OPEN ACCESS
Freely Available Online

OpenAccessPub

JOURNAL OF BIOTECHNOLOGY AND BIOMEDICAL SCIENCE
ISSN NO: 2576-6694

Review Article

DOI : 10.14302/issn.2576-6694.jbbs-18-2173

The Emerging Role of Bioinformatics in Biotechnology

Nida Tabassum Khan [1,*]

[1]Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology Engineering and Management Sciences,(BUITEMS),Quetta, Pakistan

Abstract

Bioinformatic tools is widely used to manage the enormous genomic and proteomic data involving DNA/protein sequences management, drug designing, homology modelling, motif/domain prediction ,docking, annotation and dynamic simulation etc. Bioinformatics offers a wide range of applications in numerous disciplines such as genomics. Proteomics, comparative genomics, nutrigenomics, microbial genome, biodefense, forensics etc. Thus it offers promising future to accelerate scientific research in biotechnology

Corresponding Author: Nida Tabassum Khan, Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology Engineering and Management Sciences, (BUITEMS),Quetta, Pakistan, Email: nidatabassumkhan@yahoo.com

Keywords: CisML; Possum; Clover; SeqVISTA; WebGeSTer; GeneChords.

Received: June 18, 2018 **Accepted:** Aug 02, 2018 **Published:** Aug 07, 2018

Editor: Hammad Afzal, SZABIST, Karachi.

Introduction

Bioinformatics provided computational ways for data analysis by employing informatics tools and softwares to determine protein/gene structure or sequence, homology, molecular modeling of biological system, molecular docking etc to analyze and interpret data in insilico [1].Currently bioinformatics have become a principal technology in all life sciences research. Bioinformatics has been integrated into a number of different disciplines where it assists in better understanding of the data in a shorter time frame [2]. With the massive advancement in information technology, bioinformatics is growing rapidly providing new ways and approaches for the assessment of valuable data [3].Data mining and manipulations is an important aspect of bioinformatic [4]. It allows researchers to collect, store, catalogue and analyse information in unique format that is easily manipulated for future research [5]. Some examples of data manipulation include molecular online tools and the bio extract server [6]. It is useful for accessing bimolecular data from many sources for many purposes. This is lab template for the proper accession and usage of online molecular tools like bio extract [7].

Some applications of bioinformatics in biotechnology is given below:

Genomics

To manage an escalating amount of genomic information, bioinformatic tools are required to maintain and analyze the DNA sequences from different organism [8]. Determination of sequence homology, gene finding, coding region identification, structural and functional analyses of genomic sequences etc, all this is possible by the use of different bioinformatics tools and software packages [9].

Given below is a list of few bioinformatics tools used in genomics Table 1.

Comparative Genomics

Bioinformatics plays an important role in comparative genomics by determing the genomic structural and functional relationship between different biological species [27].

Given below is a list of few bioinformatics tools used in comparative genomics Table 2.

Proteomics:

Advanced molecular based techniques led to the accumulation of huge proteomic data of protein activity patterns, interactions, profiling, composition, structural information, image analysis, peptide mass fingerprinting, peptide fragmentation fingerprinting etc [54,55]. This enormous data could be managed by using different tools of bioinformatics.

Given below is a list of few bioinformatics tools used in proteomics Table 3.

Drug Discovery

Clinical bioinformatics is an emerging new field of bioinformatics that employs various bioinformatics tool such as computer aided drug designing to design novel drugs, vaccines, DNA drug modelling ,insilico drug testing,etc to produce new and effective drugs in a shorter time frame with lower risks [62,63].

Cancer Research and Analysis

Bioinformatic tools such as NCI [64], NCIP (part of NCI) [65] and CBIIT [66] have played an important role in genomics, proteomics, imaging, and metabolomics to increase our knowledge of the molecular basis of cancer [67].

Phylogenetic studies

Using numerous bioinformatics tools, phylogenetic analysis of the molecular data can easily be achieved in a short period of time by constructing phylogenetic trees to study its evolutionary relationship based on sequence alignment [68].

Forensic Science

A number of databases consists of DNA profiles of known delinquents [69]. Advancement in microarray technology, bayesian networks, programming algorithms etc provides an effective method of evidence organization and interpretation [70, 71].

Bio-defense

Though bioinformatics has limited impact on forensic since there is a need for more advanced algorithms and computational applications so that the established databases may exhibit interoperability with each other [72].

Table 1: Bioinformatics tools/databases used in Genomics

Bioinformatics tools	**Purpose**
Carrie	Transcriptional regulatory networks database [10]
CisML	Motif detection tool [11]
ICSF	Identification of conserved structural features in TF binding sites [12]
Possum	Tool for motif searching [13]
Promoser	Promoter extraction tool from eukaryotic organisms [14]
REPFIND	Determine clustered repeats in DNA fragment [15]
Cluster-Buster	Tool for predicting motifs cluster in DNA sequences [16]
Cister	Finds regulatory regions in DNA fragments [17]
Clover	Find overrepresented motifs in DNA sequences [18]
GLAM	Tool for predicting functional motifs [19,20]
MotifViz	Identification of overrepresented motifs [21]
RANKGENE	Tool for analysing gene expression data [22]
ROVER	Predicts overrepresented motifs in DNA fragments [23]
SeqVISTA	Sequences viewer tool [24]
Tractor	Tool to find transcription factors with over-represented binding sites in the upstream regions of co-expressed human genes [25]
OHMICS	Oral human microbiome integrated computational system [26]

Table 2: Bioinformatics tools/databases used in Comparative genomics

Bioinformatics tools	**Purpose**
BLAST	DNA or protein sequence alignment tool [28]
HMMER	Homologous protein sequences searching tool [29]
Clustal Omega	Multiple sequence alignments tool [30]
Sequerome	Sequence profiling tool [31]
ProtParam	Predicts the physico-chemical properties of proteins [32]
novoSNP	Predicts single point mutation in DNA sequences [33]
ORF Finder	Find open reading frame in putative genes [34,35]
Virtual Foorprint	Analysis of whole prokaryotic genome [36]
WebGeSTer	Predicts gene termination sites during transcription [37]
Genscan	Find exon-intron sites in DNA sequences [38]
Softberry Tools	Genomes annotation tool along with the structure and function prediction of biological molecules [39]
MEGA	Study evolutionary relationship [40]
MOLPHY	Maximum likelihood based phylogenetic analysis tool [41]
PHYLIP	Tool for phylogenetic studies [42]
JStree	Tool for viewing and editing phylogenetic trees [43]
Jalview	It is an alignment editing tool [44]
DNA Data Bank of Japan	Resources for nucleotide sequences [45]
Rfam	Database contains collection of RNA families [46]
Uniprot	Protein sequence database[47]
Protein Data Bank	Database provide data on structures of nucleic acids, proteins etc [48]
SWISS PROT	Database containing the manually annotated protein sequences [49]
InterPro	Provide information on protein families, its conserved domains and actives sites [50]
Proteomics Identifications Database	Contains data on functional characterization and post -translation modification of proteins and peptides [51]
Ensembl	Database containing annotated genomes of eukaryotes including human, mouse and other vertebrates [52]
Medherb	Database for medicinally herbs [53]

Table 3: Bioinformatics tools/databases used in Proteomics

Bioinformatics tools	**Purpose**
K2 / FAST	Protein structure alignment tool [56]
SMM	Tool for determing peptides binding to major histocompatibility complex [57]
ZDOCK	Protein-protein docking tool [58]
Docking Benchmark	Tool to evaluate docking algorithms performance [59]
ZDOCK Server	An automated server for running ZDOCK [60]
Z3OnWeb.com	Proteomic analysis for analysing 2D-Gel images [61]

Nutrigenomics

Progressions in structural /functional genomics and molecular technologies such as genome sequencing and DNA microarrays generates valuable knowledge which explains nutrition in relation of an individual's genetics which directly influences its metabolism [73]. Because of the influx of bioinformatics tools, nutrition-related research is tremendously increased [74, 75].

Gene Expression

Regulation of gene expression is the core of functional genomics allowing researchers to apply genomic data to molecular technologies that can quantify the amount of actively transcribing genes in any cell at any time (e.g. gene expression arrays) [76,77].

Given below is a list of few bioinformatics tools used in gene expression study Table 4.

Food Quality

New improvements in computing algorithms and available structural simulation databases of recognized structures has brought molecular modeling into conventional food chemistry. Such simulations will make it possible to improve food quality by developing new food additives by comprehending the basis of taste tenacity, antagonism and complementation [90, 91].

Predicting Protein Structure and Function

Protein topology prediction is now so much easy thanks to bioinformatics which helps in the prediction of 3D structure of a protein to gain an insight into its function as well [92].

Given below is a list of few bioinformatics tools used in protein structure and function prediction Table 5.

Personalized Medicine

Doctors will be able to analyse a patient's genetic profile and prescribe the best available drug therapy and dosage from the beginning by employing bioinformatics tool [100].

Microbial Genome Applications

Microbes have been studied at very basic level with the help of bioinformatics tools required to analyse their unique set of genes that enables them to survive under unfavourable conditions [101].

Conclusion

Thus bioinformatics holds significant importance in countless disciplines of biotechnology such as comparative genomics, drug designing, proteomics, molecular modelling, microbial genomics etc

Table 4: Bioinformatics tools/databases used in Gene expression

Bioinformatics tools	**Purpose**
GeneChords	Conserved gene retrieval tool [78]
GENEVA	Categorizes segmentally altered genes in many complete microbial genomes [79]
HuGE Index	Human tissues gene expression database [80]
Inverted Repeats Finder	Find inverted repeats in genomic DNA [81]
ORChID	Database stores hydroxyl radical cleavage data of DNA sequences [82]
Operons	Predicts functional gene clusters [83]
Optimus	Retrieve conserved gene cluster data from numerous microbial genomes [84]
Predictome	Visualizing tool for bio complexes [85]
Tandem Repeat Database	Store information on tandem repeats in genomic DNA [86]
VisANT	Tools for visualizing and analysing many biological interactions [87]
BSG	Identification of transcription factor binding sites [88]
TFSVM	Detection of transcription factor binding site [89]

Table 5: Bioinformatics tools/databases used in Protein structure and function prediction

Bioinformatics tools	**Purpose**
CATH	Tool for the categorized organization of proteins [93]
Phyre and Phyre2	Tool for protein structure prediction [94]
HMMSTR	For the prediction of sequence-structure correlations in proteins [95]
MODELLER	Predicts 3D structure of protein [96]
JPRED/ APSSP2	Predicts secondary structures of proteins [97]
RaptorX	Predicts protein structure [98]
PHD	Predicts neural network structure [99]

References

1. Kaufman, J. (2004, October). Healthcare and life sciences standards overview-technology for life: NC symposium on biotechnology and bioinformatics. In Biotechnology and Bioinformatics, 2004. Proceedings. Technology for Life: North Carolina Symposium on (pp. 31-41). IEEE.
2. Fulekar, M. H. (Ed.). (2009). Bioinformatics: applications in life and environmental sciences. Springer Science & Business Media.
3. Gu, J., & Bourne, P. E. (Eds.). (2009). Structural bioinformatics (Vol. 44). John Wiley & Sons.
4. Piatetsky-Shapiro, G. (1996). Advances in knowledge discovery and data mining (Vol. 21). U. M. Fayyad, P. Smyth, & R. Uthurusamy (Eds.). Menlo Park: AAAI press.
5. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine, 17(3), 37.
6. Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.
7. Chen, M. S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. IEEE Transactions on Knowledge and data Engineering, 8(6), 866-883.
8. Yang, J. Y., Yang, M. Q., Zhu, M. M., Arabnia, H. R., & Deng, Y. (2008). Promoting synergistic research and education in genomics and bioinformatics. BMC genomics, 9(1), I1.
9. Sensen, C. (2005). Handbook of genome research: genomics, proteomics, metabolomics, bioinformatics, ethical and legal issues.
10. Haverty, P. M., Frith, M. C., & Weng, Z. (2004). CARRIE web service: automated transcriptional regulatory network inference and interactive analysis. Nucleic acids research, 32(suppl_2), W213-W216.
11. Haverty, P. M., & Weng, Z. (2004). CisML: an XML-based format for sequence motif detection software. Bioinformatics, 20(11), 1815-1817.
12. Burden, H. E., & Weng, Z. (2005). Identification of conserved structural features at sequentially degenerate locations in transcription factor binding sites. Genome Informatics, 16(1), 49-58.
13. Fu, Y., Frith, M. C., Haverty, P. M., & Weng, Z. (2004). MotifViz: an analysis and visualization tool for motif discovery. Nucleic acids research, 32 (suppl_2), W420-W423.

14. Halees, A. S., Leyfer, D., & Weng, Z. (2003). PromoSer: a large-scale mammalian promoter and transcription start site identification service. Nucleic acids research, 31(13), 3554-3559.

15. Hamilton, R. S., & Davis, I. (2007, April). RNA localization signals: deciphering the message with bioinformatics. In Seminars in cell & developmental biology (Vol. 18, No. 2, pp. 178-185). Academic Press.

16. Frith, M. C., Li, M. C., & Weng, Z. (2003). Cluster-Buster: Finding dense clusters of motifs in DNA sequences. Nucleic acids research, 31(13), 3666-3668.

17. Kessler, L., Fintor, L., Muha, C., Wun, L. M., Annett, D., & Mazan, K. D. (1993). The Cancer Information Service Telephone Evaluation and Reporting System (CISTERS): a new tool for assessing quality assurance. Journal of the National Cancer Institute. Monographs, (14), 61-65.

18. Frith, M. C., Fu, Y., Yu, L., Chen, J. F., Hansen, U., & Weng, Z. (2004). Detection of functional DNA motifs via statistical over-representation. Nucleic acids research, 32(4), 1372-1381.

19. Tompa, M., Li, N., Bailey, T. L., Church, G. M., De Moor, B., Eskin, E., ... & Makeev, V. J. (2005). Assessing computational tools for the discovery of transcription factor binding sites. Nature biotechnology, 23(1), 137-144.

20. Frith, M. C., Fu, Y., Yu, L., Chen, J. F., Hansen, U., & Weng, Z. (2004). Detection of functional DNA motifs via statistical over-representation. Nucleic acids research, 32(4), 1372-1381.

21. Fu, Y., Frith, M. C., Haverty, P. M., & Weng, Z. (2004). MotifViz: an analysis and visualization tool for motif discovery. Nucleic acids research, 32 (suppl_2), W420-W423.

22. Su, Y., Murali, T. M., Pavlovic, V., Schaffer, M., & Kasif, S. (2003). RankGene: identification of diagnostic genes based on expression data. Bioinformatics, 19(12), 1578-1579.

23. Colaert, N., Helsens, K., Impens, F., Vandekerckhove, J., & Gevaert, K. (2010). Rover: a tool to visualize and validate quantitative proteomics data from different sources. Proteomics, 10(6), 1226-1229.

24. Hu, Z., Frith, M., Niu, T., & Weng, Z. (2003). SeqVISTA: a graphical tool for sequence feature visualization and comparison. BMC bioinformatics, 4 (1), 1.

25. Gang-tiana, L. I. U., Bai-ningb, G. A. O., & Hui-mina, C. A. O. (2007). Realization of Tractor Product Design Management Based-on Network Technique [J]. Tractor & Farm Transporter, 4, 043.

26. Tamburini, S., Shen, N., Wu, H. C., & Clemente, J. C. (2016). The microbiome in early life: implications for health outcomes. Nature medicine, 22(7), 713-722.

27. Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor, G. L., Nelson, C. R., Hariharan, I. K., ... & Cherry, J. M. (2000). Comparative genomics of the eukaryotes. Science, 287(5461), 2204-2215.

28. Moreno-Hagelsieb, G., & Latimer, K. (2007). Choosing BLAST options for better detection of orthologs as reciprocal best hits. Bioinformatics, 24 (3), 319-324.

29. Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic acids research, 39(suppl_2), W29 -W37.

30. Sievers, F., & Higgins, D. G. (2014). Clustal Omega, accurate alignment of very large numbers of sequences. Multiple sequence alignment methods, 105-116.

31. Dinodia, A. (2015). Bioinformatics–Emerging Area in Medical and Biomedical Informatics. International Journal of Pharmaceutical Chemistry and Analysis, 2 (4), 187-191.

32. Bickmore, W. A., & Sutherland, H. G. (2002). Addressing protein localization within the nucleus. The EMBO journal, 21(6), 1248-1254.

33. Weckx, S., Del-Favero, J., Rademakers, R., Claes, L., Cruts, M., De Jonghe, P., ... & De Rijk, P. (2005). novoSNP, a novel computational tool for sequence variation discovery. Genome Research, 15(3), 436-442.

34. Rombel, I. T., Sykes, K. F., Rayner, S., & Johnston, S. A. (2002). ORF-FINDER: a vector for

high-throughput gene identification. Gene, 282(1), 33-41.

35. Tatusov, T., & Tatusov, R. (2007). ORF Finder (Open Reading Frame Finder). National Center for Biotechnology Information. National Institute of Health, 9.
36. Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., & Jahn, D. (2005). Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics, 21 (22), 4187-4189.
37. Mitra, A., Kesarwani, A. K., Pal, D., & Nagaraja, V. (2010). WebGeSTer DB—a transcription terminator database. Nucleic acids research, 39(suppl_1), D129 -D135.
38. Rehm, B. (2001). Bioinformatic tools for DNA/protein sequence analysis, functional assignment of genes and protein classification. Applied microbiology and biotechnology, 57(5-6), 579-592.
39. Coulibaly, I., & Page, G. P. (2008). Bioinformatic tools for inferring functional information from plant microarray data II: analysis beyond single gene. International journal of plant genomics, 2008.
40. Sohpal, V. K., Dey, A., & Singh, A. (2010). MEGA biocentric software for sequence and phylogenetic analysis: a review. International journal of bioinformatics research and applications, 6(3), 230-240.
41. Mehmood, M. A., Sehar, U., & Ahmad, N. (2014). Use of bioinformatics tools in different spheres of life sciences. Journal of Data Mining in Genomics & Proteomics, 5(2), 1.
42. Plotree, D. O. T. R. E. E., & Plotgram, D. O. T. G. R. A. M. (1989). PHYLIP-phylogeny inference package (version 3.2). cladistics, 5(163), 6.
43. Munir, M. (2013). Bioinformatics analysis of large-scale viral sequences: From construction of data sets to annotation of a phylogenetic tree. Virulence, 4(1), 97-106.
44. Clamp, M., Cuff, J., Searle, S. M., & Barton, G. J. (2004). The jalview java alignment editor. Bioinformatics, 20(3), 426-427.
45. Tateno, Y., Imanishi, T., Miyazaki, S., Fukami-Kobayashi, K., Saitou, N., Sugawara, H., & Gojobori, T. (2002). DNA Data Bank of Japan (DDBJ) for genome scale research in life science. Nucleic acids research, 30(1), 27-30.
46. Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S. R., & Bateman, A. (2005). Rfam: annotating non-coding RNAs in complete genomes. Nucleic acids research, 33(suppl_1), D121-D124.
47. UniProt Consortium. (2014). UniProt: a hub for protein information. Nucleic acids research, gku989.
48. Bank, P. D. (1971). Protein Data Bank. Nature New Biol, 233, 223.
49. Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., ... & Pilbout, S. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic acids research, 31(1), 365-370.
50. Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., ... & Finn, R. D. (2008). InterPro: the integrative protein signature database. Nucleic acids research, 37(suppl_1), D211-D215.
51. Vizcaíno, J. A., Côté, R., Reisinger, F., M Foster, J., Mueller, M., Rameseder, J., ... & Martens, L. (2009). A guide to the Proteomics Identifications Database proteomics data repository. Proteomics, 9(18), 4276-4283.
52. Flicek, P., Amode, M. R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., ... & Gil, L. (2011). Ensembl 2012. Nucleic acids research, 40(D1), D84-D90.
53. Rajoka, M. I., Idrees, S., Khalid, S., & Ehsan, B. (2014). Medherb: An Interactiv e Bioinformatics Database and Analysis Resource for Medicinally Importan t Herbs. Curr Bioinform, 9, 23-27.
54. Vihinen, M. (2001). Bioinformatics in proteomics. Biomolecular engineering, 18(5), 241-248.
55. Chou, K. C. (2005). Progress in protein structural class prediction and its impact to bioinformatics and proteomics. Current Protein and Peptide Science, 6 (5), 423-436.
56. Fogel, G. B., & Corne, D. W. (Eds.). (2002). Evolutionary computation in bioinformatics. Morgan Kaufmann.

57. Chen, P., Rayner, S., & Hu, K. H. (2011). Advances of bioinformatics tools applied in virus epitopes prediction. Virologica sinica, 26(1), 1-7.

58. Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., & Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. Bioinformatics, 30(12), 1771-1773.

59. Chen, R., Mintseris, J., Janin, J., & Weng, Z. (2003). A protein–protein docking benchmark. Proteins: Structure, Function, and Bioinformatics, 52(1), 88-91.

60. Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., & Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. Bioinformatics, 30(12), 1771-1773.

61. Smith, C. M. (2000). Bioinformatics, genomics, and proteomics. The scientist, 14(23), 26-26.

62. Murray-Rust, P. (1994). Bioinformatics and drug discovery. Current Opinion in Biotechnology, 5(6), 648-653.

63. Xia, X. (2017). Bioinformatics and Drug Discovery. Current topics in medicinal chemistry, 17(15), 1709-1726.

64. Wulfkuhle, J. D., Liotta, L. A., & Petricoin, E. F. (2003). Proteomic applications for the early detection of cancer. Nature reviews cancer, 3(4), 267-275.

65. Mongkolwat, P., Kleper, V., Talbot, S., & Rubin, D. (2014). The National Cancer Informatics Program (NCIP) Annotation and Image Markup (AIM) Foundation Model. Journal of digital imaging, 27(6), 692-701.

66. Machiela, M. J., & Chanock, S. J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics, 31(21), 3555-3557.

67. Lancashire, L. J., Lemetre, C., & Ball, G. R. (2009). An introduction to artificial neural networks in bioinformatics—application to complex microarray and mass spectrometry datasets in cancer studies. Briefings in bioinformatics, 10(3), 315-329.

68. Baxevanis, A. D., & Ouellette, B. F. (2004). Bioinformatics: a practical guide to the analysis of genes and proteins (Vol. 43). John Wiley & Sons.

69. Krane, D. E., Ford, S., Gilder, J. R., Inman, K., Jamieson, A., Koppl, R., ... & Thompson, W. C. (2008). Sequential unmasking: a means of minimizing observer effects in forensic DNA interpretation. Journal of Forensic Sciences, 53(4), 1006-1007.

70. Grada, A., & Weinbrecht, K. (2013). Next-generation sequencing: methodology and application. J Invest Dermatol, 133(8), e11.

71. Budowle, B., Wilson, M. R., Burans, J. P., Breeze, R. G., & Chakraborty, R. (2005). Microbial forensics. In Microbial forensics (pp. 1-25).

72. Valdivia-Granda, W. A. (2010). Bioinformatics for biodefense: challenges and opportunities. Biosecurity and bioterrorism: biodefense strategy, practice, and science, 8(1), 69-77.

73. Corthésy-Theulaz, I., den Dunnen, J. T., Ferré, P., Geurts, J. M., Müller, M., van Belzen, N., & van Ommen, B. (2005). Nutrigenomics: the impact of biomics technology on nutrition research. Annals of Nutrition and Metabolism, 49(6), 355-365.

74. Müller, M., & Kersten, S. (2003). Nutrigenomics: goals and strategies. Nature Reviews Genetics, 4(4), 315-322.

75. Werf, M. V. D., Schuren, F. H. J., Bijlsma, S., Tas, A. C., & Ommen, B. V. (2001). Nutrigenomics: application of genomics technologies in nutritional sciences and food technology. Journal of Food Science, 66(6), 772-780.

76. Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research, 37(1), 1-13.

77. de Hoon, M. J., Imoto, S., Nolan, J., & Miyano, S. (2004). Open source clustering software. Bioinformatics, 20(9), 1453-1454.

78. Xia, J., Gill, E. E., & Hancock, R. E. (2015). NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression

data. Nature protocols, 10(6), 823-844.

79. Zheng, Y., Roberts, R. J., & Kasif, S. (2004). Segmentally variable genes: a new perspective on adaptation. PLoS biology, 2(4), e81.
80. Lal, A., Lash, A. E., Altschul, S. F., Velculescu, V., Zhang, L., McLendon, R. E., ... & Papadopoulos, N. (1999). A public database for gene expression in human cancers. Cancer Research, 59(21), 5403-5407.
81. Lerat, E. (2010). Identifying repeats and transposable elements in sequenced genomes: how to find your way through the dense forest of programs. Heredity, 104(6), 520-533.
82. Greenbaum, J. A., Pang, B., & Tullius, T. D. (2007). Construction of a genome-scale structural map at single-nucleotide resolution. Genome research, 17 (6), 947-953.
83. Brouwer, R. W., Kuipers, O. P., & van Hijum, S. A. (2008). The relative value of operon predictions. Briefings in bioinformatics, 9(5), 367-375.
84. Pope, W. H., Jacobs-Sera, D., Best, A. A., Broussard, G. W., Connerly, P. L., Dedrick, R. M., ... & Rockenbach, K. (2013). Cluster J mycobacteriophages: intron splicing in capsid and tail genes. PLoS One, 8(7), e69273.
85. Mellor, J. C., Yanai, I., Clodfelter, K. H., Mintseris, J., & DeLisi, C. (2002). Predictome: a database of putative functional links between proteins. Nucleic acids research, 30(1), 306-309.
86. Le Flèche, P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., ... & Vergnaud, G. (2001). A tandem repeats database for bacterial genomes: application to the genotyping of Yersinia pestis and Bacillus anthracis. BMC microbiology, 1(1), 2.
87. Hu, Z., Mellor, J., Wu, J., & DeLisi, C. (2004). VisANT: an online visualization and analysis tool for biological interaction data. BMC bioinformatics, 5(1), 17.
88. Carriço, J. A., Sabat, A. J., Friedrich, A. W., & Ramirez, M. (2013). Bioinformatics in bacterial molecular epidemiology and public health: databases, tools and the next-generation sequencing revolution. Eurosurveillance, 18(4), 20382.
89. Holloway, D. T., Kon, M., & DeLisi, C. (2008). Classifying transcription factor targets and discovering relevant biological features. Biology Direct, 3(1), 22.
90. Desiere, F., German, B., Watzke, H., Pfeifer, A., & Saguy, S. (2001). Bioinformatics and data knowledge: the new frontiers for nutrition and foods. Trends in Food Science & Technology, 12(7), 215-229.
91. Udenigwe, C. C. (2014). Bioinformatics approaches, prospects and challenges of food bioactive peptide research. Trends in Food Science & Technology, 36 (2), 137-143.
92. Kelley, L. A., & Sternberg, M. J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nature protocols, 4(3), 363-371.
93. Orengo, C. A., Pearl, F. M., & Thornton, J. M. (2005). The CATH domain structure database. Structural Bioinformatics, Volume 44, 249-271.
94. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols, 10(6), 845-858.
95. Bystroff, C., & Shao, Y. (2002). Fully automated ab initio protein structure prediction using I-SITES, HMMSTR and ROSETTA. Bioinformatics, 18 (suppl_1), S54-S61.
96. Kuntal, B. K., Aparoy, P., & Reddanna, P. (2010). EasyModeller: A graphical interface to MODELLER. BMC research notes, 3(1), 226.
97. Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: a protein secondary structure prediction server. Nucleic acids research, 43(W1), W389-W394.
98. Peng, J., & Xu, J. (2011). RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins: Structure, Function, and Bioinformatics, 79(S10), 161-171.
99. Rost, B. (1996). [31] PHD: Predicting one-dimensional protein structure by profile-based neural networks. Methods in enzymology, 266, 525-539.

100. Fernald, G. H., Capriotti, E., Daneshjou, R., Karczewski, K. J., & Altman, R. B. (2011). Bioinformatics challenges for personalized medicine. Bioinformatics, 27(13), 1741-1748.

101. Kerkhoven, R., Van Enckevort, F. H., Boekhorst, J., Molenaar, D., & Siezen, R. J. (2004). Visualization for genomics: the microbial genome viewer. Bioinformatics, 20(11), 1812-1814.