

Multidisipliner Bir Bilim Dalı Olarak Biyoinformatiğe Genel Bir Bakış

Gülşah KEKLİK^{1*}, Bahri Devrim ÖZCAN²

^{1,2}Çukurova Üniversitesi Ziraat Fakültesi Zootekni Bölümü Biyometri ve Genetik Anabilim Dalı, Adana

¹<https://orcid.org/0000-0002-1775-2773>

²<https://orcid.org/0000-0002-9198-656X>

*Sorumlu yazar: gulsahkeklik@gmail.com

Derleme

Makale Tarihiçesi:

Geliş tarihi: 26.03.2022

Kabul tarihi: 22.05.2022

Online Yayınlanma: 18.07.2022

Anahtar Kelimeler:

Biyoinformatik

Bilişim teknolojileri

Veri tabanları

Moleküler biyoloji ve genetik

ÖZ

Biyologların, DNA'nın üç boyutlu yapısından yola çıkarak hücre içindeki bilgilerin kromozomlar tarafından kuşaklar arası taşındığını keşfetmelerinin ardından gen projeleri üzerinde çalışmaları ve bunu yaparken bilişim teknolojilerinde kullanılan araçlardan yararlanmaları kaçınılmaz bir hal almıştır. Biyoinformatik, biyolojik sistemlerden elde edilen verilerin analizinin yapılması ve değerlendirilmesi için biyoloji, tıp bilimleri, bilişim teknolojileri, matematik ve biyoistatistiğin birleşiminden doğan multidisipliner bir bilim dalıdır. Biyolojik dizilerin analizi için geliştirilen biyoinformatik bilimi; günümüzde biyoloji, genomik çalışmalar ve gen ekspresyon çalışmaları, tanı ve tedavi ile tanısı konmuş hastalıklara ilişkin ilaçlar geliştirme ve benzeri alanlarda kullanılmakta olup bilişim teknolojileri ve biyoistatistik biliminden oldukça faydalanmaktadır. Bu çalışmanın konusu kapsamında, biyoinformatik biliminin ortaya çıkışı, diğer disiplinler ile ilişkisi, uygulama alanları ve bir bilim dalı olarak gelecekte ne anlam ifade edeceği ele alınarak bu sürece ilişkin bilgiler verilecektir.

An Overview of Bioinformatics as a Multidisciplinary Science

Review

Article History:

Received: 26.03.2022

Accepted: 22.05.2022

Published online: 18.07.2022

Keywords:

Bioinformatics

Information technologies

Databases

Molecular biology and genetics

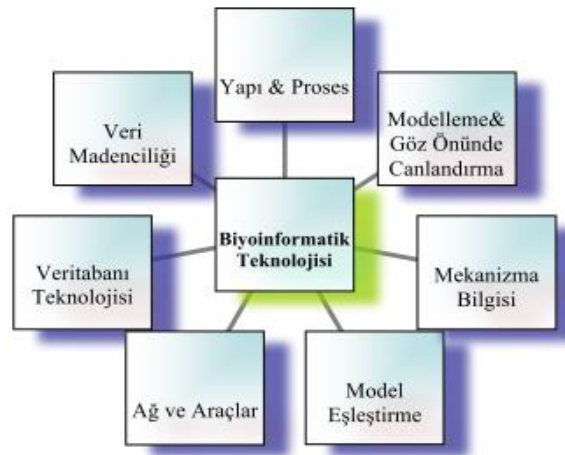
ABSTRACT

After biologists discovered that the information in the cell is carried between generations by chromosomes based on the three-dimensional structure of DNA, it has become inevitable to work on gene projects and to benefit from the tools used in information technologies while doing this. Bioinformatics is a multidisciplinary branch of science arising from the combination of biology, medical sciences, information technologies, mathematics and biostatistics for the analysis and evaluation of data obtained from biological systems. The science of bioinformatics developed for the analysis of biological sequences; today, it is used in biology, genomic studies and gene expression studies, diagnosis and treatment, drug development for diagnosed diseases and similar fields, and it makes great use of information technologies and biostatistics. Within the scope of the subject of this study, the emergence of bioinformatics, its relationship with other disciplines, its application areas and what it will mean in the future as a science will be discussed and information about this process will be given.

To Cite: Keklik G., Özcan BD. Multidisipliner Bir Bilim Dalı Olarak Biyoinformatiğe Genel Bir Bakış. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2022; 5(2): 1082-1091.

Giriş

Biyoinformatik, büyük ölçekli biyolojik verilerin analizini yapmak üzere birtakım disiplinlerin birleşiminden doğan multidisipliner bir bilim dalı olup bazı algoritmalar ve biyoistatistiksel yöntemler aracılığıyla biyolojik verilere uygulanan bilişim teknikleridir (Akın ve ark., 2014). Bilgisayar bilimi ve moleküler biyolojinin gelişiminin bir sonucu olarak biyoinformatik, biyolojik bilgilerin oluşturulması ve korunması için veri tabanları oluşturulmasıdır. Genomik veri analizinde kullanılan işlemler, veri toplanması ve bilgi oluşturulması da buna dahil, yüksek tekrarlılık ve matematiksel karmaşıklık içerdiğinden analizler yapılırken bilgisayarın etkin kullanımı önemli bir yer tutmaktadır (Tanyolaç ve ark., 2015). Biyoinformatik; bilgisayar bilimi, yazılım mühendisliği, matematik ve moleküler biyoloji disiplinlerini kapsayan bir bilim dalı olarak öne çıkmaktadır (Searls, 2000).



Şekil 1. Biyoinformatik teknolojisi (Chen ve Chang, 1998)

Tıp ve biyoloji alanındaki gelişmeler, bilişim teknolojilerinden faydalanmayı kaçınılmaz kılması ile birlikte genlerin üç boyutlu grafiklerinin çıkarılması ve kısa zaman dilimlerinde yüksek sayılarda veri üretebilen cihazlar aracılığıyla moleküler yapılarla ilişkin veri tabanlarının oluşturulması olanaklı hale gelmiştir (Eser ve Öngel, 2014). Gelişen ve giderek artan bilgiyi toplamak ve yönetmek, ona en pratik yoldan ulaşılmasını sağlamak, karmaşık yapıdaki biyolojik sistemleri inceleyerek onu çözülebilir kılmak biyoinformatiğin hedefleri arasındadır (Kumaresan ve ark., 2014).

Biyoinformatikte büyük verilerden bilgi elde etmek için makine öğrenimi yaygın olarak kullanılan başarılı bir yöntemdir. Makine öğrenimi algoritmaları; temel kalıplar ortaya çıkarmak, modeller oluşturmak ve en uygun modele dayalı tahminler yapmak için ilgili verileri kullanır. İyi bilinen bazı algoritmalar (destek vektör makineleri, gizli Markov modelleri, Bayes ağları, Gauss ağları) genomik, proteomik, sistem biyolojisi ve diğer birçok alanda uygulanmıştır (Min ve ark., 2017).

Biyoinformatiğin Tarihçesi

Tıp alanındaki çalışmaların zamanla gelişmesi, birtakım yeni yöntemlerin kullanılmasını merak konusu haline getirmiştir. Bilim insanları bir süredir birçok organizmanın tüm genom dizisini

oluşturmak için çalışmaktadır. Biyoinformatik ise bu araştırmaların bir sonucu olarak ortaya çıkmıştır (Kuonen, 2003).

1951 yılında, Pauling ve Corey, proteinlerin sekonder yapılarının doğru tahmini için bir yaklaşım geliştirmiş olup bu yaklaşımın biyoinformatiğin başlangıcı olduğu söylenebilir. 1966 yılında, moleküler grafiklerin bilgisayarla çizimine ilişkin ilk makalenin Scientific American dergisinde yayımlanmasının biyoinformatik biliminin asıl başlangıcı olduğu bilinmektedir. “Biyoinformatik” bir terim olarak 1980’li yılların ortalarından sonra kullanılmış olup moleküler biyoinformatik, computational biology (hesaplamalı biyoloji), biocomputing (biyobilişim) terimleri de biyoinformatik yerine kullanılmaktadır (Gentleman ve ark., 2004).

1988’de, temel ve karmaşık verilerin analizinin ve yorumunun yapılabilmesi için birtakım yeni yöntemlerin geliştirilmesinde bilinen “National Center for Biotechnology Information (NCBI)” kurulmuştur. Ekim 1990’da, bu alandaki en önemli projelerden biri olan İnsan Genom Projesi (İGP) ortaya çıkmıştır. İGP’nin biyoinformatik çalışmaların ilerlemesinde önemi büyüktür. 2003 yılında tamamlanan bu projeye İngiltere, Japonya, Fransa, Almanya gibi bazı ülkeler de katkı sağlamışlardır. 2005 yılında bitirilmesi beklenen bu proje, teknolojik gelişmeler sayesinde beklenenden iki yıl önce bitirilmiş olup proje verilerindeki hataların temizlenmesine ve bu verilere ilişkin analizlere devam edileceği öngörülmüştür (Collins ve ark., 2003; Polat ve Karahan, 2009).

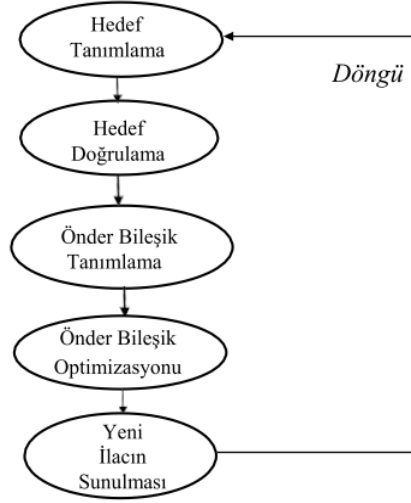
2013 yılında, Dibyajyoti Saha ve arkadaşları tarafından yapılan bir çalışmada, biyoinformatikte ilaç keşfinin maliyetine ilişkin etkiler ele alınmıştır. Maliyeti azaltma, ilacı laboratuvarından hastaya ulaştırmak için gerekli süre, hedef odaklı ve düşük yan etkileri olan ilaçlar üretmenin yollarını keşfetmek için biyoinformatik araçların çalışmalarına değinilmiştir.

Xuhua Xia tarafından yapılmış bir çalışmada, genomik ve proteomik terimlerinden bahsedilmiş ve biyoinformatik analizlerde kullanılan büyük veri yapılarına ilişkin genom mimarisi, epigenetic, genomik, exome sıralama, transkriptomik, ribozom profili, proteomik, protein (RNA) yapıları ve cistromic örnekleri verilmiştir (Xia, 2017).

Biyoinformatiğin Uygulama Alanları

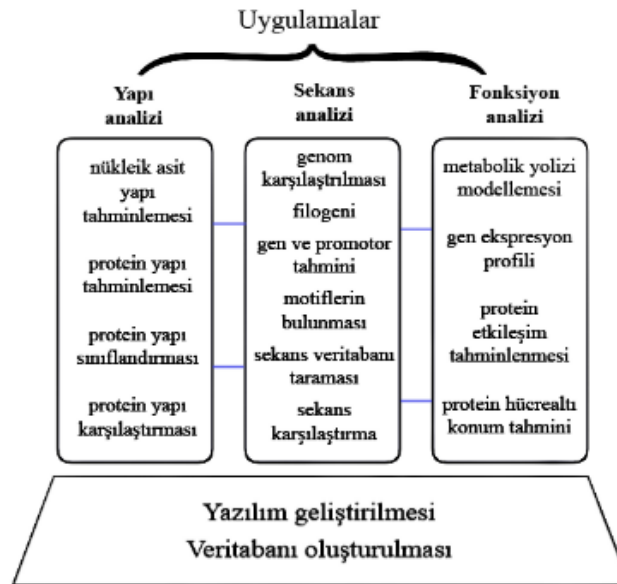
Biyoinformatiğin temelinde farklı biyomoleküller arasındaki benzerlikleri tespit etme durumu söz konusudur. Karakteristiği tam bir biçimde tespit edilememiş bir proteinin homologlarını (türdeşlerini) bulmak, proteinler arasında bilgi aktarımını sağlamak ve bunun sonucunda elde edilen bilgileri değerlendirmek biyoinformatiğin uygulama alanları arasındadır. Biyoinformatiğin tıp alanında uygulanan deneysel çalışmalar, diğer canlılarda insan proteininin yapısını modellemeye yardımcı olur. Dolayısıyla modellenen proteine bağlanabilen moleküllerin tasarlanması sağlanabilmektedir. Çeşitli hastalıklardan etkilenen hücrelerin, sağlıklı hücrelerle karşılaştırılarak aralarındaki farklılıklardan kaynaklanan hastalıkların teşhisi ve bu hastalıklara uygun ilaç tasarımları yapılabilmektedir (Attwood ve Parry-Smith, 1999; Luscombe ve ark., 2001; Tandon ve Bhattacharjee, 2009). Bu anlamda medikal tanı ve tedavi yöntemlerinin geliştirilmesi bakımından önemli kazanımlar elde edilmiştir (Hogue,

2002). Biyoinformatik, ilaç endüstrisinde genomik maddeleri bir ilaç hedefi kaynağı olarak benimsemiştir ve dolayısıyla genom çapında üretilen verileri kullanmak bakımından oldukça önemli bir yer tutmaktadır (Drews, 2000).



Şekil 2. İlaç keşif döngüsü (Blundell ve ark., 2006)

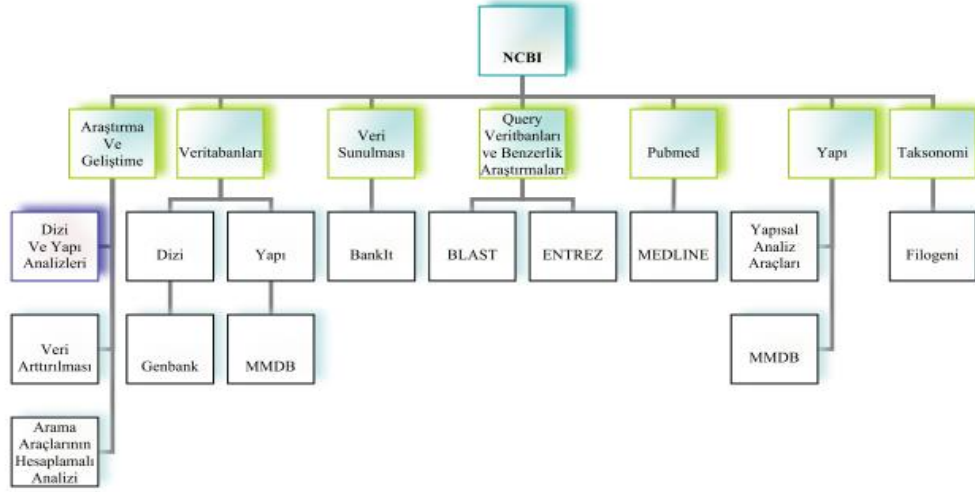
Çoklu veri kaynakları ile oluşturulan bilgilerin bir araya getirilmesi, biyoinformatiğin önemli görevlerinden birisidir. Birden çok verinin kaynağına ulaşabilecek şekilde veri kaynaklarının incelenmesi noktasında iki önemli erişim sisteminden bahsetmek mümkündür. Bunlardan ilki, “Dizi Erişim Sistemi (SRS, Sliding Rail System)” olup bu sistem kullanıcıya nükleik asit, protein dizisi, protein motifi, protein yapısı ve bibliyografik veri tabanlarına bağlantı sağlamaktadır. Diğeri ise DNA’ya, protein dizilerine, genom harita verilerine, 3 boyutlu makromoleküler yapılar ve PubMed bibliyografik veri tabanına ulaşmayı mümkün kılan “Entrez Sistemi”dir (Attwood ve Parry-Smith, 1999; Luscombe ve ark., 2001; Tandon ve Bhattacharjee, 2009).



Şekil 3. Biyoinformatiğin alt çalışma alanları (Xiong ve ark., 2006).

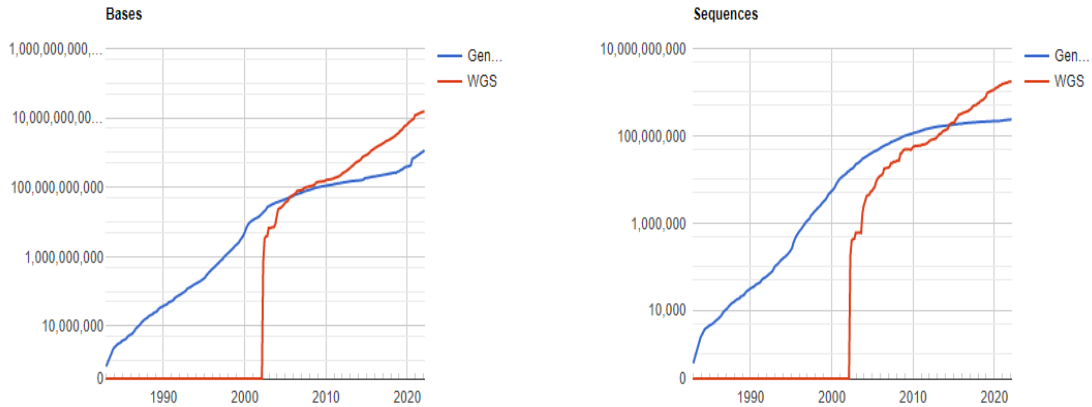
Bilgiye erişimi kolaylaştıran veri tabanlarından bazıları aşağıda belirtilmiştir:

1. Bibliyografik Veri Tabanları: MEDLINE, PUBMED.
2. Taksonomik Veri Tabanları: TÜBİTAK.
3. Nükleotid Dizisi Veri Tabanları: BLASTN, BALSTX.
4. Genomik Veri Tabanları: NCBI, UCSC, ENSEMBL.
5. Protein Dizisi Veri Tabanları: SWISSPROT, PDB.
6. DNA Veri Bankaları: GenBank (USA), EMBL (Europe), DDBJ (Japan)



Şekil 4. NCBI (Rashidi, 1999)

NCBI'dan alınmış Genbank ve Whole Genome Shotgun (WGS) veri tabanlarına ilişkin değerlerin ifade edildiği grafikler, bir örnek olarak aşağıda en güncel şekliyle verilmiştir:



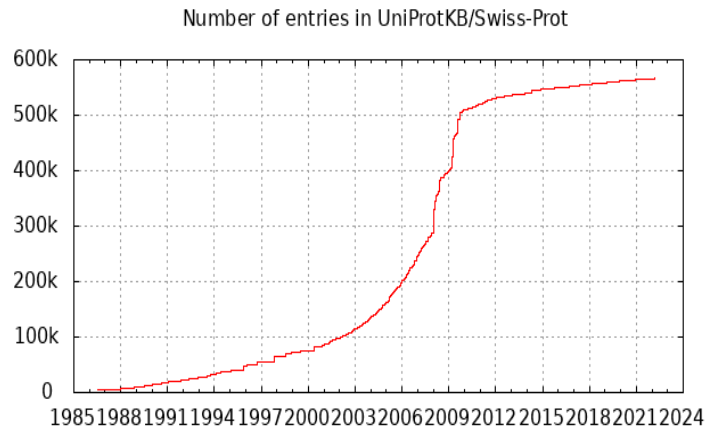
Şekil 5. GenBank ve WGS istatistikleri (www.ncbi.nlm.nih.gov)

UniProt, birçok varyantı bir arada göstermesi bakımından dünyada en yaygın kullanılan protein veri tabanlarından biridir. Bir örnek olarak; The UniProt Knowledgebase (UniProtKB)/Swiss-Prot veri

tabanlarına ilişkin veri girişlerinin yüzdesinin belirtildiği tablo ile bu girdilerin verildiği grafik aşağıda gösterilmiştir:

Tablo 1. UniProtKB/Swiss-Prot'taki veri girişleri

	Protein Varlığı	Girişler %
Protein düzeyindeki bulgular	109045	19,2
Transkript düzeyindeki bulgular	56025	9,9
Homolojiden çıkarılan sonuçlar	386901	68,2
Tahmin edilenler	13180	2,3
Belirsizler	1845	0,3



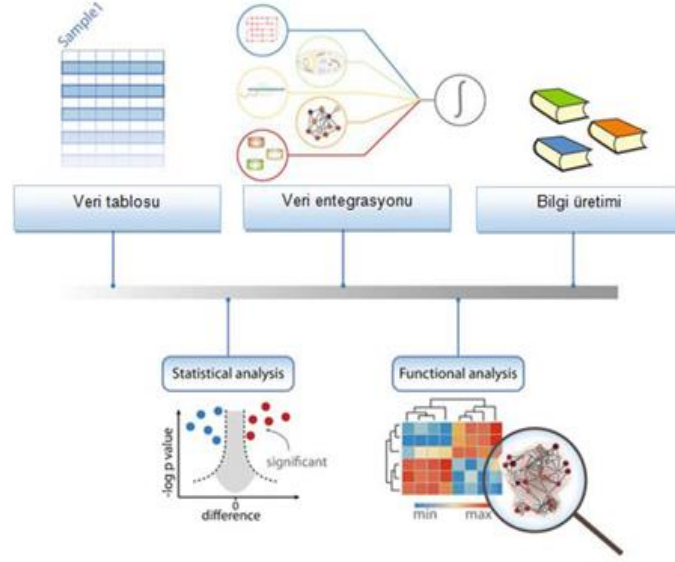
Şekil 6. UniProtKB/Swiss-Prot'taki veri girişlerinin grafiği (www.expasy.org)

Biyoinformatiğin Diğer Disiplinlerle İlişkisi

İstatistik ve bilgisayar bilimleriyle uğraşan araştırmacılar biyoinformatiğe önemli katkılarda bulunmaktadır. Verilerin yönetimi, depolanması, analizi ve yorumlanmasını sağlayan biyoinformatik, biyoistatistikten önemli bir derecede yararlanmaktadır. Fonksiyonel genomik verilerin kullanılabilir bilgi haline dönüştürülmesi biyoistatistik ile gerçekleştirilmektedir (Arhipova, 2006; Karabulut ve Karaağaoğlu, 2010). Elverişli sonuçlar elde edebilmek için moleküler biyoloji, genetik, tıp bilimleri, biyoistatistik ve bilgisayar bilimleri arasında güçlü bir ilişki olması gerekmektedir (Polat ve Karahan, 2009).

Biyoinformatik, moleküler biyolojiden oluşan verilerin analiz edilmesine imkân tanıyan matematik ve mühendislik gibi alanlardaki teori ve uygulamalardan faydalanan multidisipliner bir disiplindir. Örneğin kanser, genetik mutasyonlardan kaynaklanan bir hastalık olup insan kanserinin ilerlemesine neden olan mekanizmaların nasıl bir anlam ifade ettiği anlaşılamamıştır (Stransky ve Galante, 2009). Biyoinformatik, omik teknolojileri ve hesaplamalı biyoloji yardımıyla kanser araştırmalarında teşhis ve tedaviyi kolay bir hale getirmiştir (Wu ve ark., 2012).

Veri analizinin amacı, proteomik veriyi, klinik kararlara rehber olacak şekilde yapılandırarak alana has bilgilere dönüştürmektir. Bir istatistiksel analize başlamadan evvel, farklı örnekleri karşılaştırabilmek ve verilerin hepsini var kılan yöntemleri kullanabilmek için veri temizliği yapılması gerekmektedir (Tyanova ve ark., 2018). Kanser biyoinformatiği olarak proteomik, protein tanımlama, protein seviyelerinin ölçümü, protein-protein etkileşimi ve enzimatik aktivite üzerine araştırmalar yapmaktadır (Stransky ve Galante, 2009).



Şekil 7. Perseus'ta Tipik Bir Analiz İş Akışının Taslağı (Tyanova ve ark., 2018). Perseus, omik verilerin analizine olanak tanıyan bir yazılımdır (Tyanova ve ark., 2016).

Biyoinformatiğin Geleceği

Biyoinformatik çalışmalarının asıl amacı, çeşitli analiz yöntem ve araçlar geliştirerek kitle halindeki dizi verilerinden anlamlı küçük bilgiler elde etmektir. Çok sayıda yapısal bilgi ve evrimsel ipuçları sayesinde biyoinformatiğin önünde aşılması zor fakat imkansız olmayan hedeflere ulaşılmasının, gelecekte biyoinformatik alanında yapılacak çalışmalar açısından önemi büyüktür (Atalay, 2002). Biyoinformatik alanındaki çalışmalar sayesinde elde edilen kazanımların, hastalıklara tanımlar konulması ve bu hastalıklara yönelik tedavi yöntemlerinin geliştirilmesi bakımından kayda değer bir gelişimi söz konusudur. Bilim ve teknolojiadaki gelişmeler, biyoinformatik yazılımların gelişmesinin ve gelecekte uygulama alanlarının artacağına habercisi durumundadır. Günümüzde üniversiteler, bu alanda yetişmiş personellere ihtiyaç duymaktadır ve gelecekte de bu ihtiyacın devamına yönelik sinyaller akademik çevrelerde kendisini hissettirmektedir.

Geniş ölçekli işlevsel genom ve proteom araştırmaları sonucunda, yakın gelecekte hücresel olaylar hakkındaki bilgimiz artacaktır (Gribskov, 1999). Bu verileri depolayacak, birleştirecek ve bu verilere erişimi sağlayacak araçların geliştirilmesi önemli bir ihtiyaçtır. Herhangi bir sinyal yolağının biyolojik açıdan ne anlam ifade ettiğinin tespiti sayesinde ilaçların keşfi, genetik tanı kitleri ve gen tedavi yöntemlerinin geliştirilmesi mümkün olabilecektir (Atalay, 2002).

Sonuç

Teknolojinin günden güne gelişmesiyle birlikte biyoinformatik alanındaki kazanımların genişletilmesi kaçınılmaz bir hal almıştır. Bu sebeple biyoinformatik biliminin geliştirilmesi için yalnızca tıp doktorlarının değil, mühendislerin de katkısına ihtiyaç duyulmaktadır. Bilgisayar mühendislerinin geliştireceği bir yazılım, genler ve işleyiş biçimleri bakımından toplumu bilgilendirebilecektir. DNA haritalarından elde edilen bilgilerin kullanımını sağlayan veri tabanları aracılığıyla birtakım yeni istatistiksel yöntemlerin geliştirilmesi mümkün olacaktır (Eser ve Öngel, 2014).

Son zamanlarda artan veri artışı ile birlikte biyoinformatiğin bilim dünyasındaki yeri çok daha önem kazanmış olup biyolojik çalışmaların gelişmesi açısından (yapısal biyoloji, genomik çalışmalar ve gen ekspresyon çalışmaları, ilaç tasarlama, hastalığın tanı ve tedavisi) ilerleme kaydedilmiştir (Akın ve ark., 2014).

Biyoinformatik alanındaki araştırmalar, genetik bozukluğu olan kişilerin DNA yapılarının saptanmasıyla birlikte hastalığın karakteri hakkında bilgi edinmeyi, DNA özellikleri belirlenmiş genetik bozukluklardaki ailesel geçişlerin incelenmesini ve bu sayede hastalıklı bireylerin doğmasının önüne geçilmesini kolaylaştırmıştır. Kadın doğum kliniklerinde yapılan rutin genetik tarama testleri bu alandaki ilerlemeye iyi bir örnektir (Eser ve Öngel, 2014).

Sonuç olarak biyoinformatik alanındaki gelişmeler, gelecekteki araştırma ve çalışmalara yol gösterici olması açısından büyük bir önem taşımaktadır. Multidisipliner bir bilim dalı olan biyoinformatik, çeşitli yöntem ve yaklaşımların geliştirilmesi bakımından farkındalık oluşturmaya ihtiyaç duymaktadır.

Çıkar Çatışması Beyanı

Makale yazarları herhangi bir çıkar çatışması olmadığını beyan eder.

Araştırmacıların Katkı Oranı Beyan Özeti

Yazarlar makaleye benzer oranda katkı sağlamış olduğunu beyan eder.

Kaynakça

Akın AC., Bürçe B., Çevirici B., Şahin B., Şahin E., Şahin Y. Disiplinler arası bir bilim dalı: biyoinformatik. 2014. <https://paperzz.com/doc/5078763/p12.-disiplinler-aras%C4%B1-bir-bilim-dal%C4%B1--biyoinformatik>

Arhipova I. The role of statistical methods in computer science and bioinformatics. Latvia University of Agriculture, Latvia 2006; ICOTS-7.

Atalay RÇ. Neden biyoinformatik? Avrasya Dosyası. Moleküler Biyoloji ve Gen Teknolojileri Özel 2002; 8(3): 129-141.

Attwood TK. eind Parry-Smith, DJ: Introduction to bioinformatics, 1999.

- Chen JN., Chang JS. Topical clustering of MRD sense based on information retrieval techniques, *Computational Linguistics* 1998; 24(1): 61-95.
- Collins FS., Morgan M., Patrinos A. The human genome project: lessons from large- scale biology. *Science* 2003; 300(5617): 286-290.
- Dibyajyoti S., Bin ET., Swati P. Bioinformatics: The effects on the cost of drug discovery. *Galle Medical Journal* 2013; 18(1): 44-50.
- Drews J. Drug discovery: a historical perspective. *Science* 2000; 287(5460): 1960-1964.
- Eser U., Öngel K. *Güncel Biyoinformatik Yaklaşımlar*, 2014.
- Gentleman RC., Carey VJ., Bates DM., Bolstad B., Dettling M., Dudoit S., Zhang J. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* 2004; 5(10): 1-16.
- Gribnikov M. The new biological literature. *Bioinformatics (Oxford, England)* 1999; 15(5): 347-347.
- Hogue CW. Bioinformatics for beginners. *Trends in Biochemical Sciences* 2002; 10(27): 542-543.
- Karabulut E., Karaoğlu E. Biyoinformatik ve biyoistatistik. *Hacettepe Tıp Dergisi* 2010; 41, 162-170.
- Kumaresan V., Bhatt P., Palanisamy R., Gnanam A., Pasupuleti M., Arockiaraj J. A murrel cysteine protease, cathepsin L: bioinformatics characterization, gene expression and proteolytic activity. *Biologia* 2014; 69(3): 395-406.
- Kuonen D. Challenges in bioinformatics for statistical data miners. *Bulletin of the Swiss Statistical Society* 2003; 46, 10-17.
- Luscombe NM., Greenbaum D., Gerstein M. What is bioinformatics? A proposed definition and overview of the field. *Methods of Information in Medicine* 2001; 40(04): 346-358.
- Min S., Lee B., Yoon S. Deep learning in bioinformatics. *Briefings in Bioinformatics* 2017; 18(5): 851-869.
- Polat M., Karahan A. Multidisipliner yeni bir bilim dalı: Biyoinformatik ve tıpta uygulamaları. *SDÜ Tıp Fakültesi Dergisi* 2009; 16(3): 41-50.
- Rashidi HH., Buehler LK. *Bioinformatics basics: Applications in biological science and medicine.* CRC press, 1999.
- Searls DB. Using bioinformatics in gene and drug discovery. *Drug Discovery Today* 2000; 5(4): 135-143.
- Stransky B., Galante P. Application of bioinformatics in cancer research. *An Omics Perspective on Cancer Research* 2009; 211-233. https://doi:10.1007/978-90-481-26750_12
- Tandon P., Bhattacharjee P. *Bioinformatics: An overview.* Plant Tissue Culture and Molecular Markers, 2009.
- Tanyolaç B., Kaya HB., Soya S., Akkale C. *Biyoteknoloji ve biyoinformatik*, 2012.
- Blundell TL., Sibanda BL., Montalvo RW., Brewerton S., Chelliah V., Worth CL., Harmer NJ., Davies O., Burke D. Structural biology and bioinformatics in drug design: opportunities and

challenges for target identification and lead discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2006; 361(1467): 413-423.

Tyanova S., Temu T., Sinitcyn P., Carlson A., Hein MY., Geiger T., Mann M., Cox J. The perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 2016; 13(9): 731–740. <https://doi.org/10.1038/nmeth.3901>

Tyanova S., Cox J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Cancer Systems Biology* 2018; 133148. https://doi:10.1007/978-1-4939-7493-1_7

Wu D., Rice CM., Wang X. Cancer bioinformatics: A new approach to systems clinical medicine. *BMC Bioinformatics* 2012; 13(1): 71. <https://doi:10.1186/1471-2105-13-71>

Xia X. Bioinformatics and drug discovery. *Current Topics in Medicinal Chemistry* 2017; 17(15): 1709-1726.

Xiong J., Rayner S., Luo K., Li Y., Chen S. Genome wide prediction of protein function via a generic knowledge discovery approach based on evidence integration. *BMC Bioinformatics* 2006; 7, 268. <https://doi.org/10.1186/1471-2105-7-268>

www.ncbi.nlm.nih.gov

www.expasy.org