Biyoinformatiğe Giriş

Selman MUSLU[1]
Murat TURAN[2]
Emre İLHAN[3]

Giriş

Biyoinformatik, biyoloji, bilgi teknolojileri ve matematik alanlarını birleştiren, genomik, proteomik, metabolomik gibi özellikle moleküler genetik verilerin bilgisayar uygulamaları yardımıyla toplanması, depolanması, analiz edilmesi, yorumlanması ve görselleştirilmesine odaklanan bilim dalıdır. Bu veriler biyoinformatik araçlarının yardımıyla canlıların genetik yapısı, proteinlerin yapısı, metabolik yolakları gibi konuları anlamaya yardımcı olmaktadır. Biyoinformatik bilimi, bu verilerin anlaşılmasına yardımcı olmasının yanı sıra, canlıların evrimsel süreçleri, hastalıkların oluşumu ve tedavi edilmesi gibi konulara da odaklanmaktadır (1-6).

Ayrıca biyoinformatik bilimi, birçok alanda kullanılmaktadır (7,8);

- Tıp ve sağlık: Genetik bilgilerin analiz edilerek hastalıkların tanısı ve tedavisinde, ilaçların birbirleri ile olan etkileşimlerinin anlaşılmasında, tıbbi verilerin işlenmesi ve yorumlanmasında, sağlık verilerinin daha etkin bir şekilde yönetilmesinde kullanılmaktadır.

---

[1] Dr., Erzurum Teknik Üniversitesi, Fen Fakültesi, Moleküler Biyoloji ve Genetik Bölümü, selman.muslu25@erzurum.edu.tr, ORCID iD: 0000-0003-4777-0726

[2] Dr. Öğr. Üyesi, Erzurum Teknik Üniversitesi, Fen Fakültesi, Moleküler Biyoloji ve Genetik Bölümü, m.turan@erzurum.edu.tr, ORCID iD: 0000-0003-2900-1755

[3] Doç. Dr., Erzurum Teknik Üniversitesi, Fen Fakültesi, Moleküler Biyoloji ve Genetik Bölümü, emre.ilhan@erzurum.edu.tr, ORCID iD: 0000-0002-8404-7900

Kaynaklar

1. Franco ML, Cediel JF, Payan, C. Breve historia de la bioinformática. *Colombia Médica*. 2008; 39(1): 117-120.

2. Zhang X, Zhou X, Wang X. Chapter1: Basics for Bioinformatics. Jiang R, Zhang X, Zhang MQ (ed.) *Basics of Bioinformatics: Lecture Notes of the Graduate Summer School on Bioinformatics of China* içinde. Tsinghua University Press, Beijing and Springer-Verlag Berlin Heidelberg; 2013. p. 1-25.

3. Mirsaydaliyevich YI. History of Bioinformatics. *Journal of Social Sciences &Interdisciplinary Research*. 2022; 11(7): 72-76.

4. Sharma MK, Dhar MK, Kaul S. Bioinformatics: An introduction and Overview. *International Journal of Engineering Research and Development*. 2012; 3(8): 88-99.

5. Fenstermacher D. Introduction to Bioinformatics. *Journal of the Amerıcan Society for Information Science and Technology*. 2005; 56(5): 440-446.

6. Atalay RÇ. Neden Biyoinformatik? *Avrasya Dosyası*. 2022; 8(3): 127-139.

7. Ünel NM, Önal İK, Çetin F, et al. Biyoinformatiğe Giriş. Baloğlu MC (ed.) *Biyoinformatik Temelleri ve Uygulamaları* içinde. Pegem Akademi; 2020: p. 2-16.

8. Keklik G, Özcan BD. An Overview of Bioinformatics as a Multidisciplinary Science. *OKU Journal of The Institute of Science and Technology*. 2022; 5(2): 1082-1091.

9. Fu J, Zhang Y, Liu J, et al. Pharmacometabonomics: data processing and statistical analysis. *Briefings in Bioinformatics*. 2021; 22(5): 1-25.

10. Chen YPP. Introduction to Bioinformatics. Chen YPP. (ed.) Bioinformatics Technologies. Springer-Verlag Berlin Heidelberg: 2005: p. 1-13.

11. O'Donoghue SI. Grand Challenges in Bioinformatics Data Visualization. *Frontiers in Bioinformatics*. 2021; 1: 669186. doi: 10.3389/fbinf.2021.669186

12. Çelik Altunoğlu Y, Ceylan KB, Ceylan Y, et al. Biyolojik Veri Tabanları. Baloğlu MC (ed.) Biyoinformatik Temelleri ve Uygulamaları içinde. Pegem Akademi; 2020: p. 19-44.

13. Kamble A, Khairkar R. Basics of Bioinformatics in Biological Research. *International Journal of Applied Sciences and Bİotechnology*. 2016; 4(4): 425-429. doi:10.3126/ijasbt. v4i4.16252

14. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Research*. 2001; 11(5): 863-874. doi:10.1101/GR.176601

15. Gollery M. Bioinformatics: Sequence and Genome Analysis. *2nd ed. Clinical Chemistry*. 2005; 51(11): 2219-2220.

16. Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. *Journal of Molecular Biology*. 1990; 215(3): 403–410. doi: S0022-2836(05)80360-2

17. Elias I. Settling the intractability of multiple alignment. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*. 2006; 13(7): 1323-1339. doi: CMB.2006.13.1323

18. Wang L, Jiang T. On the complexity of multiple sequence alignment. *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology*. 1994; 1(4): 337-348. doi: CMB.1994.1.337

19. Thompson JD, Higgins DG, Gibson, TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. 1994; 22(22): 4673-4680. doi: 10.1093/NAR/22.22.4673

20. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of Molecular Biology. 2000; 302(1): 205-217. doi: 10.1006/JMBI.2000.4042

21. Ortet P, Bastien O. Where Does the Alignment Score Distribution Shape Come from? *Evolutionary Bioinformatics Online*. 2010; 6(6): 159. doi: 10.4137/EBO.S5875

22. Dayhoff MO, Eck RV. *Atlas of protein sequence and structure*. Silver Spring, Md.: National Biomedical Research Foundation; 1972.

23. Pevsner J. *Bioinformatics and functional genomics*. (3. Baskı). West Sussex, UK: John Wiley & Sons; 2015.

24. Dai X, Shen L. Advances and Trends in Omics Technology Development. *Frontiers in Medicine*. 2022; 9: 1546. doi: 10.3389/FMED.2022.911861/BIBTEX

25. Orgogozo V, Peluffo AE, Morizot B. The "Mendelian Gene" and the "Molecular Gene": Two Relevant Concepts of Genetic Units. *Current Topics in Developmental Biology*, 2016; 119: 1-26. doi: 10.1016/BS.CTDB.2016.03.002

26. Debnath M, Prasad GB, Bisen PS. Introduction to Molecular Diagnostic. Mousumi Debnath, Godavarthi B.K.S. Prasad, Prakash S. Bisen (Eds.), *Molecular Diagnostics: Promises and Possibilities* içinde. Netherlands: Springer, Dordrecht; 2010. p. 1-10.

27. Davies H. A role for "omics" technologies in food safety assessment. *Food Control*. 2010; 21(12): 1601-1610. doi: 10.1016/J.FOODCONT.2009.03.002

28. Budak ŞÖ, Dönmez S. Gıda Biliminde Yeni Omik Teknolojileri. *Gıda*. 2012; 37(3): 173-179.

29. Bilello J. The agony and ecstasy of "OMIC" technologies in drug development. *Current Molecular Medicine*, 2005; 5(1): 39-52. doi: 10.2174/1566524053152898

30. Alabert C, Groth A. Chromatin replication and epigenome maintenance. *Nature reviews Molecular Cell Biology*, 2012; 13(3): 153-167.

31. Horton P, Park KJ, Obayashi T, et al. WoLF PSORT: protein localization predictor. Nucleic acids research. 2007; 35(suppl_2): W585-W587.

32. Fischer HP. Towards quantitative biology: integration of biological information to elucidate disease pathways and to guide drug discovery. *Biotechnology Annual Review*. 2005; 11(SUPPL.): 1-68. doi:10.1016/S1387-2656(05)11001-1

33. Huang S, Chaudhary K, Garmire LX. More Is Better: Recent Progress in Multi-Omics Data Integration Methods. *Frontiers in Genetics*. 2017; 8(JUN): 84. doi: 10.3389/FGENE.2017.00084

34. Kim M, Tagkopoulos I. Data integration and predictive modeling methods for multi-omics datasets. *Molecular Omics*. 2018; 14(1): 8-25. doi: 10.1039/C7MO00051K

35. Lin E, Lane HY. Machine learning and systems genomics approaches for multi-omics data. *Biomarker Research*. 2017; 5(1): 1-6. doi: 10.1186/S40364-017-0082-Y/FIGURES/3

36. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome biology*. 2017; 18(1): 83. doi: 10.1186/s13059-017-1215-1

37. Aborode AT, Awuah WA, Mikhailova T, et al. OMICs Technologies for Natural Compounds-based Drug Development. *Current Topics in Medicinal Chemistry*. 2022; 22(21): 1751-1765. doi: 10.2174/1568026622666220726092034

38. Zhao K, Rhee SY. Omics-guided metabolic pathway discovery in plants: Resources, approaches, and opportunities. *Current Opinion in Plant Biology*. 2022; 67: 102222. doi: 10.1016/J.PBI.2022.102222

39. Tyers M, Mann M. From genomics to proteomics. *Nature*. 2003; 422(6928): 193-197.

40. Debnath M, Prasad GB, Bisen PS. Omics technology. Mousumi Debnath, Godavarthi B.K.S. Prasad, Prakash S. Bisen (Eds.), *Molecular Diagnostics: Promises and Possibilities* içinde. Netherlands: Springer, Dordrecht; 2010. p. 11-31.

41. Russell C, Rahman A, Mohammed, AR. Application of genomics, proteomics and metabolomics in drug discovery, development and clinic. *Therapeutic Delivery*. 2013; 4(3): 395-413.

42. Martin DB, Nelson PS. From genomics to proteomics: techniques and applications in cancer research. *Trends Cell Biol*. 2001; 11: 60-65.

43. Unamba CI, Nag A, Sharma RK. Next generation sequencing technologies: the doorway to the unexplored genomics of non-model plants. *Frontiers in Plant Science*. 2015; 6: 1074.

44. Horgan RP, Kenny LC. 'Omic'technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist*. 2011; 13(3): 189-195.

45. Bal SH, Budak F. Genomik, proteomik kavramlarına genel bakış ve uygulama alanları. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*. 2013; 39(1): 65-69.

46. Vallejos-Vidal E, Reyes-Cerpa S, Rivas-Pardo JA, et al. Single-nucleotide polymorphisms (SNP) mining and their effect on the tridimensional protein structure prediction in a set of immunity-related expressed sequence tags (EST) in Atlantic salmon (*Salmo salar*). *Frontiers in Genetics*, 2020; 10: 1406.

47. Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Research*. 2006; 34: D354–D357.

48. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics*. 2009; 10(1): 57. doi: 10.1038/NRG2484

49. Jiménez-Chillarón JC, Díaz R, Ramón-Krauel M. Omics Tools for the Genome-Wide Analysis of Methylation and Histone Modifications. *Comprehensive Analytical Chemistry*. 2014; 63: 81-110. doi: 10.1016/B978-0-444-62651-6.00004-0

50. Kodama Y, Shumway M, Leinonen R, et al. The Sequence Read Archive: explosive growth of sequencing data. *Nucleic Acids Research*. 2012; 40(Database issue): D54–D56. doi: 10.1093/nar/gkr854

51. Dunn MJ. Proteomics comes of age. *Proteomics*. 2014; 14(1): 1-2.

52. Macaulay IC, Carr P, Gusnanto A, Ouwehand WH, et al. Platelet genomics and proteomics in human health and disease. *The Journal of Clinical Investigation*. 2005; 115(12): 3370-3377.

53. Kenyon GL, DeMarini DM, Fuchs E, et al. Defining the mandate of proteomics in the post-genomics era: workshop report. *Molecular & Cellular Proteomics*. 2002; 1(10): 763-780.

54. Pierce JD, Fakhari M, Works KV, et al. Understanding proteomics. *Nursing & Health Sciences*. 2007; 9(1): 54-60.

55. Horgan RP, Kenny LC. 'Omic'technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist*. 2001; 13(3): 189-195.

56. Mallick P, Kuster B. Proteomics: a pragmatic perspective. *Nature Biotechnology*. 2010; 28(7): 695-709.

57. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry*. 1975; 250(10): 4007-4021.

58. Cho A, Normile D. Nobel Prize in chemistry. Mastering macromolecules. *Science*. 2002; 298: 527–528.

59. Blueggel M, Chamrad D, Meyer HE. Bioinformatics in proteomics. *Current Pharmaceutical Biotechnology*. 2004; 5(1): 79-88.

60. Ludvigsen M, Honoré B. Transcriptomics and proteomics: integration? *eLS*, 2018; 1-7.

61. Wen B, Zeng WF, Liao Y, Shi Z, et al. Deep learning in proteomics. *Proteomics*. 2020; 20(21-22): 1900335.

62. Oliver SG, Winson MK, Kell DB, Systematic functional analysis of the yeast genome. *Trends in Biotechnology*. 1998; 16(9): 373-378.

63. Wishart DS, Tzur D, Knox C, et al. HMDB: the human metabolome database. *Nucleic Acids Research*. 2007; 35(suppl_1): D521-D526.

64. Ramautar R, Demirci A, de Jong GJ. Capillary electrophoresis in metabolomics. *TrAC Trends in Analytical Chemistry*. 2006; 25(5): 455-466.

65. Monton MRN, Soga T. Metabolome analysis by capillary electrophoresis–mass spectrometry. *Journal of Chromatography A*. 2007; 1168(1-2): 237-246.

66. Cevallos-Cevallos JM, Rouseff R, Reyes-De-Corcuera JI. Untargeted metabolite analysis of healthy and Huanglongbing-infected orange leaves by CE-DAD. *Electrophoresis*. 2009; 30(7): 1240-1247.

67. Cevallos-Cevallos JM, Reyes-De-Corcuera JI, Etxeberria E, et al. Metabolomic analysis in food science: a review. *Trends in Food Science & Technology*. 2009; 20(11-12): 557-566.

68. Vautz W, Zimmermann D, Hartmann M, et al. Ion mobility spectrometry for food quality and safety. *Food Additives and Contaminants*. 2006; 23(11): 1064-1073.

69. Cho IH, Kim YS, Choi HK. Metabolomic discrimination of different grades of pine-mushroom (*Tricholoma matsutake* Sing.) using 1H NMR spectrometry and multivariate data analysis. *Journal of Pharmaceutical and Biomedical Analysis*. 2007; 43(3): 900-904.

70. Beckmann M, Enot DP, Overy DP, et al. Representation, comparison, and interpretation of metabolome fingerprint data for total composition analysis and quality trait investigation in potato cultivars. *Journal of Agricultural and Food Chemistry*. 2007; 55(9): 3444-3451.

71. van der Greef J, Davidov E, Verheij E, et al. The Role of Metabolomics in Systems Biology. Harrigan GG, Royston Goodacre (Eds), *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis* içinde. New York: Springer New York; 2003. p. 171-198.

72. Kaddurah-Daouk R, Krishnan K. Metabolomics: a global biochemical approach to the study of central nervous system diseases. *Neuropsychopharmacology*. 2009; 34(1): 173-186.

73. German JB, Hammock BD, Watkins SM. Metabolomics: building on a century of biochemistry to guide human health. *Metabolomics*. 2005; 1(1): 3-9.

74. Li S, Tian Y, Jiang P, et al. Recent advances in the application of metabolomics for food safety control and food quality analyses. *Critical Reviews in Food Science and Nutrition*. 2021; 61(9): 1448-1469.

75. Cowan MM. Plant products as antimicrobial agents. *Clinical Microbiology Reviews*. 1999; 12(4): 564-582.

76. Patwardhan B, Warude D, Pushpangadan P, et al. Ayurveda and traditional Chinese medicine: a comparative overview. *Evidence-based Complementary and Alternative Medicine*. 2005; 2(4): 465-473.

77. Saito K, Matsuda F. Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology*. 2010; 61: 463-489.

78. Rocca-Serra P, Brazma A, Parkinson H, et al. ArrayExpress: a public database of gene expression data at EBI. *Comptes Rendus Biologies*. 2003; 326(10-11): 1075-1078.

79. Barrett T, Edgar R. Gene Expression Omnibus: microarray data storage, submission, retrieval, and analysis. *Methods in Enzymology*. 2006; 411: 352-369.

80. Wise RP, Caldo RA, Hong L, et al. *Barleybase/plexdb*. In Plant bioinformatics. Humana Press; 2007. pp. 347-363.

81. Leinonen R, Sugawara H, Shumway M, et al. The sequence read archive. *Nucleic Acids Research*, 2010; 39(suppl_1): D19-D21.

82. Harishchander A. A review on application of bioinformatics in medicinal plant research. *Bioinformatics & Proteomics Open Access Journal*. 2017; 1(1): 000104.

83. Kanehisa M. Representation and analysis of molecular networks involving diseases and drugs. *Genome Informatics 2009: Genome Informatics Series*. 2009; 23: 212-213.

84. Kaul S, Koo HL, Jenkins J, et al. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000; 408(6814): 796-815.

85. Matsumoto T, Wu JZ, Kanamori H, et al. The map-based sequence of the rice genome. *Nature*. 2005; 436(7052): 793–800.

86. Skuse GR, Du C. Bioinformatics tools for plant genomics. *International Journal of Plant Genomics*, 2008; 2008: 910474.

87. Batley J, Edwards D. The application of genomics and bioinformatics to accelerate crop improvement in a changing climate. *Current Opinion in Plant Biology*. 2016; 30: 78-81.

88. Shafi A. Haq E. Fazili K.M. et al. *Impact of Bioinformatics on Plant Science Research and Crop Improvement*. Hakeem KR, Shaik NA, Banaganapalli B, et al. (Eds.), *Essentials of Bioinformatics, Volume III* içinde. Cham: Springer; 2019. p. 29-46.

89. Henry RJ. Genomics Strategies for Germplasm Characterization and the Development of Climate Resilient Crops. Santosh Kumar (Ed.) Crop breeding: bioinformatics and preparing for climate change içinde. New York: Apple Academic Press; 2016. p. 3-12.

90. Takeda S, Matsuoka M. Genetic approaches to crop improvement: responding to environmental and population changes. *Nature Reviews Genetics*. 2008; 9(6): 444-457.

91. Ahmad P, Ashraf, M, Younis M, et al. Role of transgenic plants in agriculture and biopharming. Biotechnology advances, 2012; 30(3): 524-540.

92. Tester M, Langridge P. Breeding technologies to increase crop production in a changing world. *Science*. 2010; 327(5967): 818-822.

93. Cogburn LA, Porter TE, Duclos MJ, et al. Functional genomics of the chicken-a model organism. Poultry Science, 2007; 86(10): 2059-2094.

94. AH. Genomics of *Sorghum*. *International Journal of Plant Genomics*. 2008; 2008: 362451.

95. Malav AK, Chandrawat I, Chandrawat KS. Gene pyramiding: an overview. *International Journal of Current Research inBiosciences and Plant Biology*. 2016; 3(7): 22-28.

96. Feltus FA, Wan J, Schulze SR, et al. An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. *Genome Research*. 2004; 14(9): 1812-1819.

97. Varshney RK, Graner A, Sorrells ME. Genomics-assisted breeding for crop improvement. *Trends in Plant Science*. 2005; 10(12): 621-630.

98. Pandey SP, Somssich IE. The role of WRKY transcription factors in plant immunity. *Plant Physiology*. 2009; 150(4): 1648-1655.

99. Schenk PM, Carvalhais LC, Kazan K. Unraveling plant–microbe interactions: can multi-species transcriptomics help? *Trends in Biotechnology*, 2012; 30(3): 177-184.

100. Hu H, Scheben A, Edwards D. Advances in integrating genomics and bioinformatics in the plant breeding pipeline. *Agriculture*. 2008; 8(6): 75.

101. Jorgensen WL. The many roles of computation in drug discovery. *Science*. 2004; 303(5665): 1813-1818.

102. Kavitha K, Venkataraman G, Parida A. An oxidative and salinity stress induced peroxisomal ascorbate peroxidase from Avicennia marina: molecular and functional characterization. *Plant Physiology and Biochemistry*. 2008; 46(8-9): 794-804.

103. Cabral A, Stassen JH, Seidl MF, et al. Identification of Hyaloperonospora arabidopsidis transcript sequences expressed during infection reveals isolate-specific effectors. *PLoS One*, 2011; 6(5): e19328.

104. Yuan D, Tu L, Zhang X. Generation, annotation and analysis of first large-scale expressed sequence tags from developing fiber of *Gossypium barbadense* L. *PLoS One*. 2011; 6(7): e22758.