Chapter 11
An Improved Hybridized Evolutionary Algorithm Based on Rules for Local Sequence Alignment

Jayapriya J.
National Institute of Technology, India

Michael Arock
National Institute of Technology, India

ABSTRACT

In bioinformatics, sequence alignment is the heart of the sequence analysis. Sequence can be aligned locally or globally depending upon the biologist's need for the analysis. As local sequence alignment is considered important, there is demand for an efficient algorithm. Due to the enormous sequences in the biological database, there is a trade-off between computational time and accuracy. In general, all biological problems are considered as computational intensive problems. To solve these kinds of problems, evolutionary-based algorithms are proficiently used. This chapter focuses local alignment in molecular sequences and proposes an improvised hybrid evolutionary algorithm using particle swarm optimization and cellular automata (IPSOCA). The efficiency of the proposed algorithm is proved using the experimental analysis for benchmark dataset BaliBase and compared with other state-of-the-art techniques. Using the Wilcoxon matched pair signed rank test, the significance of the proposed algorithm is explicated.

INTRODUCTION

The Ultimate goal of Bioinformatics is to better understand the functionality of living cells at the molecular level. They are basically three analysis in molecular level namely sequence, structural and functional analysis. Amongst them sequence analysis is considered as the important one as it paves way for other two analysis. The first step in this is to align the sequences locally or globally where both are regard as important for biologist. In analysis, sequences are represented as the combination of alphabet (Xiong,

2006). Figure 1 shows the overview of bioinformatics domain. This figure explains the three analysis and its applications. Apart from these applications there are many in sequence analysis. As sequence analysis is the first phase of the investigation, many researches are developed in this area.

Sequence alignment is a process in which sequences are arranged in such a way that similar residues are found in the same column. Initially, two sequences are aligned to find the similarity between them is known as pair-wise alignment (PA). As an extension of PA, Multiple Sequence Alignment (MSA) came into exists, where in this more than two sequences are aligned. These sequences can be aligned locally or globally depending upon the information needed for the analysis. The main applications of sequence alignment are to construct phylogenetic tree, to find motif (conserved patterns), gene promoter. These identified motifs give more information that is used to study the relationship between the sequences and its consequences. The different forms of sequence alignment are depicted in Figure 2. The sequences in the pair-wise alignment in the figure 2 represents protein and in the MSA, DNA sequences. Both the alignment can be done for all types of molecular sequences. MSA reveals additional biological information than pair-wise alignment. Likewise local alignment gives information as global alignment. Local alignments are more useful for dissimilar sequences that are suspected to contain regions of similarity. A locally aligned sequence is used to find the motif patterns. In MSA process, three major tasks are scoring; creating an alignment and assessing its significances. Depending upon the type of alignment scoring function are chosen. Any statistical test applied on the bases of dataset comparison.

Exhaustive and heuristic algorithms are two forms of traditional alignment approaches. Dynamic programming sequence matching algorithm is used in exhaustive method. This method was first used for pair-wise alignment. These types of methods are used only for small number of sequences dataset. Many heuristic algorithms has been encouraged to solve the above mentioned problem. This can be classified into Progressive Alignment (PA), Iterative Alignment (IA) and Block-based Alignment (BA) types. BA method (Mohsen et al., 2011) is used to find the conserved domains and motifs. It is not done through PA and IA methods.

Figure 1. Overview of bioinformatics domain
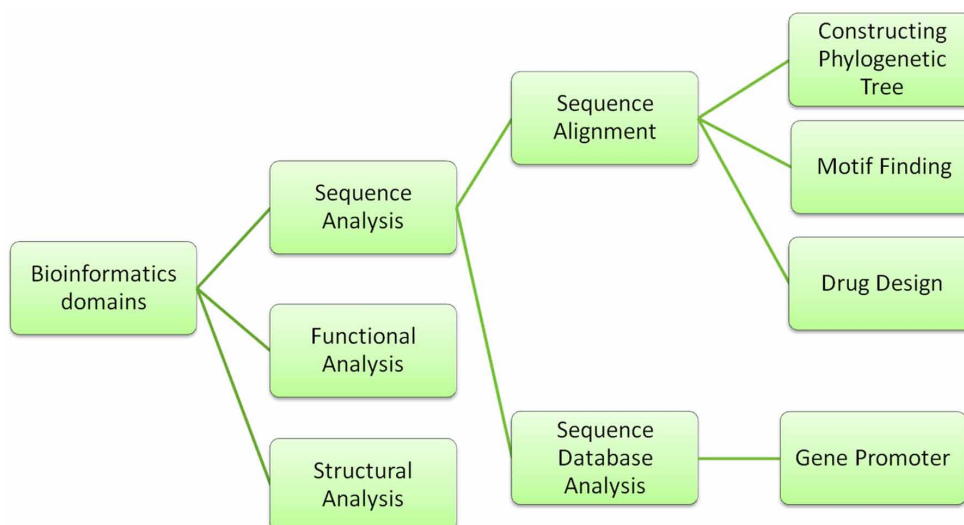
21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/an-improved-hybridized-evolutionary-algorithm-based-on-rules-for-local-sequence-alignment/208049

Related Content

Role of Explainable Artificial Intelligence (XAI) in Prediction of Non-Communicable Diseases (NCDs)
Jana Shafi, Shamayita Basuand Selvani Deepthi Kavila (2022). *Principles and Methods of Explainable Artificial Intelligence in Healthcare (pp. 113-130).*
www.irma-international.org/chapter/role-of-explainable-artificial-intelligence-xai-in-prediction-of-non-communicable-diseases-ncds/304178

Recent Studies and Research on Sickle Cell Disease: Statistical Analysis and Machine Learning Approach
Bikesh Kumar Singhand Hardik Thakkar (2020). *Handbook of Research on Advancements of Artificial Intelligence in Healthcare Engineering (pp. 237-243).*
www.irma-international.org/chapter/recent-studies-and-research-on-sickle-cell-disease/251148

Extending Loosely Coupled Federated Information Systems Using Agent Technology
Manoj A. Thomas, Victoria Y. Yoonand Richard Redmond (2007). *International Journal of Intelligent Information Technologies (pp. 1-20).*
www.irma-international.org/article/extending-loosely-coupled-federated-information/2420

Autonomous Market Segments Estimation Using Density Conscious Artificial Immune System Learner
Vishwambhar Pathak (2017). *Maximizing Business Performance and Efficiency Through Intelligent Systems (pp. 110-135).*
www.irma-international.org/chapter/autonomous-market-segments-estimation-using-density-conscious-artificial-immune-system-learner/178299

Arabic Query Expansion Using WordNet and Association Rules
Ahmed Abbache, Farid Meziane, Ghalem Belalemand Fatma Zohra Belkredim (2016). *International Journal of Intelligent Information Technologies (pp. 51-64).*
www.irma-international.org/article/arabic-query-expansion-using-wordnet-and-association-rules/164512