Volume 5 - Issue 03 - March,2025

SABUJEEMA

An International Multidisciplinary e-Magazine

Article ID: SIMM0521

Popular Article

ANALYSIS OF BIOINFORMATIC TOOLS

SWORAJ KUMAR NAYAK

2nd Yr Msc (Ag), Department of Agricultural Biotechnology,
College of Agriculture, OUAT, Bhubaneswar

Open Access

Abstract

In molecular biology, bioinformatics uses computers, statistics, and mathematics to record, collect, evaluate, and interpret biological data. Despite still being in its infancy, it has emerged as one of the fastest-growing fields and has swiftly established itself as a crucial part of any biological research endeavor. It is gaining popularity because of its rapid and economical analysis of vast amounts of biological data. A biologist can use a variety of web- and/or computer-based technologies, most of which are freely available, to help them extract useful information from biological data. A thorough overview of some of the tools that a life scientist can use to analyse biological data is provided in this article.

KEYWORDS: - Phylogeny, Bioinformatics, Protein structure, Database

INTRODUCTION

The interdisciplinary field of bioinformatics was created by combining several different fields, including biology, mathematics, computer science, and statistics, to create storage techniques, biological data retrieval and analysis [Mount DW et al., 2004]. Gene characterisation, protein structure and physiochemical property determination, phylogenetic analysis, and simulations to examine the interactions of biomolecules in living cells are all common uses for computational techniques.

EVOLUTIONARY ANALYSIS

Phylogenetic analyses are methods for reconstructing the evolutionary relationship between a group of related molecules or organisms, predicting specific features of a molecule with unknown functions, tracking gene flow, and determining genetic relatedness. Phylogenetic tools are frequently used to test various evolutionary hypotheses and have become indispensable for functional genomics, especially when a gene's functions are unknown. The fundamental idea behind phylogeny is to arrange living things in groups based on how similar they are to one another; the more related the species, the closer they would appear on a tree. To account for the absence of statistical independence between species, phylogenetic comparative analysis is frequently employed [Freckleton RP et al., 2002].

DATABASE SEQUENCING

A biological sequence database is a sizable collection of data about biological molecules, including proteins, polymers, and nucleic acids, each of which is uniquely recognized by a key. The information that has been preserved is useful for primary sequence analyses in addition to being significant for future use. As high throughput sequencing technology have advanced, full genome sequencing has been achieved, producing enormous amounts of data daily. Databases are of primary and secondary type. Data from several primary sources is included in a composite database.

PROTEIN STRUCTURE IDENTIFICATION

Protein molecules start out as shapeless strings of amino acids that eventually fold up into a three-dimensional (3D) structure to become biologically active. Since any protein must fold into the correct topology in order to perform its biological functions, knowledge of a protein's 3D structure is essential to understanding that protein's function. Bioinformatics techniques can readily identify secondary structure elements in a protein sequence, such as helices, sheets, domains, strands, and coils. Proteins adopt a specific structure because of weaker electrostatic forces, such as hydrogen bonds, between these elements. The most popular technique for forecasting the target protein's template-based structure is homology modeling. However, this is hampered by the very small amount of structures available in PDB [Yao L et al., 2010].

INTERACTION OF MOLEULES

Proteins seldom carry out their tasks alone; instead, they constantly interact with other molecules to carry out specific tasks. Recognizing the interactions between biomolecules and other molecules has several ramifications, such as for drug design, protein folding, and purification methods [Wang L et al., 2010], and as a result, it has emerged as a highly sought-after study field employing experimental or bioinformatics methodologies. Interface size, amino acid composition at the interface, chemical group types, surface complementarity, hydrophobicity, hydrogen bonding, and conformational changes during complex formation are the factors that control protein-protein interactions.

Several protein datasets are used to study these characteristics. There are two categories of in-silico methods for studying molecular interactions: homology-based and non-homology-based. As the name suggests, the homology-based approaches compare directly sequences of proteins. Collectively, the non-homology techniques consider functional relationships. Non-homology-based approaches are effective in assigning functions to genes whose homologues have not yet been identified, even if homology-based approaches are still the most popular approach.

CONCLUSION

The field of bioinformatics is relatively new, although it has advanced quickly in recent years. It has enabled us to virtually test our

theories, which enables us to take a more accurate and an well-informed choice before to beginning expensive experiments. Despite the fact that an increasing number of tools are being developed for molecular simulations, genome and proteome analysis, structure prediction, rational drug design, and more, none of them are "perfect." As a result, the search for a better package to address the issues at hand will continue. There is no doubt that the availability of databases, which may be generic or customized, will play a major role in directing future research.

REFERENCES

Mount DW (2004) Sequence and genome analysis. New York: Cold Spring.

Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. Am Nat 160: 712-726.

Yao L, Evans JA, Rzhetsky A (2010) Novel opportunities for computational biology and sociology in drug discovery. Trends Biotechnol 28: 161-170

Wang L, Huang C, Yang MQ, Yang JY (2010) BindN+ for accurate prediction of DNA and RNA-binding residues from protein sequence features. BMC Syst Biol 4 Suppl 1: S3.