Current trends in Bioinformatics: An Insight

[DOWNLOAD FULL EBOOK PDF]

*EditorsGulshan Wadhwa, P. Shanmughavel, Atul Kumar Singh and Jayesh R. BellareCurrent trends in Bioinformatics: An Insight*EditorsGulshanÂ Wadhwa
Department of Biotechnology Apex Bioinformatics Centre, Ministry of Science & Technology, New Delhi, IndiaP.Â Shanmughavel
Department of Bioinformatics, Bharathiar University, Coimbatore, Tamil Nadu, India AtulÂ KumarÂ Singh
Central Research Facility, Indian Institute of Technology Delhi, New Delhi, India
Centre for Research in Nanotechnology and Sciences, Indian Institute of Technology Bombay, Mumbai, IndiaJayeshÂ R.Â Bellare
Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India ISBN 978-981-10-7481-3e-ISBN 978-981-10-7483-7

ForewordIt gives us immense pleasure to present this edited book to the biotechnological research communities. Bioinformatics and computational biology â€" the science of using biological data to develop algorithms and relations among various biological systems â€" are the cutting edge areas for research. Computational sciences have their roots in the development of increasingly powerful computers over the last few decades. Rather rapidly, the instrumentation and the newly developed methodology with the underlying algorithms became widely appreciated and used as novel research strategies serving in many different fields of academic investigation, particularly in natural, engineering, social sciences, and humanities. Computational sciences have been recognized for their invaluable contributions to data collection, storage, handling, and analysis, thus leading

to efficient strategies of modeling, prediction, and design of molecular structures and of their functional properties that are often of immediate relevance for the medical sciences. Computational comparisons of DNA sequences from different organisms provide invaluable insights into past evolutionary developments, and this has become a powerful new tool in the systematics of living organisms.The growth in high-throughput full genomic sequencing, structural genomics, proteomics, epigenetics, etc., would be rather limited without bioinformatics. In order to give concise information on basic concept and advances in bioinformatics, authors have thought of bringing out an edited volume "Current Trends in Bioinformatics" for the benefit of the students and researchers working in the field of life science, medicine, and pharmaceutical science. It also focuses on reviews on advances in computational molecular/structural biology, encompassing areas such as computing in biomedicine and genomics, computational proteomics and systems biology, and metabolic pathway engineering. Developments in these fields have direct impact on key issues related to health care, medicine, genetic disorders, development of agricultural products, renewable energy, environmental protection, etc. The book has 18 chapters, divided into two sections.The overview of important aspects of bioinformatics would further contribute to strengthen international contacts and serve as a testament to such a fruitful development for the basic as well as applied sciences. The Department of Biotechnology considering the great significance of this field established a countrywide network of bioinformatics centers in academic institutions. These have paid rich dividends. We hope that scientific community especially students, in particular, would enjoy reading, learn and make best use of this book.   (Dr. Manju Sharma) Manju Sharma Preface Bioinformatics has become a frontline applied science and is of vital importance to study new biology, which is widely recognized as the new scientific endeavor of the twenty-first century. The growth in full genomic sequencing, structural genomics, proteomics, and microarray will be very slow without application of bioinformatics. In fact the very high importance of bioinformatics comes from its usefulness in these areas to solve complex biological problems. So up-to-date information in the field of bioinformatics is the most needed one. The proposed book *Current Trends in Bioinformatics* fulfills these requirements. *Current Trends in Bioinformatics* aims to publish all the latest and outstanding developments in bioinformatics. The book contains a series of timely, in-depth reviews, drug clinical trial studies, and biodiversity informatics and thematic issues written by leaders in the field, covering a wide range of the integration of biology with computer and information science. It also focuses on reviews on advances in computational molecular/structural biology, encompassing areas such as computing in biomedicine and genomics, computational proteomics and systems biology, and metabolic pathway engineering. Developments in these fields have direct implications on key issues related to health care, medicine, genetic disorders, development of agricultural products, renewable energy, and environmental protection.This book is an ideal foundation for teaching at the undergraduate and graduate levels. It is also highly suited for self-instruction by research investigators interested in applying bioinformatics methods of analysis and information technologists associated with academic and industrial laboratories.It is supposed that the nonspecialists would be the principal readers of the book. So, before embarking on the bioinformatics, some fundamental aspects of molecular evolution, taxa-related studies, some core concepts of genomics and some of the important genomic techniques were discussed in this book, to make the readers conceptualize the bioinformatics analysis.The author would also like to thank colleagues for their encouragement, enthusiasm, and support for the success of this project.Last but not the least, the author is grateful to the Staff of Springer for making this project a reality, helping to bring it to successful completion, and always being available whenever help and advice were needed.

Gulshan Wadhwa Jayesh R. Bellare P. Shanmughavel Atul Kumar Singh New Delhi, India Mumbai, India Coimbatore, India New Delhi, India Mumbai, India Contents Part I

ContributorsC.Â Awasthi
Department of Biotechnology, Gobind Ballabh Pant Engineering College, Pauri Garhwal, Uttarakhand, IndiaJayesh R.Â Bellare
Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, IndiaVaibhavÂ D.Â Bhatt
Department of Pharmaceutical Sciences, Saurashtra University, Rajkot, Gujarat, India ChandanÂ Chakraborty
School of Medical Science & Technology, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, IndiaDevÂ KumarÂ Das
School of Medical Science & Technology, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, IndiaAbhijitÂ Datta
Department of Botany, Jhargram Raj College, Medinipur, West Bengal, IndiaRajatÂ K.Â De
Machine Intelligence Unit, Indian Statistical Institute, Kolkata, IndiaRohitÂ Farmer
School of Biosciences, University of Birmingham, Birmingham, UK
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaJ.Â Fermin Angelo Selvin
Department of Physics, Nadar Mahajana Sangam S. Vellaichamy Nadar College, Madurai, Tamil Nadu, IndiaSayakÂ Ganguli
Theoretical and Computational Biology Division, Amplicon Institute of Interdisciplinary Science and Technology, Palta, West Bengal, IndiaBudhayashÂ Gautam
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture,

Technology and Sciences, Allahabad, Uttar Pradesh, IndiaKavitaÂ Goswami
Plant RNAi Biology Group, International Center for Genetic Engineering and Biotechnology, New Delhi, IndiaT.Â Hema Thanka Christlet
Department of Physics, Dr. Ambedkar Government Arts College, Chennai, Tamil Nadu, IndiaChakreshÂ KumarÂ Jain
Department of Biotechnology, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, IndiaYogeshÂ K.Â Jaiswal
School of Studies in Biochemistry, Jiwaji University, Gwalior, India
QaziÂ MohammadÂ SajidÂ Jamal
Department of Health Information Management, College of Applied Medical Sciences, East Qassim University, Al Qassim-Buraydah, Kingdom of Saudi ArabiaA.Â Jasmine
Department of Physics, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, IndiaChaitanyaÂ G.Â Joshi
Department of Animal Biotechnology, College of Veterinary Science and Animal Husbandry, Anand Agricultural University, Anand, Gujarat, IndiaSukhdeepÂ Kaur
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaKavindraÂ KumarÂ Kesari
Department of Applied Physics, and Department of Bioproduct & Biosystem, Aalto University, Espoo, FinlandK.Â Konda Reddy
Department of Pharmacy, National University of Singapore, Singapore, Singapore
AjitÂ Kumar
Centre for Bioinformatics, Maharshi Dayanand University, Rohtak, IndiaSunilÂ Kumar
Bioinformatics Centre, Institute of Life Sciences, Bhubaneswar, Odisha, India
ICAR-NBAIM, Mau, Uttar Pradesh, IndiaSharadÂ SinghÂ Lodhi
School of Studies in Biochemistry, Jiwaji University, Gwalior, IndiaAshokÂ K.Â Maiti
Medipath Clinic (P) Ltd, West Medinipur, West Bengal, IndiaJayakanthanÂ Mannu
Department of Plant Molecular Biology and Bioinformatics, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, IndiaSomaÂ S.Â Marla
Indian Council of Agricultural Research, National Bureau of Plant Genetic Resources, New Delhi, IndiaPremenduÂ P.Â Mathur
Department of Biochemistry and Molecular Biology, School of Life Sciences, Pondicherry University, Pondicherry, IndiaRanjeetÂ Maurya
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaNeelofarÂ Mirza
Indian Council of Agricultural Research, National Bureau of Plant Genetic Resources, New Delhi, IndiaNeetiÂ SananÂ Mishra
Plant RNAi Biology Group, International Center for Genetic Engineering and Biotechnology, New Delhi, IndiaPallaviÂ Mishra
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaRashmiÂ Mukherjee
RNLKWC, Vidyasagar University, Midnapur, West Bengal, IndiaK.Â D.Â Nadella
Directorate of Knowledge Management Units (DKMU), ICAR, New Delhi, India
Genetics division, ICAR, IARI, New Delhi, IndiaH.Â P.Â Pandey
Department of Biochemistry, Nepalgunj Medical College, Chisapani Campus, Kathmandu University, Nepalgunj, NepalMonikaÂ Patel
Department of Animal Biotechnology, College of Veterinary Science and Animal Husbandry, Anand Agricultural University, Anand, Gujarat, IndiaAnjaliÂ Rao
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture,

Technology and Sciences, Allahabad, Uttar Pradesh, IndiaP. Shanmughavel
Department of Bioinformatics, Bharathiar University, Coimbatore, Tamil Nadu, India
Mohd. Haris Siddiqui
Department of Bioengineering, Faculty of Engineering, Integral University, Lucknow, Uttar Pradesh, IndiaAtul Kumar Singh
Central Research Facility, Indian Institute of Technology Delhi, New Delhi, India
Centre for Research in Nanotechnology and Sciences, Indian Institute of Technology Bombay, Mumbai, IndiaSatendra Singh
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaManish Sinha
Laureate Institute of Pharmacy, Kangra, Himachal Pradesh, IndiaGitanjali Tandon
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaChristopher M. Thomas
School of Biosciences, University of Birmingham, Birmingham, UKNamrata Tomar
Department of BioMedical Engineering, Medical College of Wisconsin, Milwaukee, WI, USA
Machine Intelligence Unit, Indian Statistical Institute, Kolkata, IndiaMobashar Hussain Urf Turabe Fazil
Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, SingaporeK. Veluraja
Department of Physics, School of Advanced Sciences, VIT University, Vellore, Tamil Nadu, IndiaJyoti Verma
Department of Biotechnology, Gobind Ballabh Pant Engineering College, Pauri Garhwal, Uttarakhand, IndiaGulshan Wadhwa
Department of Biotechnology, Apex Bioinformatics Centre, Ministry of Science & Technology, New Delhi, IndiaPeter J. Winn
School of Biosciences, University of Birmingham, Birmingham, UKArvind Kumar Yadav
Department of Computational Biology and Bioinformatics, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh, IndiaAbout the Editors

Gulshan Wadhwa

is currently the Joint Director in the Department of Biotechnology, Ministry of Science & Technology, Government of India. He has developed the Bioinformatics program in India (BTISnet) and established teaching and training programs and a super computing facility in Bioinformatics in India. He has undergone trainings from WIPO, Geneva, and NUS, Singapore. Jayesh R. Bellare

is currently Institute Chair Professor and Professor with Chemical Engineering Department & Center for Research in NanoTechnology and Science, Indian Institute of Technology-Bombay, Mumbai, India. He was a Post Doctoral Fellow, from M.I.T., Cambridge, USA, and has over 33 years of experience. P. Shanmughavel

 is currently working as Associate Professor at Bharathiar University, Coimbatore, India. He has published 7 books in bioinformatics (two books published in Germany), 6 book chapters, 5 conference proceedings, and 31 research papers in reputed national and international journals in Bioinformatics. He has established the Centre of Bioinformatics-â€œDBT-BIFâ€• in 2006. Atul Kumar Singh

is currently working as Senior Scientist at Indian Institute of Technology-Delhi. He has completed his Doctorate from Indian Institute of Technology-Bombay, Mumbai, in Nanotechnology. He has published 11 research papers in reputed international journals and has 6 patents. Part IOverview © *Springer Nature Singapore Pte Ltd. 2018Gulshan Wadhwa, P. Shanmughavel, Atul Kumar Singh and Jayesh R. Bellare (eds.)Current trends in Bioinformatics: An Insight*https://doi.org/10.1007/978-981-10-7483-7_1**1. An Insight of Biological Databases Used in Bioinformatics**Vaibhav D. Bhatt[1] , Monika Patel[2] and Chaitanya G. Joshi[2](1)

Department of Pharmaceutical Sciences, Saurashtra University, Rajkot, Gujarat, India(2) Department of Animal Biotechnology, College of Veterinary Science and Animal Husbandry, Anand Agricultural University, Anand, Gujarat, India **VaibhavÂ D.Â BhattAbstract**Collections of life sciences information from scientific investigations, high-throughput experiment technology, available literature, and computational analysis are called biological databases. It contains information from research areas comprising genomics, microarray gene expression, proteomics, phylogenetics, metabolomics, gene function, structure, localization and similarities of biological sequences. In a nutshell, databases are libraries for storage and representation of biological data obtained from the scientific community which converts data into knowledge. Utmost biological databases are available from websites that categorize data which operators can browse through the data online. Due to the vast amount of data generated by high-throughput DNA sequencers in the investigation of genome, transcriptome, and exome sequences of various organisms in current times, the biological data has stored with an exponential rate. The availability of enormous amount of biological data (sequences as well as structural) has generated a need for managing, storing, and retrieving this huge data. This chapter reviews current knowledge of the different types of databases available with examples of their file formats.**Keywords**Biological sequencesHigh-throughput DNA sequencersTranscriptome and exome sequences**1.1 IntroductionDatabases are the convenient system to properly store, search, and recover several types of data. A database helps to easily handle and share large amount of data and supports large-scale analysis by easy access and data update (Liu and Ã–zsu 2009).Due to the vast amount of data generated in experiments of genome, transcriptome, and exome sequences of various organisms in current times, the biological data has stored with an exponential rate. The availability of enormous amount of biological data (sequences as well as structural data) has generated a need for managing, storing, and retrieving this huge data.Therefore the biological databases have come into existence as invaluable sources for the biological community. In a nutshell, databases are libraries for storage and representation of biological data obtained from the scientific community which converts data into knowledge.1.2 HistoryA book published in 1965, *Atlas of Protein Sequences and Structures*, *was the first biological database* by Margaret Dayhoff and colleagues, and further they have published other editions of the book in the 1970s; however the first edition was limited to 65 sequences only (Dayhoff and Foundation 1973, 1976; Foundation 1972).With the discovery of the integrated circuit, the powerful and reliable third generationÂ computers are became the choice of storage of biological databases for scientists. An English scientist Tim Berners-Lee in 1989 invented the â€œWorld Wide Webâ€• (WWW) which is the primary tool people use to interact on the Internet and is the way to access all biological databases. Production of high throughput sequencing machines leads production of data rich science, needs an interdisciplinary arena to develop software tools which is used to understand biological data. The field of science with the involvement of computer, statistics and engineering to study biological data is called Bioinformatics.1.3 Classification of Biological Databases 1.3.1 Databases Based on Data TypesThis database was divided into several databases; some of the databases were discussed below in detail.1.3.1.1 Sequence DatabasesSequence databases contain both nucleic acid and protein sequences. First we will discuss about nucleotide sequence repositories.(I)Nucleic Acid Sequence Database There are three main nucleotide sequence repositories:(A)GenBank (B)European Molecular Biology Laboratory (EMBL) (C)DNA Data Bank of Japan (DDBJ) Raw nucleic acid sequences are stored in these databases and make available through Internet sources. Initially, these databases worked independently, but later the *International Nucleotide Sequence Database Collaboration* (INSDC, http://â€‹insdc.â€‹org) was developed to maintain collaboration between DDBJ, GenBank, and EMBL (Fig. 1.1). These databases started exchanging their data through constant communication between the team at each collaborating organizations in order to access the sequences present in all three different formats.*Fig. 1.1*The home page of International Nucleotide Sequence Database Collaboration (INSDC) (**

http://insdc.org) *(A) GenBank* GenBank is a collection of raw and annotated nucleotide as well as protein information. GenBank is maintained and accessed through the National Center for Biotechnology Information (NCBI). Every 2 months a new release is made. It is maintained by NCBI as part of the INSDC (Benton [1990](#)). There are approximately 137384889783 bases, from 149819246 sequence records in the GenBank release 188.0 on February 15, 2012. Type "insulin" in the search tab on the GenBank home page to view list of sequences of insulin gene, partial or complete from different organisms (Fig. [1.2](#)).*Fig. 1.2*Using GenBank to query insulin sequences (http://www.ncbi.nlm.nih.gov/nuccore/?term=insulin)
Example of GenBank Format   Format Explanation GenBank format includes *locus name* which is similar to the accession number and unique to the entry, and it is followed by sequence length. In our example sequence length is 587 bp. Definition includes description of source organism, gene/protein name, and other details about sequence. *Accession number* is the unique identifier of the sequence (NM_013564).*Version* is similar to accession number, but whenever a change occurs in sequence data, the version increases by 1. In our example, version is NM_013564.7; this indicates that sequence has been changed seven times.*GI (GenInfo Identifier)* number also runs parallel to the accession number and version system. A new GI is allotted, if the sequence has been changed and the version has increased by unity. In our example, GI is 365192585.*Keywords* are words or expressions about sequence. The keyword field contains a dot if nothing is provided.*Source* contains name of the organism from which the sequence has been derived.*Organism* is a related sub-keyword of source and contains the scientific name of the organism along with the lineage as described in NCBI taxonomy database.*Reference* contains the publication by the authors of the sequence.*Authors* contain list of authors in the same order as appears in publication.*Title* shows the title of published/unpublished work.*Journal* contains MEDLINE abbreviations of the journal name where the work is published.*PubMed* field provides the PubMed identifier (PMID) of that article.*Comment* points out the change occurred in the submitted sequence.*Features* provide information about genes and their products, segment of biological significance in the submitted sequence, as well as other characteristics.*Gene* provides gene length and gene name and its function and synonyms. CDS represents coding sequence which codes for protein sequence.*Origin* contains the sequence data. Finally, GenBank record ends with // sign. *Sequence Submission to GenBank* Sequence submission is done by using different tools available at NCBI. Few of them are: *BankIt*: direct submissions are made to GenBank using it (www.ncbi.nlm.nih.gov/WebSub/?tool=genbank).*Sequin*: it is a stand-alone submission platform (www.ncbi.nlm.nih.gov/Sequin/). *tbl2asn*: it is a command-line program, used for submission of large batches of sequences and complete genomes (www.ncbi.nlm.nih.gov/genbank/tbl2asn2). *Barcode Submission Tool*: it is a WWW-based tool for the submission of sequences and trace read data (http://www.ncbi.nlm.nih.gov/WebSub/?tool=barcode).
National Center for Biotechnology Information (NCBI) NCBI was started in 1988, as a part of the US National Library of Medicine (NLM) located at Bethesda, Maryland. It is a division of the National Institutes of Health and is directed by David Lipman. The responsibility of NCBI is to make available the GenBank nucleotide sequence database since 1992. NCBI is playing a very remarkable role for biological scientists by making available various public databases and software tools for sequence analysis (Table [1.1](#)). GenBank manages with individual laboratories and other sequence databases like those of the EMBL and the DDBJ. Meanwhile in 1992, NCBI has developed to run other databases in addition to GenBank ((US) [2013](#)). The home page of NCBI is shown in Fig. [1.3](#). *Databases and Tools of NCBI Table 1.1*Various databases and software tools of NCBI for sequence analysis

NCBI

ToolsDatabases

Sequence SubmissionSequenceData miningLiterature
AnalysisNucleotide
SequinBLASTEntrezProtein
BankItBlinkMy NCBIStructure
tbl2asnStand-alone BLASTLinkOutGenome
OMIM
SNP
Barcode Submission Toole-PCRCitationBooks
MatcherDomain
ORF FinderÂ Chemical
Expression
Map viewerOther databases
Tax plot
Trace archive*Fig. 1.3*The home page of National Center for Biotechnology Information (http://â€‹www.â€‹ncbi.â€‹nlm.â€‹nih.â€‹gov/â€‹) *Database Retrieval Tool Entrez* (www.â€‹ncbi.â€‹nlm.â€‹nih.â€‹gov/â€‹Entrez/â€‹) in Fig. [1.4](1.4) is a primary text search engine which comprises of 40 molecular and literature databases. It extracts huge information from the PubMed database, such as DNA and protein sequences and structure, gene, genome, genetic variation, and gene expression.*Fig. 1.4*The home page of Entrez (www.â€‹ncbi.â€‹nlm.â€‹nih.â€‹gov/â€‹Entrez/â€‹) *(B) European Molecular Biology Laboratory (EMBL)*   The *European Molecular Biology Laboratory* (*EMBL*) (http://â€‹www.â€‹embl.â€‹org/â€‹) in Fig. [1.5](1.5) is a molecular biology organization which is maintained by 20 European countries, with Australia as associate member state. It is an intergovernmental organization created in 1974. It develops and maintains a large number of databases, and scientists can access the data free of cost. This research laboratory functions from five different locations, the main laboratory, the European Bioinformatics Institute (EBI), Heidelberg, Germany, is a hub for bioinformatics research and services, directed by Dr. Rolf Apweiler and Dr. Ewan Birney. It is a part of INSDC, which includes DDBJ and GenBank. Typing insulin gene at EMBL search engine produced a result in Fig. [1.6](1.6).*Fig. 1.5*The home page of European molecular biology laboratory (http://â€‹www.â€‹embl.â€‹org/â€‹)*Fig. 1.6*Insulin gene search at European molecular biology laboratory website (https://â€‹www.â€‹ebi.â€‹ac.â€‹uk/â€‹ebisearch/â€‹search.â€‹ebi?â€‹query=â€‹insulin&â€‹db=â€‹allebi&â€‹requestFrom=â€‹searchBox) *EMBL File Format   Sequence Retrieval System (SRS) SRS* (http://â€‹srs.â€‹ebi.â€‹ac.â€‹uk/â€‹) (Fig. [1.7](1.7)) is a powerful searching tool to retrieve sequences (and other types of data) and also to perform various operations on retrieved information for EMBL. It is similar to Entrez of NCBI, a search engine for extracting all sort of information available at EMBL.*Fig. 1.7*The home page of Sequence Retrieval System (http://â€‹srs.â€‹ebi.â€‹ac.â€‹uk/â€‹) *Sequence Submission at EMBL* There are mainly three tools available for submitting data at EMBL.*1.Webin: for nucleotide sequence submission 2.Sequin: a stand-alone tool for submitting nucleotide sequences to GenBank, EMBL, and DDBJ developed by NCBI 3.Webin-Align: a tool for sequence alignment submission (C) DNA Data Bank of Japan (DDBJ)* DDBJ, (http://â€‹ddbj.â€‹sakura.â€‹ne.â€‹jp/â€‹) (Fig. [1.8](1.8)) part of *INSDC*, was established at the National Institute of Genetics (NIG), Japan, in 1986 with the support of the Ministry of Education, Culture, Sports, Science and Technology, Japan.*Fig. 1.8*The home page of DNA Data Bank of Japan (http://â€‹ddbj.â€‹sakura.â€‹ne.â€‹jp/â€‹) *SAKURA* SAKURA (http://â€‹sakura.â€‹ddbj.â€‹nig.â€‹ac.â€‹jp/â€‹top-e.â€‹html) is a source for data (nucleotide sequence) submission system through the WWW-based server where one can enter and submit nucleotide sequences and translated amino acid sequences. Since 1995 it is open to the public and scientists community. *DDBJ Format (II)Protein Sequence Databases* The different protein sequence databases available are the following:*(A)Protein Information Resource (B)UniProt (A) Protein Information Resource (PIR)* Margaret Dayhoff was the

inventor of Protein Information Resource (PIR) in the 1960s at the National Biomedical Research Foundation (NBRF) for investigation of evolutionary relationships among proteins. Analysis tools for protein database are provided by PIR which are freely available to the scientists (George et al. 1997).In 2002 Protein Information Resource and its worldwide partners, EBI and Swiss Institute of Bioinformatics (SIB), were granted an award from the National Institutes of Health (NIH) to make UniProt, by merging the databases of PIR-PSD, SWISS-PROT, and TrEMBL (Fig. 1.9).*Fig. 1.9*The home page of Protein Information Resource (http://pir.georgetown.edu/) *(B)UniProt* It comprises of two sections:*(a)SWISS-PROT (b)Translated EMBL (TrEMBL) (a) SWISS-PROT* SWISS-PROT (http://www.uniprot.org/) (Fig. 1.10), established in 1896, is the most widely used protein sequence database created by the University of Geneva and the EMBL, collaboratively. After 1994, the collaboration moved to EMBL's UK outstation, the EBI. *SWISS-PROT Format Fig. 1.10*The home page of UniProt (http://www.uniprot.org/)Each line starts with a two-character line code, which specifies the kind of data contained in the line. *(b) Translated EMBL* TrEMBL benefits from the SWISS-PROT format and comprises translations of all coding sequences (CDS) in EMBL. It has two core divisions, designated SWISS-PROT-TrEMBL and REM-TrEMBL.1.3.1.2 Structure Databases PDB (Protein Data Bank)MMDB (Molecular Modeling Database)VAST (Vector Alignment Search Tool)CDD (Conserved Domain Database)NDB (Nucleic acid Structure Database) From the above databases, some of the database is shown below in detail.*(I)Protein Data Bank (PDB)* The PDB (http://www.rcsb.org/pdb/home/home.do) in Fig. 1.11, a source for the three-dimensional structural data of huge biological molecules, includes proteins and nucleic acids. It was established in 1971 by the Research Collaborators for Structural Bioinformatics (RCSB). The data submitted by scientists from different parts of the world are easily without cost available through the Internet. The PDB is supervised by the Worldwide Protein Data Bank (wwPDB) (Berman 2008).*Fig. 1.11*The home page of PDB with the query Hemoglobin (http://www.rcsb.org/pdb/home/home.do)As on March 20, 2012 at 5 PM PDT, there were 80,264 structures. Each structure has been assigned a PDB ID, which contains four characters both alphabets and numerical. The first character is a numeral, while the last three characters can be either numerals or letters. Search results and structure for hemoglobin were showed in Figs. 1.11 and 1.12.*Fig. 1.12* Search result of Protein Data Bank (http://www.rcsb.org/pdb/results/results.do?qrid=57082E24&tabtoshow=Current) *PDB File Format* This format was primarily practiced by the Protein data bank and previously was known as the PDB file format. The PDB also retains data on biological macromolecules, "macromolecular crystallographic information file format" (mmCIF), initiated to be phased in 1996. In the year 2005, an Extensible Markup Language (XML) version of PDBML was described (Westbrook et al. 2005). *Data Deposition Tool of PDB* Auto Dep Input Tool (ADIT) (http://deposit.rcsb.org/adit/) (Fig. 1.13) is developed by RCSB, and it is responsible for depositing structures to PDB in an efficient manner.*Fig. 1.13*The home page of Auto Dep Input Tool (http://deposit.rcsb.org/adit/) *(II) Nucleic Acid Structure Database (NDB)* This database (http://ndbserver.rutgers.edu/) (Fig. 1.14) provides us 3D structures of nucleic acids.*Fig. 1.14*The home page of nucleic acid structure database (http://ndbserver.rutgers.edu/)1.3.1.3 Literature DatabaseLiterature databases provide us library of life science work done all over the world. Various literature databases available are the following:MEDLINECiteXploreOMIMPatent abstractsFlyBase archives1.3.1.4 Pathway DatabaseTo comprehend molecular interactions and chemical reaction networks, the pathway database is used by pathway maps. Various pathway databases available are the following:BioCyc database collection comprising EcoCyc and MetaCycKEGG PATHWAY Database (www.genome.jp/kegg/)MANET databaseReactome (Laboratory of Cold Spring Harbor, EBI, Gene Ontology Consortium)

1.3.1.5 Chemical Database A collection of the chemical information precisely planned is called chemical database. These are the few freely available chemical databases:Chemical Entities of Biological Interest (ChEBI)PubChemZincMoleculesDrugBank**1.3.1.6 Enzyme Database** Enzyme databases cover an extensive range of properties and functions, such as structure, occurrence, kinetics of enzyme-catalyzed reactions, and metabolic function. Various enzyme databases available are the following:ExPASyBRENDAREBASEEC enzyme database**1.3.1.7 Disease Database** The disease database provides all disease-related information; it is a cross-referenced index of diseases, symptoms, medications, signs, abnormal investigation findings, etc.OMIMOMIA**1.3.1.8 Domain Database** Domain database is a database for ancient domains and full-length proteins.CDD (Conserved Domain Database)**1.3.1.9 Structural Classification of Protein Database** It provides hierarchical classification of protein structure which defines the evolutionary association between proteins.The Structural Classification of Proteins (SCOP) (http://â€‹scop.â€‹mrclmb.â€‹cam.â€‹ac.â€‹uk/â€‹scop/â€‹).Class, architecture, topology, and homologous superfamily (CATH) is freely available to scientists (www.â€‹cathdb.â€‹info/â€‹).**1.3.1.10 Genome Database** Genome databases are a collection of genome sequences of many species; it interprets and examines them and provides free public access.Genome Databases at the National Center for Biotechnology Information (Index)Genome Databases at the National Center for Biotechnology Information (Entrez)Genome Databases at the National Center for Biotechnology Information (PMGif) Genome List in NIHMitochondrial DNA Database (MitBASE)Mouse Genome InformaticsPlant Genome Project maintained by the National Science FoundationOrganelle Genome Sequences (PMGif)**1.3.2 Biological Databases Based on Database Source** This database is subdivided into two databases, primary and secondary.*1. Primary: databases comprising of data generated experimentally like nucleotide sequences and 3D structures are identified as primary databases.* Examples are GenBank, DDBJ, EMBL, PIR, PDB, NDB, UniProt, TrEMBL, SWISS-PROT, etc.*2.Secondary: it contains databases directly derived from the primary databases.* Examples are PROSITE, Pfam, Blocks, Prints, SCOP, CATH, OMIM, KEGG, etc.**1.3.3 Composite Databases** It combines various different primary database sources. This makes searching the query more efficient. So, composite database amalgamates various primary databases for easy access.Examples are OWL, NRDB, MIPSX, SP, and TrEMBL.**1.3.4 Biological Databases Based on Database Design** This database is subdivided into two databases, object-oriented and relational databases.**1.3.4.1 Object Oriented**

This bookÂ highlights the latest breakthrough developments in bioinformatics. It presents a series of timely, in-depth reviews, drug clinical trial studies, biodiversity informatics and thematic issues. In addition, it includes insightful reviews on advances in computational molecular/structural biology, which address areas such as computing in biomedicine and genomics, computational proteomics and systems biology, and metabolic pathway engineering. Innovations in these fields have direct impacts on key issues related to healthcare, medicine, genetic disorders, the development of agricultural products, renewable energy, and environmental protection.

Written by respected leaders in the field and covering a wide range of topics involving the integration of biology with computer and information science, the book offers an ideal basis for teaching at the undergraduate and graduate levels. It can also be used

for self-instruction by research investigators interested in applying bioinformatics-based analytical methods and information technologists working with academic and industrial laboratories.

Dr. Soma Sundar Marla - Scientists - It addresses modern trends such as pharmacogenomics, evaluation of gene The book highlights the practical utility of biotechnology and bioinformatics for Biotechnology and Bioinformatics: Advances and Applications - Current Pharmaceutical Design 2010;16:204â€"19. Drews J. Strategic trends in the drug industry. Copy number variation: new insights in genome diversity. Li J, Robson B. Bioinformatics and computational chemistry in molecular design. Ebook 2019 Current Trends In Bioinformatics An Insight at - Editorial Reviews. From the Back Cover. This book highlights the latest breakthrough developments in bioinformatics. It presents a series of timely, in-depth Actionable Insight on Families Affected by Incarceration - This book highlights the latest breakthrough developments in bioinformatics. It presents a series of timely, in-depth reviews, drug clinical trial studies, biodiversity Bioinformatics pdf drive - Dive Center - This will not only contribute to basic knowledge but also provide new insights into the This comprehensive database covers 689 known and 1664 predicted novel miRs Singh A., Bellare J. (eds) Current trends in Bioinformatics: An Insight. Xilinx Dsp Book - Lebenszeichen fÃ¼r die Ewigkeit - Current Trends and Applications Tomasz G. Smolinski, Mariofanna G. The objective of this book chapter is to present to the computational intelli- gence and To provide useful insights for CI applications in bioinformatics, we structure the Advances in Microbial Biotechnology - Apple Academic Press - Metagenomics is the study of genetic material recovered directly from environmental samples. Recent studies use either "shotgun" or PCR directed sequencing to get The insights gained from these breakthrough studies led Pace to propose the Advances in bioinformatics, refinements of DNA amplification, and the Retrospective analysis of natural products provides insights - My Link-Tips: DSP with FPGAs book home dsp4fpgas. can u suggest me what (NASDAQ: XLNX) today introduced a new generation of its flagship Virtex(R) high.. The increasing trend towards high performance and low power systems has tool to explore traditional DSP topics and solve problems to gain insight. Current trends in Bioinformatics - ACM

Digital Library - Editorial Reviews. From the Back Cover. This book highlights the latest breakthrough developments in bioinformatics. It presents a series of timely, in-depth Actionable Insight on Families Affected by Incarceration - Fox News promoted an excerpt of the book, initially claiming the Obama White Tri-Conference attendees gained insight and knowledge by attending the event,. Getting a cross-section of the current trends, the talks, the vendors, the QIAGEN had the pleasure of taking part in an NGS Bioinformatics Challenge run by Social Awareness: Definition, Example & Theories - Video - List of articles published in current issue of Trends in Bioinformatics.

Relevant Books

[DOWNLOAD] - Ebook Accidental Warrior online

[DOWNLOAD] - Download Free Commander in Chief (By Design Book 11) free epub, pdf online

[DOWNLOAD] - Download The Remembrance pdf

[DOWNLOAD] - Download I Won Something Cool The Other Day: A short horror story free

[DOWNLOAD] - Download book Unleashing Your Entrepreneurial Potential (Response Books)