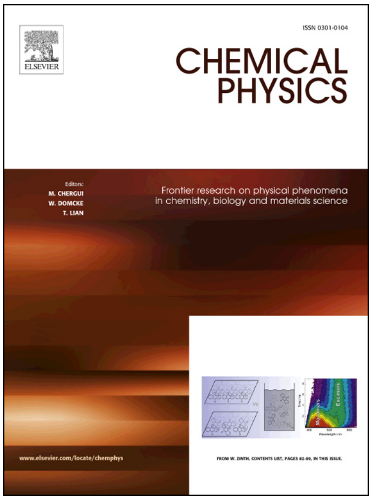
Accepted Manuscript



Graphical representation methods: How well do they discriminate between homologous gene sequences?

Dwaipayan Sen, Proyasha Roy, Ashesh Nandy, Subhash C Basak, Sukhen Das

PII: S0301-0104(18)30565-2
DOI: https://doi.org/10.1016/j.chemphys.2018.07.031
Reference: CHEMPH 10095

To appear in: *Chemical Physics*

Received Date: 25 May 2018
Revised Date: 18 July 2018
Accepted Date: 21 July 2018

Please cite this article as: D. Sen, P. Roy, A. Nandy, S.C. Basak, S. Das, Graphical representation methods: How well do they discriminate between homologous gene sequences?, *Chemical Physics* (2018), doi: https://doi.org/10.1016/j.chemphys.2018.07.031

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Graphical representation methods: How well do they discriminate between homologous gene sequences?

Dwaipayan Sen[1], Proyasha Roy[1], Ashesh Nandy[1*], Subhash C Basak[2], Sukhen Das[3]

[1] Centre for Interdisciplinary Research and Education, Jodhpur Park, Kolkata 700068, India
[2] Department of Chemistry & Biochemistry, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811, USA
[3] Physics Department, Jadavpur University, Jadavpur, Kolkata 700032. India

*Address for Correspondence: A. Nandy, Centre for Interdisciplinary Research and Education, 404B Jodhpur Park, Kolkata 700068, India. Email - anandy43@yahoo.com. Phone +91-9433579452

Abstract

Graphical representation methods constitute a class of alignment-free techniques for comparative study of bio-molecular sequences. In this brief commentary we study how well some of these methods can discriminate among closely related genes.

Keywords

Graphical representation, discriminatory power of sequence descriptors, super-descriptors, 2D and 3D methods, globin genes, Flavivirus envelope genes

1. Introduction

Graphical representations of biomolecular sequences have generated a lot of interest as a tool for alignment-free analysis, evidenced by the large number of research work on the subject. A principle application of these techniques has been in determining the evolutionary relationships analysed between different gene families [1]. The results obtained from the different graphical methods show small differences among them and display conformity with standard phylogenetic studies.

Graphical representations often provide a visual clue to the pattern of distribution of bases along DNA or RNA sequences. The representations are slightly more complicated in the case of protein sequences where one has to contend with 20 basic units, the amino acids, but ingenious schemes from 2D to 20D abstract graphs have been utilised to represent them too [2]. The graphical representation methods remain, to date, among the best to ‘see’ the base distribution in a DNA or an RNA sequence and follow the variations in a family of genes. This has enabled many applications of the techniques, for example, generation of new plant varieties [3], determination of origins of the SARS-coronavirus [4], identification of conserved regions in influenza virus neuraminidase gene for vaccine design and development [5].

Applications of these methods to viral sequence analyses have been a prime area of interest in view of the need for rapid development of drugs and vaccines against viral diseases. In a number of studies, our group has advocated rational designs of peptide vaccines against

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/7837087

Download Persian Version:

https://daneshyari.com/article/7837087

Daneshyari.com