

**REVIEW ARTICLE**ISSN:2394-2371
CODEN (USA):IJPTIL**A Review on Biomedical Importance of computational Biology and Bioinformatics**Shashank Shekher Mishra^{1*}, Neeraj Kumar², C S Sharma¹, Sajal Dixit³¹Department of Pharmaceutical Chemistry, B. N. College of Pharmacy, Udaipur-313001, India²Department of Pharmaceutical Chemistry, Geetanjali University, Udaipur-313001, India**ABSTRACT**

This review provides the summary of application of computational biology and bioinformatics with recent changes and modifications. This review also covers the database and software's which are applied in life science research. The biomedical importance of bioinformatics and computational biology is closely related to enhancement of efficacy and effectiveness and speed of drug development. This entails informatics not only at cellular level, where knowledge and techniques are advancing rapidly but also at the level of biological systems. Bioinformatics is very essential in molecular biology and medicine for understanding the biological result. In this review we also describe the application of computational biology and bioinformatics in the drug design and development process.

Keywords: - Genomics, Molecular biology, Transcriptome, Drug-design, Homology-modeling.

INTRODUCTION

The fields of Computational Biology and Bioinformatics have a large number of transformations during the last few decades. There are many bioinformatic tools and software's are used widely in life science research. A biomedical scientist uses an interdisciplinary approach in calculating and

understanding the biological data. These approaches based on statistics, computer science, chemical science, biological science and engineering. The biological system is a complex system [1]. In biological system a various functionally diverse and frequently multifunctional, element set interact selectively and nonlinearly to produce coherent rather than complex behaviors [1]. The bioorganic compounds such as nucleotides, amino acids and monosaccharide's are the monomeric units of complex biomolecules- nucleic acids (DNA and RNA), proteins and polysaccharides, respectively [2]. Today, bioinformatics

Corresponding Author:*Shashank Shekher Mishra**Department of Pharmaceutical Chemistry
Bhupal Nobles' College of Pharmacy, Udaipur-313001,
Rajasthan, IndiaEmail- mishrashashankshekher@gmail.com

Article Received on: 22-09-2015

Published on: 30-09-2015

becomes as an essential tool in major areas of biology. Bioinformatic methods such as signal processing are widely applied in the molecular biology. It also play a vital role in the sequence analysis of genome in the genetics and genomics, development of biological and gene ontologies to organize and query biological data and analysis and regulation of gene and protein expression. There are various uncovered biological aspects of molecular biology, such as genome sequences and protein properties, but this alone is not sufficient for interpreting biological systems [1]. In structural and experimental biology, bioinformatics helps in the simulation and modelling of nucleic acids, and protein structure with molecular interactions.

Cell biologists searched how all the extraordinary and specific properties of the cell arise from underlying molecular events: the assembly of macromolecules, binding of macromolecules to each other, catalytic effects that enhance particular chemical reaction, and the distribution of information carried by giant molecules [3]. The primary objective of computational biology is to design and develop computer models that apply physical, chemical, and biological principles that mirror the behavior of biologic molecules and methods [4]. Computational biology exploring the

virtual experiments and analyze the biological results by predictive models. Computational biology divided into two different branches: knowledge discovery, or data-mining, which extracts the hidden patterns from huge quantities of experimental data, to construct the hypothesis as a result; and simulation-based analysis, which tests hypothesis with *in silico* experiments, providing predictions to be tested by *in vitro* and *in vivo* studies [1]. In computational biology, there are various computational tools and methods are used for analyzing and interpretation of the biology result.

Need and development of Bioinformatic Resources and computational biology-

In the sequence data analysis, the generation of biological data provides the major opportunities for the development of new methods to problems in computational biology [5]. Biomedical scientist faced the challenges in resolving the multifactorial diseases that demands a large increase in the depth of our knowledge of living organisms [4]. In the post-genome era, biological scientist also posed the significant challenge of utilizing information churned out by multiple 'omics' technologies (genomics, transcriptomics, proteomics, glycomics, metabolomics) [6]. So, biomedical scientists therefore have turned to sophisticated

computational tools to collect and evaluate biologic information on a mass scale [4]. There are various computational models available to resolve the problems of statistical modeling and database searching. One of the computational models is Hidden Markov Models (HMMs) that is applied for database searching and multiple alignments of protein families and protein domains [5]. Other models are Neural Networks, Support Vector Machines, Decision Trees and Random Forests.

Bioinformatics exploits the tedious information storage and processing capabilities of the computer to develop tools for the collection, retrieval, and analysis of biologic data on a mass scale. There are various bioinformatic resources available that can be accessed via the internet, which provides them with global reach and impact. The central objective of a typical bioinformatics project is to assemble all of the available information relevant to a particular topic in a single location, often referred to as a library or database, in a uniform format that renders the data amenable to manipulation and analysis by computer algorithms [4]. The capabilities of bioinformatic databases can vary depending upon the scope and nature of their objectives [2]. First, biomedical information comes in a various forms. For example, the coding

information of a genome is composed of simple linear sequences of four nucleotide bases. While the number of amino acid residues that define a protein's primary structure is minute relative to the number of base pairs in a genome, a description of a protein's x-ray structure requires that the location of each atom be specified in three-dimensional space [4]. Second, anticipating the manner in which users may wish to search or analyze the information within a database, and devising algorithms for coping with these variables, can prove extremely challenging. For example, even the simple task of searching a gene database commonly employs, alone or in various combinations, criteria as diverse as the name of the gene, the name of the protein that it encodes, the biologic function of the gene product, a nucleotide sequence within the gene, a sequence of amino acids within the protein it encodes, the organism in which it is present [4].

Applications of Computational Biology and Bioinformatics- Bioinformatics have wider application in molecular biology and medicine research. In molecular biology for understanding the biological processes, computational tools are used for pattern recognition, data mining, machine learning algorithms and visualization of results.

Resources of Bioinformatic database and software's-

1. Useful bioinformatic websites (available freely on the internet)- [7, 8]

S.No.	Websites
1.	National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) -maintains bioinformatic tools and databases
2.	National Center for Genome Resources (www.ncgr.org/) -links scientists to bioinformatics solutions by collaborations, data, and software development
3.	Genbank (www.ncbi.nlm.nih.gov/Genbank) -stores and archives DNA sequences from both large scale genome projects and individual laboratories
4.	Unigene (www.ncbi.nlm.nih.gov/UniGene) -gene sequence collection
5.	European Bioinformatic Institute (www.ebi.ac.uk) -centre for research and services in bioinformatics; manages databases of biological data
6.	Ensembl (www.ensembl.org) -automatic annotation database on genomes
7.	BioInform (www.bioinform.com) -global bioinformatics news service
8.	Swissprot (www.expasy.org/sprot/) -important protein database with sequence data from all organisms
9.	International Society for Computational Biology (www.iscb.org/) -aims to advance scientific understanding of living systems through computation
10.	IUBio Archive, Indiana University (www.iubio.bio/) -centre for research and services in bioinformatics

2. Bioinformatic Catalogues- [7, 8]

S. No.	Bioinformatic Catalogues
1.	BioHunt, ExPASy http://www.expasy.org/BioHunt/ Automatic robot updated list
2.	Bioinformatics.ca http://www.bioinformatics.ca/links_directory
3.	Bioinformatics.net http://www.bioinformatics.net/
4.	Bioinformatik.de http://www.bioinformatik.de/
5.	BioNetbook, Pasteur Institute http://www.pasteur.fr/recherche/BNB/bnb-en.html Semi-automatic updates
6.	CSM Molecular Biology Resource, SDSC http://restools.sdsc.edu/

7.	GenomeWeb	http://www.hgmp.mrc.ac.uk/GenomeWeb/
8.	Open Directory Project Submitted links	http://dmoz.org/Science/Biology/Bioinformatics/
9.	SouthWest Biotechnology and Informatics Center	http://www.swbic.org/

3. News and discussion groups- [8]

S. No.	Bioinformatic News and Discussion groups		
1.	BIOSCI/Bionet	http://www.bio.net/	Biologist and bioinformatics focus
2.	Bioinformatics.org	http://bioinformatics.org/	Bioinformatics focus
3.	Bioinformatics.net	http://bioinformatics.net/	Biologist focus

4. Selected bioinformatics publications for software tools- [8]

S. No.	Bioinformatic Publications		
1.	BioInform	http://www.bioinform.com/	Bioinformatics news briefs
2.	Bioinformatics	http://bioinformatics.oupjournals.org/	Bioinformatics focus
3.	BMC Bioinformatics	http://www.biomedcentral.com/bmcbioinformatics/	Bioinformatics focus

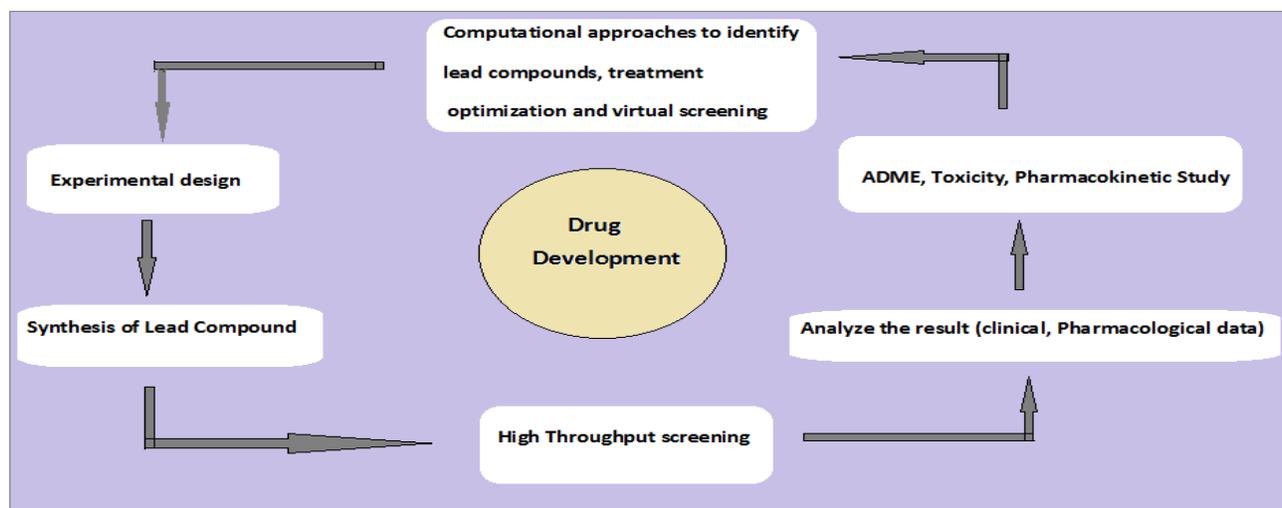


Fig 1.1 Overview of Drug-design and development process

The major research efforts in the field include sequence alignment, gene finding, genome assembly, drug design, drug discovery, drug development, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, genome wide association studies, and the modelling of evolution. Prominent in life science research today has been the introduction and application of various next-generation sequencing platforms to a wide range of research problems from rapid de novo sequencing, to genomic evaluation of polymorphisms and other mutation events, and analyses of the transcriptome of various RNA macromolecules and the epigenome [9].

Computational biology provides the tools for understanding and interpretations of the biological result. Advances in the computer technology that provide faster computations fuelled the expansion of computer biology into many different scientific areas, as did the development of new mathematical algorithms that allowed highly sophisticated problems to be solved quickly. Computational biology and bioinformatics provide basic framework for development of new drug or drug-design and drug-development. These bioinformatic tools reduce the costs and time of experimental hypothesis.

1. Translational Bioinformatics-

Translational bioinformatics is an innovative area that focuses on the computational challenges in biomedical research and the analysis of the huge amount of clinical data generated from tools [10]. In the translational bioinformatics, disease genes are searched and then compared with normal genes. The relations between diseases, phenotype and mechanisms have been explained in an attempt to identify potential new applications for approved drugs, which could lead and raise the speed of drug development and reduce overall costs [11]. The study of disease genes has evolved from basic assumptions that genes follow Mendelian laws to modern computational methods that are capable of providing insight on hundreds of genes and discriminate particular mutations associated with diseases [12]. The major approach in the field have lead to basic knowledge of the functional, networking and evolutionary properties of disease genes as well as to the identification of genes for specific diseases [12].

2. Structural Bioinformatics-

In structural bioinformatics, protein structure is predicted. The primary structure of protein in which amino acid is sequenced is determined from the sequence on the gene that encodes it. The

secondary, tertiary and quaternary structure covers the structural information of proteins. The prediction of protein structure remains an open problem. Proteomics is a scientific endeavor that attempts to study the sum total of all the proteins in a cell from the point of view of their individual functions and how the interaction of specific proteins with other cellular components affect the functions of these proteins [13]. High-throughput high-resolution mass spectrometry allows the amino-acid sequences of proteins to be determined very quickly.

Structure based discovery of drugs or inhibitors requires 3-D protein structures. SWISS-MODEL workspace is a web-based integrated service dedicated to protein structure homology modelling that assists and guides the user in building protein homology models at different levels of complexity [14].

3. Sequence Analysis- A thousands of DNA sequences is decoded and this information is analyzed to determine genes, RNA genes, regulatory sequences, etc. For sequence analysis various techniques are used such as shotgun sequencing technique, annotation and genomics. Sequencing technologies are becoming affordable and are replacing the microarray-based genotyping methods, which were limited to interrogating regions of known

variation [15]. The field of computer science uses genetic algorithms is different from computational evolutionary biology. Bioinformatics also translates discoveries to the clinic by disseminating discoveries through curated, searchable databases like PharmGKB, dbGaP, PacDB and FDA AERS [15].

4. Feature Selection Techniques- This technique is more valuable and has various advantages. The objectives of this technique is to avoid overfitting and improve model performance, i.e. prediction performance in the case of supervised classification and better cluster detection in the case of clustering, to provide faster and more cost effective models and to gain a deeper insight into the underlying processes that generated the data [16].

5. Network Biology- Network biology focuses on molecular interactions that determine the function of this complex machinery [17]. The biological networks may be constructed from a single molecule or entity (such as genes), network biology associated with many different data types, such as proteins, small molecules, gene expression data, and others, which are all connected physically, functionally, or both.

CONCLUSION

The aim of this paper was to increase the familiarity of computational biology techniques

and their biomedical importance in the bioinformatics community and to provide useful information from literature and also give information about available software. Bioinformatics is a computational discipline that's central objective is to manage and analyze the biological data [18]. Clinical bioinformatics is a new and related emerging subfield that combines clinical informatics, bioinformatics, medical informatics, information technology, mathematics, and omics science together [19]. Computational biology focuses on the algorithmic aspects. In conclusion, I am delighted to report that the application of computational biology and bioinformatics is a thriving field of research [20]. The notion of computing in biology is now enthroned as the pillar of new biology. The merging of computation into the fabric of biosciences and biomedicine by 2020, while acquiring a key, critical position amongst other technologies in the toolkit, will possibly necessitate a redefinition of computational biology as a distinct discipline in the not so distant future [21]. In the last 20 years, bioinformatics has evolved from protein sequences alignment, through the analysis of 2-DE gel images. Certainly, improvements directed towards increasing performance, reducing operation time and making automation a more realistic goal are still needed

[22]. The development of computational biology and bioinformatic tools has vital biomedical importance in the design of new molecular entity.

REFERENCES

1. Kitano, H. (2002). Computational systems biology. *Nature*, 420(6912), 206-210.
2. Satyanarayan, U., & Chakrapani, U. (1999). *Textbook of Biochemistry*.
3. Lodish, M. H., Berk, M. A., & Matsudaira, P. (2005). *Biologie moléculaire de la cellule*. De Boeck Supérieur.
4. Granner, R. K. M. D. K., & Rodwell, P. A. M. V. W. (2006). *Harper's Illustrated Biochemistry (Harper's Biochemistry)*. McGraw-Hill Medical.
5. Krogh, A., Brown, M., Mian, I. S., Sjölander, K., & Haussler, D. (1994). Hidden Markov models in computational biology: Applications to protein modeling. *Journal of molecular biology*, 235(5), 1501-1531.
6. Lee, W. P., & Tzou, W. S. (2009). Computational methods for discovering gene networks from expression data. *Briefings in bioinformatics*, 10(4), 408-423.

7. Bayat, A. (2002). Science, medicine, and the future: Bioinformatics. *BMJ: British Medical Journal*, 324(7344), 1018.
8. Gilbert, D. (2004). Bioinformatics software resources. *Briefings in bioinformatics*, 5(3), 300-304.
9. Brazas, M. D., Yamada, J. T., & Ouellette, B. F. (2009). Evolution in bioinformatic resources: 2009 update on the Bioinformatics Links Directory. *Nucleic acids research*, 37(suppl 2), W3-W5.
10. Butte, A. J. (2008). Translational bioinformatics: coming of age. *Journal of the American Medical Informatics Association*, 15(6), 709-714.
11. Wang, K., Nemenman, I., Banerjee, N., Margolin, A. A., & Califano, A. (2006, January). Genome-wide discovery of modulators of transcriptional interactions in human B lymphocytes. In *Research in Computational Molecular Biology* (pp. 348-362). Springer Berlin Heidelberg.
12. Kann, M. G. (2010). Advances in translational bioinformatics: computational approaches for the hunting of disease genes. *Briefings in bioinformatics*, 11(1), 96-110.
13. Beale, J. M., Block, J., & Hill, R. A. (2010). *Organic medicinal and pharmaceutical chemistry*. Philadelphia: Lippincott Williams & Wilkins.
14. Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2), 195-201.
15. Fernald, G. H., Capriotti, E., Daneshjou, R., Karczewski, K. J., & Altman, R. B. (2011). Bioinformatics challenges for personalized medicine. *Bioinformatics*, 27(13), 1741-1748.
16. Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *bioinformatics*, 23(19), 2507-2517.
17. Barabasi, A. L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature reviews genetics*, 5(2), 101-113.
18. Merelli, E., Armano, G., Cannata, N., Corradini, F., d'Inverno, M., Doms, A., ... & Luck, M. (2007). Agents in bioinformatics, computational and systems biology. *Briefings in bioinformatics*, 8(1), 45-59.

19. Wang, X., & Liotta, L. A. (2011). Clinical bioinformatics: a new emerging science. *J. Clinical Bioinformatics*, 1, 1.
20. Lio, P. (2003). Wavelets in bioinformatics and computational biology: state of art and perspectives. *Bioinformatics*, 19(1), 2-9.
21. Ouzounis, C. A. (2012). Rise and demise of bioinformatics? Promise and progress. *PLoS Comput Biol*, 8(4), e1002487.
22. Palagi, P. M., Hernandez, P., Walther, D., & Appel, R. D. (2006). Proteome informatics I: bioinformatics tools for processing experimental data. *Proteomics*, 6(20), 5435-5444