



Structural and Functional Bioinformatics

Nida Tabassum Khan*

Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology Engineering and Management Sciences,(BUIITEMS),Quetta, Pakistan

*Corresponding author: Nida Tabassum Khan, Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology Engineering and Management Sciences, (BUIITEMS), Quetta, Pakistan, Tel: 03368164903; E-mail: nidatabassumkhan@yahoo.com

Received date: January 26, 2018 Accepted date: February 22, 2018 Published date: February 26, 2018

Citation: Khan, N.T. Structural and Functional Bioinformatics. (2018) Lett Health Biol Sci 3(1): 7- 11.

Abstract:

Structural and functional bioinformatics help us to design and formulate prognostic computational models and frameworks that exploit our growing knowledge of biological macromolecules in terms of their structural organization and functional capabilities. Integration of structural and functional biochemistry of macromolecules with informatics empowers significant progress in understanding the fundamentals of biology.

Keywords: Cath; Scop; Haddock; Prosite; Vanted; Brenda

Introduction

Structural and functional bioinformatics aimed to unravel biological problems by analyzing sequences of biological molecules such as DNA and protein using computational algorithms, informatics tools and software's to assess molecular data^[1]. Some of the applications of this novel field are given below:

Prediction of protein structure: Understanding the correlation between amino acid sequence and the three dimensional structure of protein, it can be helpful for determining protein structure from its amino acid sequence^[2]. Numerous bioinformatics tools could be utilized for protein structure and function prediction including secondary structure prediction^[3], homology modeling^[4], protein threading^[5], ab initio methods^[6], prediction of motif^[7], domain^[8], transmembrane helix^[9], signal peptide^[10] etc. Some of these tools and databases used are summarized below in Table 1.

Table 1: Protein structure/function prediction software's and databases.

S.no	Tools and Databases	Purpose
1	PDB (Protein Data Bank)	Universal storage of 3D structure data of macromolecules. Providing methods for visualizing the structure and downloading structural information ^[11]
2	MMDB (NCBI Structure Database)	Includes database of 3D structure of biomolecules. Provides information on biological functions of proteins, on mechanisms related to their functions and on relationship between biomolecules and their evolutionary history ^[12]
3	BLAST (Basic Local Alignment Search Tool)	Searches for regions of local similarity in a protein or DNA sequence ^[13]
4	SWISS PDB Viewer	Analyze structural alignment of proteins and provide comparison of their active sites, their amino acid mutations angles, distances and H bonds between their atoms ^[14]
5	PDBsum	Consists of images of structure, detailed structural analysis derived from PROMOTIF program, schematic graphs of interactions, summary PROCHEK results etc ^[15]

6	Ligplot	Determine interaction between proteins, ligands, hydrogen atoms, hydrophobic interactions etc ^[16]
7	SCOP	Database containing detailed information of protein structure and phylogenetics ^[17]
8	CATH (Class, Architecture, Topology and Homologous super family)	Database that stores hierarchical classification for domain structures of proteins ^[18]
9	Composition Profiler	Bioinformatics tool for determining amino acids enrichment or depletion based on their physicochemical or structural features ^[19]
10	Prosite/ Interpro	Determines protein families, conserved domains, motifs etc ^[20]
11	Pfam	Determines protein families ^[21]
12	SWISS PROT	Database of annotated protein sequences ^[22]
13	PRIDE (Proteomics Identifications Database)	Contain information on functional characterization and post-translation modification of proteins and peptides ^[23]
14	SGMP (Signaling Gateway Molecule Pages)	Database contains information on functional states of proteins involved in signal transduction pathways ^[24]
15	ProtParam	Tool to find proteins physico-chemical properties ^[25]
18	SMART(Simple Modular Architecture Retrieval Tool)	Defines multiple information about the query protein ^[26]
19	Auto Dock	Determines protein-ligand interaction ^[27]
20	HADDOCK	Determines modeling and bio-molecular ^[28]
21	BIND	Database that provides information on molecular interaction of biological molecules ^[29]
22	APSSP2	Determines proteins secondary structure ^[30]
23	MODELLER	Determines 3D structure of protein ^[31]
24	Phyre and Phyre2	Find protein structure ^[32]

Few bioinformatics tools and databases for functional analysis of large gene are summarized below in Table 2.

Table 2: Tools for structural and functional analysis of gene.

S.no	Tools and Databases	Purpose
1	Gene Ontology	Systematically dissect large gene lists ^[33]
2	Onto-Express, MAP Finder, Go Miner, DAVID, EASE, Gene MergeFuncAssociate etc	Analysis of gene-annotation enrichment ^[34-40]
3	REPAIRtoire(Examples Repair GENES, Human DNA Repair Genes, Repair-Fun Map and GeneSNPs etc)	Contains data on all DNA repair systems and proteins from model organisms ^[41,42]
4	BRENDA (BRAunschweig Enzyme Database)	Database contains information on enzymes properties etc ^[43]
5	Pathway Commons	Contains data on biological pathways including macromolecule interactions, biochemical reactions, complex assembly, transport, catalysis events, etc ^[44]
6	OriDB (DNA replication), Data base (Replication Domain, apoptosis), Telomerase database (telomere maintenance), REBASE (DNA restriction and modification), and DANcER (epigenetic/chromatin modification)	Databases contains data on DNA metabolism ^[45-49]
7	JIGSAW	Determine genes, splicing sites etc ^[50]
8	novoSNP	Find point mutation in DNA sequence ^[51]
9	PPP(Prokaryotic promoter prediction)	Tool to determine promoter in a gene ^[52]
10	WebGeSTer	Database to find the termination sites during transcription in the genes ^[53]
11	Genscan	Determine exon-intron sites in sequences ^[54]

Few bioinformatics tools and databases for analysis of lipids are summarized below in Table 3.

Table 3: Tools for analysis of lipids.

S.no	Tools and Databases	Purpose
1	magnet	Software enables retrieval and visualization of biological relationships across heterogeneous data sources from an integrated database[55]
2	VANTED	Enable importing and customization of KEGG lipid-specific pathways[56]
3	GOLD LIPID MAPS Proteome Database (LMPD), LIPID BANK, LIPIDAT and LMSD	Database of genomics of lipid-associated disorders[57-60]

Conclusion

Progress in the field of structural and functional bioinformatics includes future contribution to structural and functional understanding of the macromolecules such as DNA, proteins, lipids etc for the better apprehension of the biological processes and pathways on which the origin of life rely.

References

- Chou, K.C. Structural bioinformatics and its impact to biomedical science. (2004) *Current medicinal chemistry* 11(16): 2105-2134.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Rokde, C.N., Kshirsagar, M. Bioinformatics: Protein structure prediction. (2013) In *Computing, Communications and Networking Technologies (ICCCNT)*, 2013 Fourth International Conference on IEEE: (1-5).
[Pubmed](#) | [Crossref](#) | [Others](#)
- Barton, G.J. Protein secondary structure prediction. (1995) *Current opinion in structural biology* 5(3): 372-376.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Bordoli, L., Kiefer, F., Arnold, K., et al. Protein structure homology modeling using SWISS-MODEL workspace. (2009) *Nat Protoc* 4(1): 1-13.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Cristobal, S., Zemla, A., Fischer, D., et al. A study of quality measures for protein threading models. (2001) *BMC bioinformatics* 2(1): 5.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Chivian, D., Robertson, T., Bonneau, R., et al. Ab initio methods. (2003) *Structural Bioinformatics*: 547-557.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Rost, B., Liu, J., Nair, R., et al. Automatic prediction of protein function. (2003) *Cell Mol Life Sci* 60(12): 2637-2650.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Rost, B., Sander, C. Prediction of protein secondary structure at better than 70% accuracy. (1993) *J Mol Biol* 232(2): 584-599.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Krogh, A., Larsson, B., Von Heijne, G., et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. (2001) *J Mol Biol* 305(3): 567-580.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Emanuelsson, O., Nielsen, H., Brunak, S., et al. Predicting sub cellular localization of proteins based on their N-terminal amino acid sequence. (2000) *J Mol Biol* 300(4): 1005-1016.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Berman, H.M., Westbrook, J., Feng, Z., et al. The Protein Data Bank, 1999-. (2006) In *International Tables for Crystallography Volume F: Crystallography of biological macromolecules*: 675-684.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Chen, J., Anderson, J.B., DeWeese-Scott, C., et al. MMDB: Entrez's 3D-structure database. (2003) *Nucleic Acids Res* 31(1): 474-477.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Altschul, S.F., Gish, W., Miller, W., et al. Basic local alignment search tool. (1990) *J Mol Biol* 215(3): 403-410.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Kaplan, W., Littlejohn, T.G. Swiss-PDB viewer (deep view). (2001) *Brief Bioinform* 2(2): 195-197.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Laskowski, R.A. PDBsum: summaries and analyses of PDB structures. (2001) *Nucleic Acids Res* 29(1): 221-222.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Wallace, A.C., Laskowski, R.A., Thornton, J.M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. (1995) *Protein Eng* 8(2): 127-134.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Murzin, A.G., Brenner, S.E., Hubbard, T., et al. SCOP: a structural classification of proteins database for the investigation of sequences and structures. (1995) *J Mol Biol* 247(4): 536-540.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Orengo, C.A., Pearl, F.M., Thornton, J.M. The CATH domain structure database. (2005) *Structural Bioinformatics* 44: 249-271.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Vacic, V., Uversky, V.N., Dunker, A.K., et al. Composition Profiler: a tool for discovery and visualization of amino acid composition differences. (2007) *BMC Bioinformatics* 8: 211.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Hulo, N., Sigrist, C.J., Le Saux, V., et al. Recent improvements to the PROSITE database. (2004) *Nucleic Acids Res* 32: 134-137.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Bateman, A., Birney, E., Cerruti, L., et al. The Pfam protein families' database. (2002) *Nucleic acids research* 30(1): 276-280.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Boeckmann, B., Bairoch, A., Apweiler, R., et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. (2003) *Nucleic Acids Res* 31(1): 365-370.
[Pubmed](#) | [Crossref](#) | [Others](#)
- Martens, L., Hermjakob, H., Jones, P., et al. PRIDE: the proteomics identifications database. (2005) *Proteomics* 5(13): 3537-3545.
[Pubmed](#) | [Crossref](#) | [Others](#)

24. Dinasarapu, A.R., Saunders, B., Ozerlat, I., et al. Signaling gateway molecule pages—a data model perspective. (2011) *Bioinformatics* 27(12): 1736-1738.
[Pubmed](#) | [Crossref](#) | [Others](#)
25. Bendtsen, J.D., Jensen, L.J., Blom, N., et al. Feature-based prediction of non-classical and leaderless protein secretion. (2004) *Protein Eng Des Sel* 17(4): 349-356.
[Pubmed](#) | [Crossref](#) | [Others](#)
26. Schultz, J., Copley, R.R., Doerks, T., et al. SMART: a web-based tool for the study of genetically mobile domains. (2000) *Nucleic Acids Res* 28(1): 231-234.
[Pubmed](#) | [Crossref](#) | [Others](#)
27. Goodsell, D.S., Morris, G.M., Olson, A.J. Automated docking of flexible ligands: applications of AutoDock. (1996) *J Mol Recognit* 9(1): 1-5.
[Pubmed](#) | [Crossref](#) | [Others](#)
28. De Vries, S.J., Van Dijk, M., Bonvin, A.M. The HADDOCK web server for data-driven bimolecular docking. (2010) *Nat Protoc* 5(5): 883-897.
[Pubmed](#) | [Crossref](#) | [Others](#)
29. Bader, G.D., Betel, D., Hogue, C.W. BIND: the biomolecular interaction network database. (2003) *Nucleic acids research* 31(1): 248-250.
[Pubmed](#) | [Crossref](#) | [Others](#)
30. Zhang, H., Tang, H., Zhang, L., et al. Evaluation on prediction methods of protein secondary structure. (2003) *Computers and Applied Chemistry* 20(6): 735-740.
[Pubmed](#) | [Crossref](#) | [Others](#)
31. Webb, B., Sali, A. Protein structure modeling with MODELLER. (2014) *Protein Structure Prediction*: 1-15.
[Pubmed](#) | [Crossref](#) | [Others](#)
32. Kelley, L.A., Mezulis, S., Yates, C.M., et al. The Phyre2 web portal for protein modeling, prediction and analysis. (2015) *Nat Protoc* 10(6): 845-858.
[Pubmed](#) | [Crossref](#) | [Others](#)
33. Harris, M.A., Clark, J., Ireland, A., et al. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. (2004) *Nucleic Acids Res* 32(1): 258-261.
[Pubmed](#) | [Crossref](#) | [Others](#)
34. Khatri, P., Draghici, S., Ostermeier, G.C., et al. Profiling gene expression using onto-express. (2002) *Genomics* 79(2): 266-270.
[Pubmed](#) | [Crossref](#) | [Others](#)
35. Doniger, S.W., Salomonis, N., Dahlquist, K.D., et al. MAPP Finder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. (2003) *Genome Biol* 4(1): R7.
[Pubmed](#) | [Crossref](#) | [Others](#)
36. Zeeberg, B.R., Feng, W., Wang, G., et al. GoMiner: a resource for biological interpretation of genomic and proteomic data. (2003) *Genome Biol* 4(4): R28.
[Pubmed](#) | [Crossref](#) | [Others](#)
37. Huang da, W., Sherman, B.T., Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. (2009) *Nat Protoc* 4(1): 44-57.
[Pubmed](#) | [Crossref](#) | [Others](#)
38. Gentleman, R.C., Carey, V.J., Bates, D.M., et al. Bioconductor: open software development for computational biology and bioinformatics. (2004) *Genome Biol* 5(10): R80.
[Pubmed](#) | [Crossref](#) | [Others](#)
39. Castillo-Davis, C.I., Hartl, D.L. GeneMerge—post-genomic analysis, data mining, and hypothesis testing. (2003) *Bioinformatics* 19(7): 891-892.
[Pubmed](#) | [Crossref](#) | [Others](#)
40. Berriz, G.F., King, O.D., Bryant, B., et al. characterizing gene sets with FuncAssociate. (2003) *Bioinformatics* 19(18): 2502-2504.
[Pubmed](#) | [Crossref](#) | [Others](#)
41. Milanowska, K., Krwawicz, J., Papaj, G., et al. REPAIRtoire—a database of DNA repair pathways. (2010) *Nucleic acids res* 39(1): 788-792.
[Pubmed](#) | [Crossref](#) | [Others](#)
42. Milanowska, K., Rother, K., Bujnicki, J.M. Databases and bioinformatics tools for the study of DNA repair. (2011) *Mol Biol Int*: 475718.
[Pubmed](#) | [Crossref](#) | [Others](#)
43. Schomburg, I., Chang, A., Ebeling, C., et al. BRENDA, the enzyme database: updates and major new developments. (2004) *Nucleic Acids Res* 32(1): 431-433.
[Pubmed](#) | [Crossref](#) | [Others](#)
44. Cerami, E.G., Gross, B.E., Demir, E., et al. Pathway Commons, a web resource for biological pathway data. (2011) *Nucleic acids res* 39(1): 685-690.
[Pubmed](#) | [Crossref](#) | [Others](#)
45. Nieduszynski, C.A., Hiraga, S.I., Ak, P., et al. OriDB: a DNA replication origin database. (2006) *Nucleic Acids Res* 35(1): 40-46.
[Pubmed](#) | [Crossref](#) | [Others](#)
46. Milanowska, K., Rother, K., Bujnicki, J.M. Databases and bioinformatics tools for the study of DNA repair. (2011) *Mol Biol Int* 2011: 475718.
[Pubmed](#) | [Crossref](#) | [Others](#)
47. Podlevsky, J.D., Bley, C.J., Omana, R.V., et al. The telomerase database. (2008) *Nucleic Acids Res* 36(1): 339-343.
[Pubmed](#) | [Crossref](#) | [Others](#)
48. Roberts, R.J., Vincze, T., Posfai, J., et al. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. (2009) *Nucleic Acids Res* 38(1): 234-236.
[Pubmed](#) | [Crossref](#) | [Others](#)
49. Rother, M.B., van Attikum, H. DNA repair goes hip-hop: SMARCA and CHD chromatin remodellers join the break dance. (2017) *Philos Trans R Soc Lond B Biol Sci* 372(1731): 20160285.
[Pubmed](#) | [Crossref](#) | [Others](#)
50. Allen, J.E., Salzberg, S.L. JIGSAW: integration of multiple sources of evidence for gene prediction. (2005) *Bioinformatics* 21(18): 3596-3603.
[Pubmed](#) | [Crossref](#) | [Others](#)
51. Doran, A.G., Creevey, C.J. Snpdat: easy and rapid annotation of results from de novo snp discovery projects for model and non-model organisms. (2013) *BMC bioinformatics* 14(1): 45.
[Pubmed](#) | [Crossref](#) | [Others](#)
52. Sun, P., Ju, H., Liu, Z., et al. Bioinformatics resources and tools for conformational B-cell epitope prediction. (2013) *Comput Math methods med*: 2013: 943636.
[Pubmed](#) | [Crossref](#) | [Others](#)

53. Kumar,V., Sharma,R.M., Thakur,R.S. Big Data Analytics: Bioinformatics Perspective.(2016) *Inte Jo Inno Adv Com Sci* 5(6): 8-14.
[Pubmed](#) | [Crossref](#) | [Others](#)
54. Kruglyak, L., Nickerson, D.A. Variation is the spice of life. (2001) *Nat genet* 27(3): 234-236.
[Pubmed](#) | [Crossref](#) | [Others](#)
55. Sharman, J.L., Gerloff, D.L. MaGnET: a software tool for integrated visualization of functional genomic data relating to the malaria parasite. (2007) *BMC Systems Bio*1 (1): 32.
[Pubmed](#) | [Crossref](#) | [Others](#)
56. Junker,B.H., Klukas, C., Schreiber, F. VANTED: a system for advanced data analysis and visualization in the context of biological networks.(2006) *BMC bioinformatics*7 (1): 109.
[Pubmed](#) | [Crossref](#) | [Others](#)
57. Howard, K. The bioinformatics gold rush. (2000) *Sci Am* 283(1): 58-63.
[Pubmed](#) | [Crossref](#) | [Others](#)
58. Sud, M., Fahy, E., Cotter, D., et al. Lmsd: lipid maps structure database. (2007) *Nucleic acids Res* 35(1): 527-532.
[Pubmed](#) | [Crossref](#) | [Others](#)
59. Watanabe, K., Yasugi, E., Oshima, M. How to search the glycolipid data in “LIPIDBANK for Web”, the newly developed lipid database in Japan. (2000) *Trends in Glycoscience and Glycotechnology* 12(65): 175-184.
[Pubmed](#) | [Crossref](#) | [Others](#)
60. Yetukuri, L., Ekroos, K., Vidal-Puig, A., et al. Informatics and computational strategies for the study of lipids. (2008) *Mol BioSyst* 4(2): 121-127.
[Pubmed](#) | [Crossref](#) | [Others](#)

Submit your manuscript to Ommega Publishers and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in all major indexing services
- Maximum visibility for your research

Submit your manuscript at



<https://www.ommegaonline.org/submit-manuscript>