How to Cite:
Tyagi, A., Sharma, A., & Bhardwaj, M. (2022). Future of bioinformatics in India: A survey. *International Journal of Health Sciences*, *6*(S2), 13767–13778. https://doi.org/10.53730/ijhs.v6nS2.8624

Future of bioinformatics in India: A survey

Anushka Tyagi
Department of Computer Science and Information Technology, KIET Group of Institutions, Delhi-NCR, Muradnagar, India

Aman Sharma
Department of Computer Science and Information Technology, KIET Group of Institutions, Delhi-NCR, Muradnagar, India

Dr. Manish Bhardwaj
Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Muradnagar, India

***Abstract*---**Bioinformatics is an interdisciplinary field that develops methods and software tools for analyzing biological data in protein and nucleic acid structures. Currently, it is the most effective method to deal with the flood of information and further reduces risks, time, and cost due to the transition of wet lab work into dry lab work. India- a biotech growth catalyst, is among the top 12 destinations for biotech in the world yet holds only a 3% share of the global biotech industry. Thus, this paper deals with the future analysis of bioinformatics in India. It introduces bioinformatics and leads to an emphasis on its importance and the opportunities it opens. Furthermore, contemplates the present challenges and possible solutions that could be made to establish this field as a proliferating industry in the Indian and global market.

***Keywords*---**bioinformatics, human genome project, computational biology, biotech industry.

Introduction

Bioinformatics is a dynamic and emerging field that is used widely in drug production, gene therapy, genetic testing, agriculture, livestock, etc. It is necessarily an amalgamation of several domains like molecular Biology, Chemistry, Mathematics, Statistics, and computer science. There is no unified definition of bioinformatics however, "Bioinformatics is a subdiscipline of biology and computer science concerned with the capture, storage, analysis, and transmission of biological data, most typically DNA and amino acid sequences," according to the National Human Genome Research Institute [1].

Conventionally, biology is considered a haven for microscopes and the study of the living environment. With the advancement of physics, chemistry, and computer technology there has been an increase in sophisticated biological information. The Human Genome Project [2] was a massive blow of information in which the entire human genome of 3.3 billion base pairs was arranged and acted as a possible solution to genetic therapy and hazardous genetic diseases. Therefore, a computational infrastructure was needed for genomic data analysis and thus led to the growth of bioinformatics. Companies [3] related to drugs, agriculture, chemicals, and hybrid plants have taken bioinformatics as a considerable method to produce new drugs with limited use of natural resources. Bioinformatics is an emerging and young field regarding market size and has a long way to become a fully established sector. It must develop new and improved algorithms for data mining, analysis, and comparisons etc. This survey paper will target the future analysis of bioinformatics and its growth rate in India and the world and will deal with a discussion over the opportunities, challenges, implications, etc.

Literature Review

Bioinformatics is the ultimate solution to discover new previously unknown and hidden patterns within each protein and nucleic acid. It deals with highly complex data via computer programming as a part of its methodology. The data from the human genome project [2] was so large that it could fill 200 volumes of 1000 pages and reading it would take 26 years of round-the-clock work. As a result, bioinformatics enables fascinating tasks like comparing DNA sequences and producing potentially important and industrially valuable outcomes. It is an important field in today's world that can provide capabilities for integrating genetic data with existing and future experimental results in both graphical and text-based genomic mapping. The genomic testing capabilities will improve with the use of bioinformatics tools and next-generation sequencing. However, in order to extract a significant amount of data, we need a proper computing infrastructure for genomic data analysis in order to translate complex genomic data into useful insights.

Bioinformatics Tools and Applications [4]

1. Sequence analysis: It entails submitting RNA, peptide sequences, and DNA to a variety of analytical techniques. The software's monitoring feature assists in the examination of the following: -
   - *Origin*
   - *Evolution*
   - *Structure of Biological Database*
2. Molecular modeling: It makes use of computational and theoretical methodologies for the analysis of the behavior of molecules. "Molecular modeling comprises all methods, theoretical and computational, used to model or simulate the behavior of molecules," according to Wikipedia [5].
3. Molecular dynamics: It determines the physical motion of biological atoms and molecules. In this, the constituent particles are allowed to interact with each other for a fixed interval of time to monitor their physical moment.

4. FASTA in Bioinformatics: FASTA is a text-based peptide and nucleotide sequence representation system. The tools aid in DNA alignment sequencing and protein sequencing.

Bioinformatics is the only way by which we can cope with this flood of information. Now, with the advancement of computation technology and the fusion of Biology, Statistics, and Computer Science, researchers and scientists can look for divergence and similarities between human and other animal species. The risks [5] involved in the in vitro (experimental model creation) and in vivo (clinical trials) laboratory processes for drug discovery are largely removed by bioinformatics. With the total world population of 7.9 Billion people and every two individuals differing in 3 million base pairs, it becomes a challenge to track the sequencing of the entire human genome(having 3 billion base pairs in 23 chromosomes), which leads to the mapping of 100,000 genes. UT South Western and the University of Washington researchers founded more than 100 probable protein complexes for the first time and provided structural models for more than 700 proteins that were earlier uncharacterized. This research used evolutionary computer-based techniques like Artificial Intelligence to produce 3-D models of Eukaryotic protein [6] interaction. The research would provide new ways a protein or group of proteins can fit together and can carry out distinguished metabolic activities in our biological bodies. This led to the discovery of new drug targets concerning different cellular processes. "Our results reflect a significant achievement in the new age in structural biology in which computation plays a vital role," says Qian Cong, Ph.D., Assistant Professor in the Eugene McDermott Center for Human Growth and Development with a secondary appointment in Biophysics.

In recent context, the most popular example can be seen in the case of covid19, within which genome sequencing was studied using bioinformatics tools. Covid-19 [4] has protein binding on its membrane and is primarily an RNA virus. It is because of these components the virus has acquired the character it has and is fast mutating. The arrangement of DNA, RNA, and Protein in the virus further helps in the understanding of changing nature. Therefore, helps in knowing a potential drug for medicine and vaccination preparation. Figure: 1 shows the types of Genomic Testing [4] could be of four major types: -Pharmacogenomic testing, Tumor Testing, Diagnostic Testing, Clinical predictive testing. Genome-wide analysis and testing help in monitoring and further can prevent any diseases

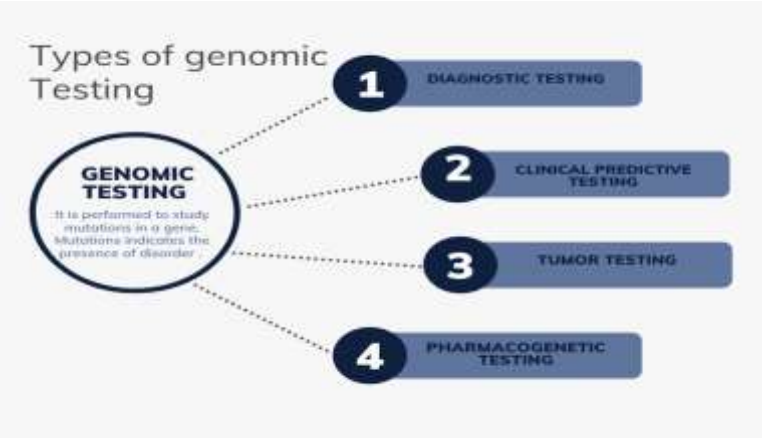

Figure:1 Types of Genomic Testing

In silico modelling [2], the fastest-growing area, aids in the testing of substances by mimicking the reactions of cells and animals using computer programmes. This reduces the amount of time and money spent in the lab. Google, Microsoft, Ginkgo Bioworks, Codexis, Twist Bioscience, and other prominent players [7] in computer and medicine have made significant investments in computational biology. The mortality rate is high in regions like India and other developing countries, and this is typically attributable to a lack of a comprehensive computational biology network. As a result, predicting the source and nature of illnesses is challenging. Figure 2 depicts the distribution of Thalassemia in important areas of developing countries, as depicted by the accompanying map [8] (by the National Institute of Health, United States of America).

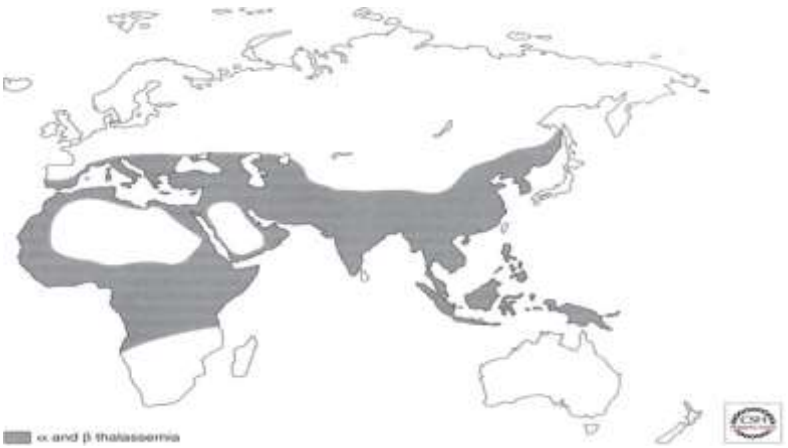

Figure:2 Distribution of alpha and beta-thalassemia

India-a Biotech Growth Catalyst

India is one of the world's top 12 biotech [9] locations. Many pharmaceutical, information technology, and biotechnology companies have entered the bioinformatics sector as a result of increasing volumes of genomic data and an increasing number of participants contracting work to Indian companies. Tata

Consultancy Services (TCS), Cognizant Technologies, Infosys, and WIPRO are among the Indian companies [10] that have already established bioinformatics divisions. Despite the rapid growth of biotechnology in India, the country only has a 3% share in worldwide biotech industries [9]. The use of bioinformatics in drug delivery and development is expected to cut the annual cost of developing a new drug by 33% and the time it takes to discover a new drug by 30%. "Bioinformatics contributes to the smallest part of the biotechnology industry in India," says Rashmi Sanyal [11]. This sector is concerned with the creation and maintenance of large electronic databases on various biological systems, as well as the tools and software used to analyze the data obtained from these databases. India has long been known for its strong IT infrastructure. As a result, it puts India in a better position than other developing countries to become a leading destination for bioinformatics work. To take a competitive advantage in the drug discovery process, pharmaceutical and life science companies emphasize bioinformatics as it reduces the time and costs of developing medicine because of its facilitation with filtering data.

The Issue to be Resolved Includes the Following

There are a slew of roadblocks impeding the growth of computational biology in India, and the issues that need to be resolved are as follows: -

- Due to a lack of interest and awareness, India's software industry does not contribute much to bioinformatics. As a result, there is no efficient database integration, making it difficult to query multiple databases at once.
- Inability to maintain quality checks and gain rapid approval for products has resulted in a lack of public-private relationships [11] and a lack of angel funding.
- The lack of a bridge between IT leaders and biologists is a major barrier to optimal bioinformatics exploration for solving life sciences issues.

Take, for example, data translation [12]: Biological data is available as a variety of web services all over the world, providing biologists with useful information. However, the disparate data formats present a technical barrier to full utilisation. They must use Bio Java on their computer to convert the documents into XML formats. Protein folding [12]: The resolution obtained is usually low and as a result, it is difficult to understand or predict the folding rates of protein from protein 3-D structure.

Possible Solutions [10]

A possible solution is to have enough leaders with experience in both life sciences and information technology, as well as a strong institutional network that effectively connects specialists from both fields. "One of the major challenges in the optimal exploration of bioinformatics for solving life science issues is the formulation of appropriate computational biology problems that can be addressed through IT tools," according to the department of biotechnology. The following are the conditions for resolving the issues: -

- Consolidation of all data into a single source and development of user-friendly software Almost all disease advances necessitate the use of computer-aided technologies.
- Because it is a multidisciplinary field, computer and other information science students must have a basic understanding of biological science and sequences.
- Appreciation of the scope and strength of bioinformatics as a field by biologists, in order to raise public awareness of the merits and opportunities offered by this field.

India is unquestionably one of the most promising countries in the fields of biotechnology and computational biology. In general, laboratory system management is an ideal, flexible, and multifunctional tool for managing and improving the efficiency and quality of analytic studies in laboratories that operate in a variety of fields. Because of the paradigm shift in biological research, bioinformatics, and computational biology centers involving data mining, artificial intelligence, machine learning, and other techniques have become necessary.

Investment Report and Recent Development

"The need for Bioinformatics tools and skills has increased as genome sequencing initiatives have resulted in the exponential growth of sequence databases," according to a recent report from the Department of Biotechnology [10]. BITSNet[10] has established itself as a highly dependable bioinformatics infrastructure with a wide range of scientific features and communication capabilities. Cancer, pulmonary diseases, diabetes, tuberculosis, and cardiovascular diseases all have established programs. The motivations for any country to invest [9] in India are explained in Figure 3 above. A project on imaging biobanks [9] for cancer is also being financed by Tata Memorial Centre for Cancer Research (TMC-ACTREC), Mumbai, AIMS, Delhi, IIT(Bombay), RGCIRC; a unit of the Indraprastha Cancer Research Society, Delhi, with the goal of developing AI tools.

Figure: 3 Advantages in India

Biotechnology [9] is regarded as one of the primary contributors to India's goal of a USD 5 trillion GDP by 2024. With a population of 1.3 billion people, half of whom are under the age of 25, India is one of the most youthful and well-equipped countries in the world. According to the India Brand Equity Foundation, the country's biotech industry includes more than 2,700 start-ups and 2,500 corporations. In the United States, India has 665 FDA-approved factories, 44 percent of global abbreviated new drug applications (ANDA), and over 1400 WHO-compliant manufacturing units. The country is also the second-largest producer of BT cotton and the third-largest producer of recombinant Hepatitis B vaccine [9]. (genetically modified pest-resistant plant cotton). In numerous aspects, computer science is linked to biotechnology, particularly in the field of bioinformatics, which entails the creation of software tools and methods for analyzing biological data. As a result, plants seeds for new innovative jobs [13] in software development, machine learning, artificial intelligence, and other fields in India and other emerging countries.

- Computational Biologists
- Bioinformatics Analysts
- Bioinformatics Engineers

India-Global Market Comparison

The Bioinformatics market research details the global, regional, and country-level market sizes, as well as the strategic gameplay. Forecasts from 2021 to 2026, according to Bioinformatics Market Report Forecasts [14]. Over the projected period, the global Bioinformatics market is predicted to increase at a compound annual rate of 19.36%, from US$ 10.729 billion in 2019 to US$ 37.029 billion in 2026. According to Acute Market Reports [15], Inc.'s latest market report, "Global Bioinformatics Market – Growth, Share, Opportunities, Competitive Analysis, and Forecast, 2016 – 2024," the bioinformatics market was valued at USD 8,580.0 Mn in 2015 and is expected to reach USD 38,614.0 Mn by 2024, growing at a CAGR of 20.3 percent between 2016 and 2024. From 2015 to 2026, Figure 4 depicts the Global Bioinformatics Market's explosive expansion.

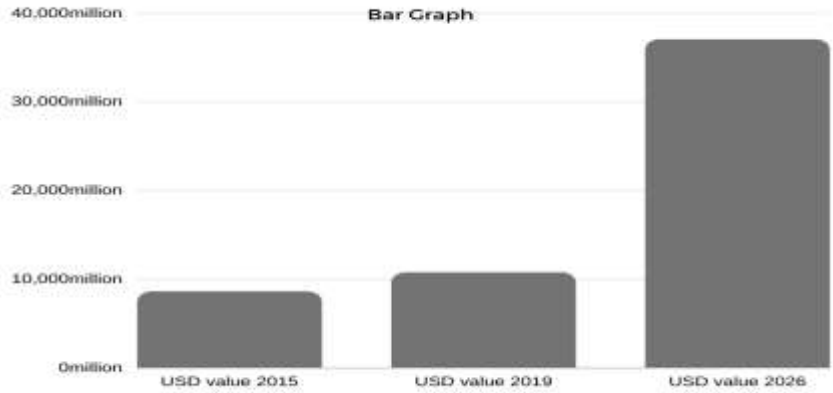

Figure: 4 Growth Rate of Global Bioinformatics Market

According to a report published by the India Brand Equity Foundation [9], the Indian biotechnology industry was worth US$ 63 billion in 2019 and is expected to grow at a CAGR of 16.4% to US$ 150 billion by 2025. The Indian biotechnology industry is predicted to increase its share of the global biotechnology market to 19% by 2025, up from 3% in 2017. Biopharmaceuticals, the largest segment in the Indian biotechnology market, accounted for 62 percent in 2020 and 58 percent in 2019. From 2019 to 2025, the Indian biologics market is predicted to grow at a CAGR of 22%, reaching US$ 12 billion. Clinical trials, contract research, and manufacturing operations are increasingly being conducted in bio-services, which will account for 15% of the biotechnology industry by 2020.

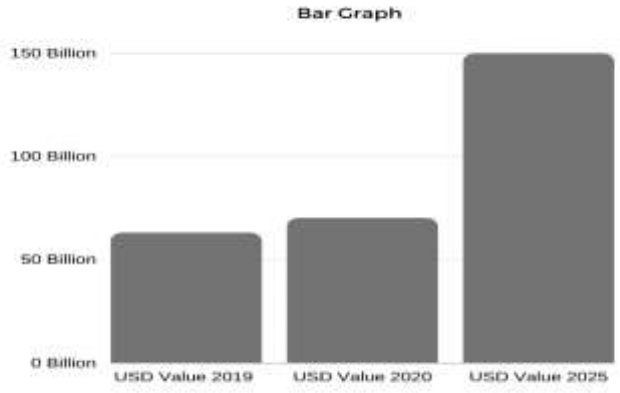

Figure:5 Growth Rate of Bioinformatics Market in India

Both Figures 4 and 5 reveal the rampant growth of Bioinformatics in the global and Indian market in the recent as well as upcoming years.

Conclusion

Bioinformatics is a new developing science in various countries and is a unique, efficient combination of Computer Science and Molecular Biology. Software used in this field helps to analyze and view biological data like DNA, RNA, and various other genomic sequences. Furthermore, the information gathered via this analysis could be used in numerous sectors like Drug Production, Gene Therapy, Gene Testing, and Diagnostics. Another paramount application of bioinformatics is Bio-economics, which deals with the investment in applied Biotechnology to increase the economic input of a country. Therefore, creates new jobs in the IT and Computation sector. Indian position as an IT nation is globally recognized and has one of the youngest populations to strengthen a plethora of computation tools and software to make bioinformatics a well-established field. Indian bioinformatics is expected to grow exponentially in the next few years and further will cause the creation of new opportunities for Indian software and AI Companies to grow their market size. This paper gives a detailed explanation about Bioinformatics, its application, and new upcoming innovations in the world. It further emphasizes its importance and opportunities provided, with the growing trends in India and International level.

Conflict of Interest: The author does not have any conflict of interest

References

1. Bioinformatics - National Human Genome Research,www.genome.gov/genetics-glossary/Bioinformatics
2. Dr. Pankaj M. Madhani; “Indian Bioinformatics: Growth Opportunities and Challenges,”pg-2; ICFAI Business School
3. Manisha Mathur; “Bioinformatics challenges: A review,”pg-1; International Journal of Advanced Scientific Research; Volume 3; Issue 6; November 2018; Page No. 29-33
4. 15 Open Source and Free Bioinformatics Tools List for genomic Testing,www.techjockey.com/blog/bioinformatics-tools;
5. Dr. Pankaj M. Madhani; “Indian Bioinformatics: Growth Opportunities and Challenges,”pg-3; ICFAI Business School
6. Molecular Modeling, https://en.wikipedia.org/wiki/Molecular_modelling
7. Artificial intelligence successfully predicts protein interactions,https://www.utsouthwestern.edu/newsroom/articles/year-2021/artificial
8. M.S. in Computer Biology, https://www.cmu.edu/ms-compbio/prospective-students/career.html#
9. World Distribution, Population Genetics, and Health Burden of the Hemoglobinopathies, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3426822/
10. Biotechnology industry in India, https://www.ibef.org/industry/biotechnology-india.aspx
11. Bioinformatics, AI & Big Data,https://dbtindia.gov.in/schemes-programmes/research-development/knowledge-generation-discovery-research-new-tools-and-0
12. Biotech Sector India and USA - How much behind is India to the US? https://www.biotecnika.org/2020/01/biotech-sector-india-and-usa-how-much-behind-is-india-to-the-us/?amp
13. Manisha Mathur; “Bioinformatics challenges: A review,”pg-4
14. The Top 3 Best Computer Science Jobs in Biotechnology, www.biotech-careers.org/articles/top-3-best-computer-science-jobs-biotechnology
15. Bioinformatics Industry Outlook, 2026 - Featuring Ayasdi, SciGenome Labs, TeselaGen, Hybrigenics, Era7 Information Tech, Lumenogix, Agilent Tech, Thermo Fisher Scientific, DNAnexus, and Biobase, https://finance.yahoo.com/news/bioinformatics-industry-outlook-2026-featuring-125800801.html
16. Bioinformatics Market (20.3% CAGR) 2015 to 2022: Global Industry Size, Share, Growth, Trends and Forecast, http://www.mobilecomputingtoday.co.uk/11694/bioinformatics-market-20-3-cagr-2015-2022-global-industry-size-share-growth-trends-forecast/
17. Bioinformatics, https://en.wikipedia.org/wiki/Bioinformatics
18. Jonathan CF, Pierre K, Holger D, Kristoffer F, International Journal of Advanced Scientific Research Alexandros S, Joseph B, et al. Biggest challenges in bioinformatics. EMBO Rep. 2013; 14(4):302-304. doi: 10.1038/embor.2013.34.

19. Al Kawam, Sen A, Datta A, Dickey N. Understanding the Bioinformatics Challenges of Integrating Genomics Into Healthcare. IEEE Journal of Biomedical and Health Informatics. 2018; 22(5):1672-1683. doi: 10.1109/JBHI.2017.2778263.
20. Guy Haskin Fernald, Emidio Capriotti, Roxana Daneshjou, Konrad J Karczewski, Russ B Altman.
21. Bioinformatics challenges for personalized medicine, Bioinformatics. 2011; 27(13):1741-1748. https://doi.org/10.1093/bioinformatics/btr295
22. Khalid Raza. Application of data mining in bioinformatics; Indian Journal of Computer Science and Engineering. 1(2):114-118
23. Mohammed J Zaki, George Karypis, Jiong Yang. Data Mining in Bioinformatics (BIOKDD) Algorithms Mol Biol. 2007; 2:4. Published online 2007 Apr 11. doi: 10.1186/1748-7188-2-4 PMCID: PMC1852315
24. ¬Raut S.A.; Sathe S.R.; Raut A.;"Bioinformatics: Trends in gene expression analysis," Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference on, vol., no., pp.97-100, 16-18 April 2010
25. See-Kiong Ng; Limsoon Wong; "Accomplishments and challenges in bioinformatics," IT Professional , vol.6, no.1, pp. 44- 50, Jan.-Feb. 2004
26. Chavan PR; “Application of Bioinformatics in the Field of Cancer Research”, 11th Workshop on Medical Informatics & CME on Biomedical Communication, vol., no., 20-22 November 2008.
27. Ackovska N.; Madevska Bogdanova A.; "Teaching Bioinformatics to Computer Science Students," Computer as a Tool, 2005. EUROCON 2005.The International Conference on Computers as a Tool, vol.1, no., pp.811-814, 21-24 Nov. 2005
28. Zadeh J.; "An undergraduate program in bioinformatics," Potentials, July-Aug. 2006; IEEE, vol.25, no.3, pp.43-46.
29. Kasabov N.; "Bioinformatics: a knowledge engineering approach," Intelligent Systems, 2004 Proceedings. 2004 2nd International IEEE Conference, June 2004; vol.1, no., pp. 19-24 Vol.1, 22-24.
30. Fulekar M.H. and J. Sharma. ; “Bioinformatics Applied in Bioremediation”. Innovative Romanian Food Biotechnology. Vol. 2 No. 2. pp 28-36.
31. David Fenstermacher. ; Introduction to bioinformatics: Research Articles, Journal of the American Society for Information Science and Technology, March 2005; v.56 n.5, p.440-446.
32. Achuthsankar S Nair.; “Computational Biology & Bioinformatics – A gentle Overview”, Communications of
33. Computer Scioety of India, January 2007. Doom T.; Raymer M.; Krane D.; Garcia O.; "Crossing the interdisciplinary barrier: a baccalaureate computer science option in bioinformatics," Education, IEEE Transactions on , Aug. 2003; vol.46, no.3, pp. 387- 393.
34. Jana R. ; Aqel M. ; Srivastava P. and Mahanti P. K. ; Soft Computing Methodologies in Bioinformatics, European Journal of Scientific Research, 2009, Vol. 26, No. 2, pp. 189-203.
35. Jacques Cohen. ; Computer science and bioinformatics, Communications of the ACM, March 2005; v.48 n.3, p.72-78.
36. DOMOKOS A.; Bioinformatics and Computational Biology. Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca. Horticulture, North America, 6527, p. 571 – 574, 09 2008.

37. Poe D.; Venkatraman N.; Hansen C.; Singh, G.; "Component Based Approach for Educating Students in Bioinformatics," Education, IEEE Transactions on , Feb. 2009; vol.52, no.1, pp.1-9,
38. Bayat A.; Science, medicine, and the future: Bioinformatics. BMJ. 2002; 324:1018–1022
39. National Center for Biotechnology, "Bioinformatics Factsheet," http://www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html, last accessed June 13, 2012.
40. Gavin J. Gordon.; Bioinformatics in Cancer and Cancer Therapy (Cancer Drug Discovery and Development) [Kindle Edition], ISBN: 978-1-58829-753-2 e-ISBN: 978-1-59745-576-3, Library of Congress Control Number: 2008931368.
41. Jacques Cohen.; Computer science and bioinformatics, Communications of the ACM, March 2005; v.48 n.3, p.72-78,
42. Umarji M.; Seaman C.; Koru A.G.; Hongfang Liu; "Software Engineering Education for Bioinformatics," Software Engineering Education and Training, 2009. CSEET '09. 22nd Conference on, Feb. 2009; vol., no., pp.216-223, 17-20.
43. Simon R.; Bioinformatics in cancer therapeutics-hype or hope? Nat Clin Pract Oncol. 2005;2:223 Goldin L.; "Bioinformatics Integration for Cancer Research Goal Question analysis," Information Technology: Research and Education, 2006. ITRE'06. International Conference on, Oct. 2006; vol., no., pp.248-252, 16-19.
44. Kihara D.; Yang YD.; Hawkins T.; Bioinformatics resources for cancer research with an emphasis on gene function and structure prediction tools. Cancer Inform. 2007; 2:25-35.
45. Ardekani AM.; Aslani F.; Lakpour N.; Application of genomics and proteomics technologies to early diagnosis of reproductive organ cancers. J Reprod Infertil. 2007; 8(3):259-278.
46. Attwood TK, Flower DR, Lewis AP, Mabey JE, Morgan SR, Scordis P and et al. (1999). PRINTS prepares for the new millennium. Nucleic Acids Res. 27(1): 220-225. DOI: 10.1093/nar/27.1.220
47. Attwood TK, Gisel A, Eriksson, NE, Bongcam-Rudloff E (2011) Concepts, Historical Milestones and the Central Place of Bioinformatics in Modern Biology: A European Perspective". Bioinformatics - Trends and Methodologies. InTech. Retrieved 8 Jan 2012. DOI: 10.5772/23535
48. Baker W, Van den Broek, A, Camon E, Hingamp P, Sterk P, Stoesser G and et al. (2000) The EMBL nucleotide sequence database. Nucleic Acids Res 28(1):19-23. DOI: 10.1093/nar/28.1.19
49. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL (2000) The Pfam protein families database. Nucleic Acids Res. 28(1): 263-266. DOI: 10.1093/nar/28.1.263
50. Benson DA, Karsch-Mizrachi I, Lipman, DJ, Ostell J, Rapp BA and Wheeler DL (2000) GenBank. Nucleic Acids Res. 28(1): 15-18. DOI: 10.1093/nar/28.1.15
51. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. Nucleic Acids Res. 41(Database issue): D36-42. DOI: 10.1093/nar/gks1195. DOI: 10.1093/nar/gks1195

52. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H and et al. (2000) The Protein Data Bank. Nucleic Acids Res. 28(1): 235-242. DOI: 10.1093/nar/28.1.235
53. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T and et al.(1992).The Nucleic Acid Database. A comprehensive relational database of three dimensionalstructures of nucleic acids. Biophysics Journal 63(3): 751-759. DOI: 10.1016/S0006-3495(92)81649-1