide substitutions and to estimate their relative frequencies. Our method starts from a model of context-independent nucleotide substitution, then context-dependent substitution parameters are successively added to increase the likelihood of the model in describing given data. We applied the method on the dataset from major eukaryotic lineages and revealed contrasting patterns in dinucleotide substitution.