Chapter 9
Role of Supercomputers in Bioinformatics

Anamika Singh
Department of Botany, Maitreyi College, University of Delhi, India

Rajeev Singh
Division of RCH, Indian Council of Medical Research, India

Neha Gupta
Northeastern University, USA & Osmania University, India

ABSTRACT

Due to the involvement of effective and client-friendly components (i.e. supercomputers), rapid data analysis is being accomplished. In Bioinformatics, it is expanding many areas of research such as genomics, proteomics, metabolomics, etc. Structure-based drug design is one of the major areas of research to cure human malady. This chapter initiates a discussion on supercomputing in sequence analysis with a detailed table summarizing the software and Web-based programs used for sequence analysis. A brief talk on the supercomputing in virtual screening is given where the databases like DOCK, ZINC, EDULISS, etc. are introduced. As the chapter transitions to the next phase, the intricacies of advanced Quantitative Structure-Activity Relationship technologies like Fragment-Based 2D QSAR, Multiple-Field 3D QSAR, and Amino Acid-Based Peptide Prediction are put forth in a manner similar to the concept of abstraction. The supercomputing in docking studies is stressed where docking software for Protein-Ligand docking, Protein-Protein docking, and Multi-Protein docking are provided. The chapter ends with the applications of supercomputing in widely used microarray data analysis.

INTRODUCTION

A supercomputer is a computer with high speed and is calculation efficient. Supercomputers first came in practice during 1960s. These supercomputers are normal and are similar to other computers but have more processors making the speed high. Presently, supercomputers are replaced by parallel supercomputers in which thousands of processors were connected to a single computer (Hoffman et al., 1990, Hill et al., 1999 & Prodan et al., 2007). In this chapter we are focusing on some novel areas

DOI: 10.4018/978-1-4666-7461-5.ch009
Role of Supercomputers in Bioinformatics

Figure 1. A brief history of supercomputing

of biological research where supercomputing is playing a vital role. These few areas are:

**SUPERCOMPUTING IN SEQUENCE ANALYSIS**

Sequence analysis is actually used to explore the DNA, RNA and protein sequences in such a way that it gives all the information about the organism, source, phylogeny, function and structure, and other characteristics. Methodologies used include sequence alignment, searches against biological databases and others. Mostly it is required to search a DNA, a protein or genome database for sequence locations that are similar to that of some query sequence.

These databases already have billions of sequences with characteristics and this sequence information is increasing day by day. Manual searching is tough, a time consuming process and the efficiency of result is questionable. So to look

Source: [www.biomedicalcomputationreview.org](http://www.biomedicalcomputationreview.org)
Related Content

Design and Implementation of Hybrid Time (HT) Group Communication Protocol for Homogeneous Broadcast Groups
[www.igi-global.com/article/design-implementation-hybrid-time-group/55420](www.igi-global.com/article/design-implementation-hybrid-time-group/55420?camid=4v1a)

Hardware Transactional Memories: A Survey
[www.igi-global.com/chapter/hardware-transactional-memories/159039](www.igi-global.com/chapter/hardware-transactional-memories/159039?camid=4v1a)

Trusted Data Management for Grid-Based Medical Applications
[www.igi-global.com/chapter/trusted-data-management-grid-based/54930](www.igi-global.com/chapter/trusted-data-management-grid-based/54930?camid=4v1a)

Testing-Effort Dependent Software Reliability Model for Distributed Systems
[www.igi-global.com/article/testing-effort-dependent-software-reliability/78150](www.igi-global.com/article/testing-effort-dependent-software-reliability/78150?camid=4v1a)