Constructing phylogenetic relationship based on the independent selection law of genome sequences

Xiaolong Li

Inner Mongolia University <https://orcid.org/0000-0003-3637-295X>

Hong Li (✉ ndlihong@imu.edu.cn)

Inner Mongolia University School of Physical Science and Technology <https://orcid.org/0000-0001-6819-3038>

Zhenhua Yang

Inner Mongolia University School of Physical Science and Technology

Zefeng Zhang

Inner Mongolia University School of Physical Science and Technology

Research article

Keywords: Genome sequence, Phylogenetic relationship, 8-mer spectra, Independent selection law, Feature set

Posted Date: April 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-458329/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background: Exploring the composition and evolution regularity of genome sequences and constructing phylogenetic relationship by alignment-free method in genome level are high-profile topics. Our previous researches discovered the CG and TA independent selection laws existed in genome sequences by analysis on the spectral features of 8-mer subsets of 920 eukaryote and prokaryote genomes. We found that the evolution state of genomes is determined by the intensity of the two independent selections and the degree of the mutual inhibition between them.

Results: In this study, the two independent selection patterns of 22 primate and 28 insect genome sequences were analyzed further. The two complete 8-mer motif sets containing CG or TA dinucleotide and their feature of relative frequency are proposed. We found that the two 8-mer sets and their feature are related directly to sequence evolution of genomes. According to the relative frequency of two 8-mer sets, phylogenetic trees were constructed respectively for the given primate and insect genomes. Through analysis and comparison, we found that our phylogenetic trees are more consistent with the known conclusions.

Conclusions: The two kinds of phylogenetic relationships constructed by CG 8-mer set and TA 8-mer set are similar in insect genomes, but the phylogenetic relationship constructed by CG 8-mer set reflect the evolution state of genomes in current age and phylogenetic relationship constructed by TA 8-mer set reflect the evolution state of genomes in a slight earlier period. We thought it is the result that the TA independent selection is repressed by the CG independent selection in the process of genome evolution. Our study provides a theoretical approach to construct more objective evolution relationships in genome level.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures

Figure 1

The separability distributions (left) and the standard deviations of the separability and conservatism values in primate and insect genome groups (right). (A) The separability distributions of CG0 and TA0 8-mer spectra. (B) The separability distributions of CG1 and TA1 8-mer spectra. (C) The separability distributions of CG2 and TA2 8-mer spectra. (D) The standard deviations of the separability and conservatism values for three CG and three TA 8-mer spectra in primate group. (E) The standard deviations of the separability and conservatism values for three CG and three TA 8-mer spectra in insect group.

Figure 2

Phylogenetic trees of primates. (A) Total 8-mers. (B) 8-mers in CG1 subset. (C) 8-mers in CG2 subset. (D) 8-mers in CG1+CG2 subsets. (E) Reference phylogenetic tree.

Figure 3

Phylogenetic tree of insecta. (A) total 8-mers. (B) 8-mers in CG1+CG2 subset. (C) 8-mers in TA1+TA2 subset.

Figure 4

8-mer spectra of Pan troglodytes genome sequence. The vertical bar is the average frequency of total 8-mers. (A) Spectra of three CG 8-mer subsets and total 8-mers. (B) Spectra of three GC 8-mer subsets and total 8-mers.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalFile.pdf](#)