

## İNSAN GENOM PROJESİ

### ***Işık YULUĞ\****

*In this article, the author has summarized the Human Genome Project and what it aims to achieve. Following the description of the human genome the history, applications, aims, strategy and the ethical and legal issues of the project are discussed in detail together with a section that explains how the data will be analyzed.*

***I***nsan genom projesi biyoloji ve biyotıp bilimlerindeki en önemli projedir. Yüzyıllardır bilim dünyasını meşgul eden genetik bilginin içeriğinin anlaşılması 20. yüzyılın başında Mendel'in kalıtımla ilgili sunmuş olduğu prensiplerle başlamıştır. Genetik biliminde süregelen bilimsel gelişmeler kabaca dört yarım yüzyıla bölünebilir. İlk çeyrek yüzyıl kalıtımın hücresel temelini kurması, ikinci çeyrek yüzyıl kalıtımın moleküler temelini atması, üçüncü çeyrek yüzyıl kalıtımın bilgisel temelini çözülmesi, hücrelerin genlerdeki bilgiyi nasıl okuduğunu gösteren biyolojik mekanizmaların bulunması, rekombinant DNA teknolojisinin gelişmesi ile genlerin klonlanması ve dizilerinin belirlenmesi gibi önemli gelişmeleri içermektedir. Son çeyrek yüzyıl ise önce ilk bulunan genlerin daha sonra da tüm genomun deşifre edilmesi ve genomik biliminin gelişmesine tanıklık etmiştir. Tüm insan genom dizisinin belirlenmesi fikri ilk defa 1984-1986 arasında Amerika Birleşik Devletleri'nde Enerji Bakanlığında (DOE) bilimsel toplantılarda tartışılmıştır. Enerji Bakanlığı ve Ulusal Sağlık Enstitüsü (National Institutes of Health) tarafından yönetilen ve 1990 yılında 15 yıllık bir süre için başlatılan projenin ilk beş yıl için plânlanan bölümü (1993-1998) tamamlandıktan sonra 1998-2003 yıllarını kapsayan ikinci bir beş yıllık plân yapılmıştır. Proje, ABD (NIH, DOE), Avrupa Birliği Ülkeleri, Rusya, Japonya, Kanada, Çin, Avustralya gibi ülkelerin katılımı ve elde edilen verilerin paylaşılması sonucunda hızla ilerlemiştir. Uluslar arası işbirliği ile yürütülen bu projeye Amerika Birleşik Devletleri'nde özel bir kuruluş olan Celera Genomics firması da bağımsız olarak katılmış ve ileri teknolojik gelişmeleri kullanarak projeye hız kazandırmıştır. Uluslar arası Konsorsiyum ve Celera Genomics şirketi arasındaki rekabet sayılabilecek bir süreç sonunda 2001 yılının Şubat ayında insan DNA dizisinin o güne kadar elde edilen sonuçları bir basın toplantısı ile tüm dünyaya duyuru-

\* Doç. Dr., Bilkent Üniversitesi, Moleküler Biyoloji ve Genetik Bölümü

Avrasya Dosyası, Moleküler Biyoloji ve Gen Teknolojileri Özel, Sonbahar 2002, Cilt: 8, Sayı: 3, s. 7-23.

rulmuştur. Temel bilimlerin en önde gelen dergilerinden biri olan Nature dergisinin Şubat ayında yayınlanan özel sayısında Amerika Birleşik Devletleri'nden yirmi grup ve İngiltere, Japonya, Fransa, Almanya ve Çin devletlerinin de içinde bulunduğu "Uluslar arası İnsan Genom Dizi Belirleme Konsorsiyumu" adı altında "İnsan Genomunun İlk Dizileri ve Analizleri" başlığıyla genom projesinde kullanılan yaklaşımlar ve bugüne kadar elde edilen sonuçlar özetlenmiştir. Gene 2001 yılının Şubat ayında Science dergisinde büyük bir çoğunluğunu Celera Genomics şirketinin araştırmacılarının oluşturduğu "İnsan Genom Dizisi" başlığı altında bu şirketin insan genomu dizisini belirlemek için kullandığı yaklaşım ve dizi analizlerinin sonuçları özet olarak yayınlanmıştır.

İnsan genomu insan gelişimi, fizyolojisi, tıp ve evrim hakkında olağanüstü bilgileri bulundurmaktadır ve insan bugüne kadar bu kadar kapsamlı olarak genom dizisi belirlenen ilk omurgalıdır.

### **GENOM NEDİR?**

Genom bir organizmanın tüm DNA'sıdır. Her organizma yaşamını oluşturmak ve devam ettirmek için gereken biyolojik bilgiyi kapsayan bir genom içerir. Hücreler her canlının temel birimleridir ve işlevlerini deoksiribonükleik asit (DNA) kapsamındaki bilgiler belirler. Genom çoğunlukla DNA yapısında olmakla birlikte bazı viruslar RNA'yı genetik materyal olarak bulundurlar. Genom içinde bulunan bilgi DNA veya RNA moleküllerinin nükleotid dizilerinde kodlanmış olarak bulunur ve "gen" adı verilen birimlere bölünmüştür.

İnsanlarda ve birçok yüksek organizmada bir DNA molekülü birbiri etrafına sıkıca dolanmış iki iplik yapısındadır ve her bir iplik adenin (A), guanin (G), sitozin (C), ve timin (T) adlı dört bazın sıralanmasıyla oluşmuştur (Şekil 1). Bu dört bazın genomdaki sırası büyük önem taşımaktadır ve her üç bazın biraraya gelerek oluşturdukları kombinasyonlar değişik aminoasitlerin kodlanmasını sağlayarak hücrenin işlevlerini kontrol eden proteinlerin oluşmasını sağlarlar. Genin içindeki bilgiler genomun belirli bölgelerine bağlanan proteinler aracılığı ile okunur ve bu proteinler "gen ifadesi" (gene expression) diye adlandırdığımız bir seri biyokimyasal reaksiyonları başlatırlar.

Genom materyali DNA olan organizmalarda bu olay iki aşamada gerçekleşir:

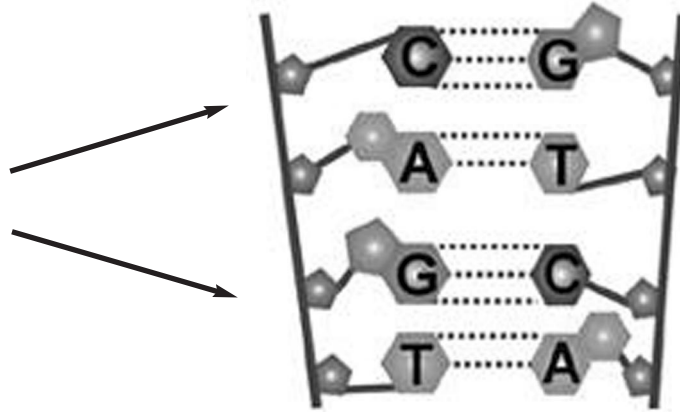
Transkripsiyon; genin RNA kopyasının çıkartılmasıdır.

Translasyon; RNA transkriptindeki amino asit dizisini belirleyecek olan genetik kodun kullanılarak proteinin sentezlenmesidir.

Bu tanım, bakteri gibi basit organizmalardaki gen ifadesinin tanımını vermemekle beraber, memeli gibi yüksek organizmalarda genom bilgisinin işlevsel proteinlere çevrilmesinde yer alan olayların karmaşıklığını tam olarak tanımlamaktadır. Gen ifadesindeki anahtar noktalar bu bilginin akışı sırasındaki kontrol noktalarıdır.

Her hücrenin bölünmesinde genomun tamamı kopyalanmalıdır. DNA replikasyonu adını verdiğimiz bu olay genom kopyalarında mutasyonların oluşmasını engellemek için çok doğru yapılmalıdır.

DNA replikasyonu sırasında veya yapısında direkt olarak değişiklikler yapabilen kimyasal veya fiziksel mutajenlere maruz kaldığı zaman bazı mutasyonlar oluşabilir. DNA tamir enzimleri bu yanlışlıkların çoğunu düzeltmekle görevlidirler. Tamir mekanizmasından kaçanlar ise bu mutant genomdan orijin olarak gelecek nesillere aktarılırlar. Genom değişimi (rearrangement) ile birlikte olan "rekombinasyon" ve mutasyonlar "moleküler evrimi" oluştururlar ve yaşayan canlıların evriminin arkasındaki temel olaylardır.



Şekil 1: DNA dört çeşit harfle (kimyasal molekül) yazılmıştır.

### İNSAN GENOMU

Bilinen tüm genomlar içinde bizi en çok insan genomunun ilgilendirmesi doğal görülmelidir. "İnsan Genom Projesi"nin ana hedefi insan genomunun yapısı ve organizasyonunu ortaya çıkarmak, buna paralel olarak diğer model organizmaların genomlarını çalışarak benzer veya farklı genleri belirlemek ve biyolojik işlevlerin nasıl çalıştığına açıklık getirmektir.

## İNSAN GENOMUNUN FİZİKSEL YAPISI

İnsan genomu iki farklı kısımdan oluşur:

**Çekirdek genomu:** İnsan çekirdek genomu yaklaşık üç milyar baz çiftinden oluşan DNA'dır. Değişik organizmaların genomunun büyüklüğü farklıdır. İnsanda çekirdek genomu kromozom adı verilen 24 DNA'ya bölünmüştür. Her kromozom 50 ile 250 milyon baz çifti kapsamaktadır (Tablo 1). En uzun kromozom yaklaşık 250 Mb, en kısıması yaklaşık 55 Mb'dır. 24 kromozomun 22'si otozomal kromozomlar, diğer ikisi ise seks kromozomları X ve Y dir. İnsan kromozomlarının yaklaşık uzunlukları Tablo 1'de gösterilmiştir.

**Mitokondri genomu:** Dairesel bir DNA molekülüdür ve 16,569 baz çifti içerir. Enerji metabolizmasından sorumlu olan bir mitokondri organelli çok sayıda mitokondri DNA'sı bulundurur.

Yetişkin insan vücudunda bulunan yaklaşık 1013 hücre kendi genom kopyasını bulundurur. Eritrositler gibi tamamen farklılaşmış kırmızı kan hücreleri ise DNA bulundurmazlar.

Hücrelerin çoğu diploiddir. Yani her otozomal kromozomdan iki kopya ve XX veya XY iki seks kromozomu olmak üzere toplam 46 kromozom bulundurlar. Bu hücrelere somatik hücreler adı verilir.

Seks hücreleri veya gametler ise haploiddir. Yani her otozomal kromozomdan bir kopya ve bir seks kromozomu olmak üzere 23 kromozom bulundurlar.

Bir somatik veya seks hücresinde yaklaşık her mitokondri organelinden 10 veya daha fazla olmak üzere toplam 8,000 kopya mitokondriyel genom bulunur.

Genler görünümümüzü, boyumuzu, cinsiyetimizi ve vücudumuzun değişik etkilere nasıl tepki vereceğini belirler ve organizmanın beslenmesini, görünümünü, enfeksiyonla savaşmasını sağlayan gerekli proteinlerin yapımı ile ilgili bilgileri taşırlar. Dünyada altı milyar insan olmasına rağmen hepsinin birbirinden farklı olmasını sağlayan genlerdir. İnsan genlerinin büyüklüğü çok değişkendir ve sıklıkla bir gen binlerce baz üzerine dağılmıştır. Genomun sadece %2'sinin protein kodlama ile ilgili olduğu düşünülmektedir ve bu bölgelere ekzon adı verilir, geri kalan kısım ise kodlama yapmayan bölgelerdir (genlerin ifadesini kontrol eden diziler ve geniş bölgeler) ve işlevleri tam bilinmemektedir.

### İNSAN GENOM PROJESİNİN HEDEFLERİ NELERDİR?

Moleküler yaşam bilimi adı altında topladığımız moleküler biyoloji ve genetik, biyokimya, hücre biyolojisi ve fizyolojinin gelişmesindeki kilit nokta genom dizisinin bilinmesidir. Yeni bir yüzyılın başlangıcında moleküler biyolojinin en önemli hedefi, bir çok değişik genomun DNA dizisinin tam olarak çıkarılmasıdır. İnsan Genom Projesi'nin ana amacı insan genomundaki yaklaşık 3 milyar DNA bazını ve yaklaşık 30-40 bin olduğu tahmin edilen genleri tanımlamaktır. Genom projesinin başarıya ulaşması için izlenen yoldaki ana başlıklar şu şekildedir: Genetik Harita, Fiziki Harita, Transkripsiyon Haritası, DNA Dizi Analizi, Model Organizmalar, Biyoinformatik, Etik, Hukukî ve Sosyal Boyutlar, Araştırma, Eğitim, Teknoloji Geliştirme ve Transfer.

Projenin plânlanan hedefleri içinde en başta insan DNA dizisinin tamamlanması ve araştırmacılara yüksek kalitede genom dizisinin sunulabilmesi gelmektedir.

İstenilen amaca ulaşmak için araştırmacılar öncelikle genomun genetik ve fiziksel haritasını çıkartmışlardır. Genetik harita belli genlerin belli bir kromozom üzerindeki yerlerinin birbirlerine olan göreceli mesafelerini belirlemek ve genleri lineer bir düzene oturtmak amacıyla hazırlanan grafiksel bir haritadır. Fiziksel haritalama genomik DNA'nın tümünü parçalara ayırarak kozmid, maya yapay kromozomları (YAC) veya bakteri yapay kromozomları (BAC) gibi vektörlere aktararak her kromozomun kitaplığının oluşturulması ve tüm kromozomlar düzeyinde birbirini takip eden klonların belirlenmesi ile oluşturulan bir haritadır. Bu ha-

**Tablo 1: İnsan Kromozomlarının Yaklaşık Uzunlukları**

<b>Kromozom Uzunluk</b>	<b>(Mb=1000 kilo baz)</b>
1	263
2	255
3	214
4	203
5	194
6	183
7	171
8	155
9	45
10	144
11	144
12	143
13	114
14	109
15	106
16	98
17	92
18	85
19	67
20	72
21	50
22	56
X	164
Y	59
<b>Toplam</b>	<b>3286</b>

ritalamaların hazırlanması genomdan ifade edilen genlerin yerlerinin tespit edilerek transkripsiyon haritası çıkarılması işlemini çok kolaylaştırmıştır.

Projenin ilerleme aşamalarında teknolojik gelişmeye en çok ihtiyaç duyulan kısım DNA dizisinin belirlenmesi ve sonuçlarının analizi için kullanılan araçlar ve yöntemler olmuştur. Genom dizisinin belirlenmesi geliştirilen otomatik makineler ve robot sistemleri sayesinde çok hızlı bir şekilde yapılabilmekte ve belirli merkezler her yıl milyonlarca baz çiftinin sırasını ortaya çıkarabilmektedirler. Ayrıca, bu bilgiler veri bankalarında toplanmakta ve yeni bilgisayar ve bilgisayar yazılım programlarındaki gelişmeler sayesinde bu verilerin hızla değerlendirilmesi yapılabilmektedir.

Bu çalışmalar insan genomu için devam ederken diğer yandan evrimsel olarak değişik türleri temsil etmeleri nedeniyle seçilen model organizmaların genomlarının da haritaları oluşturulmuştur. Bu organizmalar: Memeli *Mus Musculus* (Fare), Omurgasız *Drosophila Melanogaster* (meyve sineği), Nematod *Caenorhabditis Elegans* (kurtçuk), Protist *Saccharomyces Cerevisiae* (Maya), Bakteri *Escherichia Coli*, Bitki *Arabidopsis Thaliana*, ve Diğer Omurgalılar grubundan *Pufferfish* (*Fugu Rubripes Rubripes*) ve *Zebrafish*'tir (Tablo 2). Bugüne kadar maya, kurtçuk, meyva sineği ve birçok bakteri ve virus genom dizileri tamamen belirlenmiştir.

Son olarak da projenin sonuçlarının getireceği etik, hukukî ve sosyal konuların değerlendirilmesi proje kapsamındadır.

### **KİMİN GENOMUNUN DNA DİZİSİ BELİRLENİYOR?**

Tüm insanlar aynı temel genleri ve genom düzenleme bölgelerini kullanmaktadırlar. Bu nedenle referans olarak alınacak dizinin tek bir insanın genomunun tam bir kopyası olması gerekmemektedir. Birçok kan ve sperm örneği toplanmış olmakla beraber kimlikleri örnekleri veren veya örnekleri çalışan bilim insanları tarafından bilinmeyen 20 kadar bireyin genom örnekleri esas alınarak çalışılmaktadır. Araştırmacılar çok değişik popülasyonları temsil eden vericilerden DNA ile çalışmaktadırlar. Örneğin Celera Genomics İspanyol, Asya, beyaz ve Afrika-Amerika kökenli kişilerden kan almıştır. Dizi analizlerinde kullanılan gen kütüphaneleri anonim vericilerden alınan kandan oluşturulmuştur. Kişilere bir genetik danışman açıklama yaptıktan sonra imzaları alınmış ve alınan kanlar daha sonra birkaç kez değişik şekillerde rasgele olarak numaralarla etiketlenip tüm kayıtlar yok edilerek kanların kimden geldiğinin bilinmesi imkânsız hâle getirilmiştir.

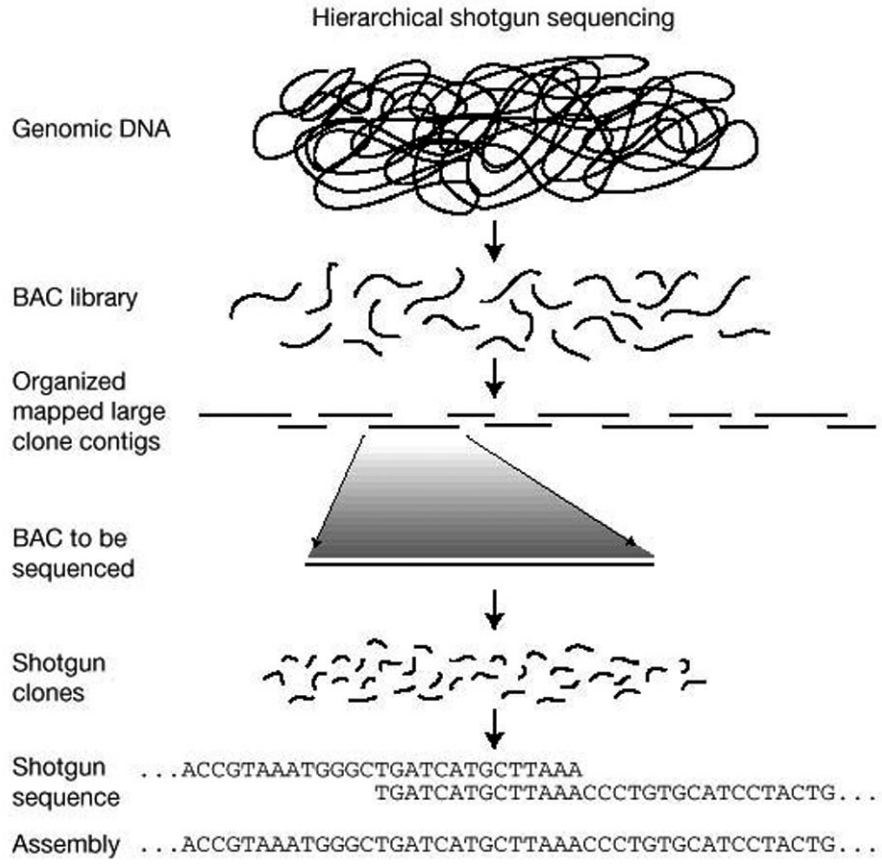
### **DNA DİZİSİNİN BELİRLENMESİNDE KULLANILAN STRATEJİ**

İnsan genom DNA'sı küçük parçalara bölündükten sonra çift iplikli DNA'lar her iki ucundan dizi analizine tâbi tutulmuştur. Sonraki aşamada bu parçalar bir araya getirilmiş ve aralarında açık kalan diziler belirlenmiş ve kalan belirsizlikler yönlendirilmiş analizler ile sonuçlandırılmaya çalışılmıştır. Bu yaklaşım sadece genlerin değil gen düzenleyici bölgeler ve kromozomun da yapısı hakkında bilgiler vermiştir. DNA'yı parçalamak ve klonlamak için protokollerin giderek gelişmesi "shotgun" adı verilen bu yöntem ile genomun daha homojen parçalar olarak elde edilebilmelerini sağlamıştır. Celera Genomics insan genomunu tüm genom shotgun yaklaşımıyla dizilemiştir. Ayrıca Uluslar arası İnsan Genom Dizileme Konsorsiyumu'ndan elde edilen hiyerarşik shotgun verilerini de kullanmışlardır (Şekil 3).

Genom dizisinin ilk öncü dizileri oluşturulurken bilim adamları her kromozomal bölgedeki baz çifti sırasını en az dört kere tanımlayarak datanın doğru olmasını sağlamışlardır. Genomun tekrar tekrar dizilmesine "depth of coverage" adı verilmektedir. Öncü dizi datası genel olarak 10 bin baz çifti büyüklüğünde ve kromozomal yeri yaklaşık olarak bilinen parçalardan oluşmaktadır.

İnsan Genom Projesi için bitmiş dizi için kabul edilen hata oranı her 10 bin bazda tek bir hatadır. Yüksek kalitede elde edilen dizi insan biyolojisinin tam olarak anlaşılabilmesi ve kalp hastalıkları, kanser ve diyabet gibi durumlara yatkınlığa neden olan kombinasyonların çözülebilmesi için gereklidir. Bitirilmiş durumda her kromozomun DNA dizisi 8-9 kez taranmış olacaktır.





Şekil 2: Genom parçalarını klonlama ve dizi analizi için izlenen yol



## ANALİZ SONUÇLARI

Şubat 2001'de İnsan Genom Projesi ve Celera Genomics DNA dizisinin o güne kadar elde edilen sonuçları yayınlanmıştır. Yayınlanan kısım insan genomunda genlerin çoğunluğunu bulandıran kısmın %90'ını kapsamaktadır.

İnsan genomu daha önce dizisi belirlenen genomlardan yaklaşık 25 kat daha büyüktür. Genomda kodlanan genlerin insan DNA'sının sadece küçük bir bölümünü oluşturduğu belirlenmiştir; fakat, bu bölüm genomun biyolojik işlevlerini gören ve araştırmacıların esas olarak ilgilendiği kısımdır. İnsan genomundan kodlanan gen sayısının 30-40 bin civarında olduğu düşünülmektedir. Bu sayı solucan veya sinek genomunda bulunan gen sayısının sadece iki katı kadardır. Ancak insan genleri daha karmaşık yapıdadır ve bir gen birden fazla seçenekli kesim bölgesi bulundurabilmesi nedeniyle daha fazla sayıda protein ürünü verebilmektedir.

İnsan genomunda halen klonlanamayan geniş DNA bölgeleri vardır. Bu bölgeler kromozomların orta ve uç bölgelerinde bulunan DNA dizileridir. Bu bölgeler çok az sayıda gen bulandırmakta ve çok sayıda kendini tekrar eden parçalardan oluşmaktadır. Bu tekrar dizilerinin doğrudan bir işlevlerinin olmadığı ama genoma şekil vererek yeni genlerin oluşmasını veya mevcut genlerin değişmesini sağladıkları düşünülmektedir. Genom projesindeki esas amaç, genden zengin olan bölgelerin dizisinin belirlenmesi olduğu için Şubat 2001 yılında yayınlanan ilk genom dizisinde bu bölgeler pek yer almamaktadır. Projenin dizi analizinin bitirilmesi hedeflenen 2003 yılında bu bölgelerin dizilerinin belirlenmesi muhtemel görünmektedir.

Örneğin insan 22. kromozomunun %60'lık kısmının DNA dizisi belirlenmiş olmasına rağmen tamamen dizinlendiği bildirilmiştir. Çünkü 22. kromozomun %60'lık bir kısmı genden zengindir ve bu bölgenin de %97'si dizi analizi kapsamındadır. Benzer şekilde *Drosophila* için hemen hemen tüm genden zengin bölgelerin dizisi belirlenmiştir.

İnsan genomu tüm dokulardaki hücrelerde aynı olmasına rağmen yüzlerce değişik hücre tipinin gelişmesini sağlayabilmektedir. İnsan genomundan kodlanan proteinlerin bilinmesi insan genlerinin işlevlerinin çalışılması için çok önemlidir. Proteinlerin bir kısmı diğer organizmaların proteinlerine benzerlik göstermekle beraber daha karmaşık yapıdadır ve daha yüksek sayıdadır.

İnsan Genom Projesi'nin önde gelen amaçlarından biri kişiler arasında farklılık gösteren ve yaklaşık 1.5 milyondan fazla olduğu düşünülen tek nükleotid polimorfizlerinin (single nucleotide polymorphism (SNP)

tespit edilmesidir. DNA'daki nükleotid bazlarının %99.9 gibi bir kısmının dizisi tüm insanlarda aynıdır. Bireyler arasındaki farklılığı oluşturan DNA dizisindeki bu lokalize değişiklikler bireylerin pek çok hastalığa ve ya ilaca karşı farklı duyarlılık göstermelerine neden olmaktadır.

Bu güne kadar 5, 16, 19, 21 ve 22. insan kromozomlarının dizileri tamamlanmıştır.

Meme kanseri, kas hastalığı, sağırılık ve körlükle ilgili bazı genlerin yeri saptanmıştır. Ayrıca kardiyovasküler hastalık, diyabet, artrit ve kanser gibi sık görülen hastalıkların genlerinin bulunmasına çalışılmaktadır.

Model organizma olarak kullanılan Arabidopsis Thaliana bitkisi, Drosophila Melanogaster meyva sineği, Caenorhabditis Elegans solucanının tüm genomu belirlenmiştir. İnsan, fare ve pufferfish genomlarının tamamlanması için çalışmalar devam etmektedir.

### **İNSAN GENOM PROJESİ'NİN YARARLARI NELERDİR?**

İnsan Genom Projesindeki hızlı gelişmeler ve getireceği potansiyel uygulamalar göz önüne alındığında, biyolojinin 21. yüzyılın bilimi olacağına şüphe yoktur.

Genomdaki her gen dizisinin ve genin işlevlerini düzenleyen bölgelerin ne olduğunu gösteren bir katalogun bulunması, genlerin işlevleri bilinmese dahi son derece yararlıdır.

Genom dizisi, hücrelerin moleküler aktivitelerinin ve bu aktivitelerin nasıl kontrol edildiğinin tanımlanmasına yol açacaktır. Gen katalogu ayrıca insan kalıtsal hastalıklarından sorumlu olan genler veya protein ürünleri ve endüstriyel önemi olan bakteriyel genler gibi önemli genlerin izolasyonu ve kullanımı ile ilgili son derece önemli bir konuya açıklık getirecektir.

Aslında bu genler genomun tümü tanımlanmamış olsa bile izole edilebilir. Fakat bu işlem uzun zaman almakta ve her gen için farklı projeler düzenlenmektedir. Bu nedenle tüm genom dizisi bilinirse, istenilen genin elde edilmesi ve katalogdan bulunması çok daha kolay olacaktır.

### **GENOM ARAŞTIRMALARININ OLASI UYGULAMALARI**

Moleküler Tıp: İnsan genom araştırmalarındaki anahtar uygulama, hastalıklara neden olan genlerin ve işlevlerinin bulunmasıdır. Günümüzde 30'dan fazla hastalığa neden olan genler bilinmektedir. Son yıllarda genom dizi verilerin doğru bir şekilde veri tabanlarına ulaşması

sonucu bu sayının yakın bir zamanda hızla artacağından hiç şüphe yoktur. Hastalıkların erken tanısı ve tanımı kolaylaştırma, hastalıklara genetik yatkınlıkların tanımlanması, daha etkin ilaçlar ve kontrol sistemlerinin geliştirilip bu ilaçların kimlerde etkili olabileceğinin önceden belirlenmesi ve gen terapisi gibi uygulamalarda hızlı bir gelişim sağlanacaktır. Şimdiden myotonik distrofi, fragile X sendromu, nörofibromatosis, ailesel kolon kanseri, Alzheimer hastalığı ve ailesel meme kanseri gibi hastalıkların genleri hakkında birçok bilgi elde edilmiştir.

Genom araştırmalarının sonuçları ayrıca aşağıdaki amaçlarla kullanılabilir:

**Mikrobiyel genomik:** patojenlerin çabuk saptanıp tedavi edilmesi, yeni yakıt ve enerji kaynakları bulunması, hava kirliliği nedenlerinin saptanması, insanların kimyasal ve biyolojik ajanlara karşı korunması, toksik atıkların temizlenmesi, radyasyonun yaptığı hasarın ve mutajenik kimyasalların vücuda etkilerinin saptanması

**Antropoloji:** çağlar boyunca evrimin ve mutasyonların incelenmesi.

**DNA tanımlanması:** suçluların doğru ve çabuk olarak teşhis edilmesi, kazada veya savaşta ölenlerin tanımlanması, aile bağlarının ispatlanması, organ nakli için uygun kişilerin saptanması.

**Tarım ve hayvancılık:** hastalığa dirençli bitki ve hayvanların geliştirilmesi, verimin artırılması.

Proje sonuçlarının kullanılması sırasında en hassas konu toplumla ilgili uygulamalardır. Bu konuda cevap verilmesi gereken pek çok soru vardır. Genetik bilgi kime aittir? Genetik bilgiye kimler ulaşabilmelidir? Genetik bilgi kişinin toplumdaki yerini nasıl değiştirebilir? Doktorlar ve diğer sağlık personeli bu konularda nasıl eğitilebilir ve doğru kararı verdiklerinden emin olunabilir mi? Kimler bu teknolojiden yarar görecektir? Bir hastalığın tedavisi yoksa da o hastalık için test yapılmalı mıdır? Çocuklar ancak ileride ortaya çıkabilecek hastalıklar için testten geçirilmeli midir? Genlerimiz davranışlarımızı etkiler mi ve kontrol edilebilirler mi? Genetik olarak modifiye edilmiş yiyecekler insan ve çevre için sağlıklı mıdır? DNA dizilerinin patentlenmesi doğru mudur?

Genom dizisinin bilinmesinin şu anda tahmin edemediğimiz yararlarının olması mümkündür. Örneğin, insan genomu çok sayıda kodlanmayan DNA bölgesi içermektedir. Bu bölgelerin hiçbir işlevi olmadığı düşünülmekle beraber belki de bu düşünce bizim bilgimizin azlığından ileri gelmektedir. Bu kodlanmayan bölgelerin işlevi var mıdır? Bu sorunun yanıtını vermek farklı genomların kodlanmayan bölgelerinin tam bir tanımını yapmak ve ortak bilgileri bularak analiz etmekle yanıtlanabilir.

Genom projesinin yapılmasının bir başka nedeni de şu andaki teknolojinin limitlerini aşabilmek ve yeni teknolojileri bu çalışmanın getireceği şevkle geliştirebilmektir.

### **BIYOİNFORMATİK**

Genom projesiyle elde edilen bilgi çoğaldıkça bu bilginin değerlendirilmesi de önem kazanmaktadır. Bu kadar çok miktardaki bilgiyi depolamak ve yorumlamak biyoinformatik adlı bir bilim dalının ilgi sahasıdır. Bu bilginin mantıklı bir şekilde incelenebilmesi ve insan genom haritasının daha düşük organizmalarda başarıldığı gibi düzenli bir haritasının çıkarılabilmesi için çok kuvvetli bilgisayarlar ve yeni programlara ihtiyaç vardır. Genom veritabanı dünyanın her köşesinden genom haritalama bulgularının depolandığı yerdir. Bu veritabanına internet ile ulaşılabilmekte ve böylece tüm dünyadan bilim adamları gelişmeleri günü gününe izleyebilmektedirler. Tüm insan genomunun dizisi üç gigabyte büyüklüğündedir ve bu datanın yorumlanması bilginin kendisinin yüzlerce katı yer tutacaktır. Bir yandan büyük laboratuvarlar ve merkezlerde biriken bilginin hızla işlenebilmesi sağlanmalı, bir yandan da küçük merkezler ve bireysel araştırmacılara bu bilgiden hızlı bir şekilde yararlanabilmeyi sağlayacak araçlar hazırlanmalıdır. Ayrıca insan genomu ile diğer organizmaların DNA'larının bilgisayar programları ile karşılaştırılmaları bilgilerimize büyük katkıda bulunacaktır.

### **MODEL ORGANİZMALAR**

İnsana çok benzer genetik yapı taşıyan model organizmalar üreme ve çoğalma sürelerinin kısalığı sayesinde birçok nesli çabuk ve ucuz olarak inceleme olanağı vermektedir. Tam genomik dizisi belirlenen model organizmalardan bazıları *Escherichia Coli* (bakteri), *Saccharomyces Cerevisiae* (maya), *Drosophila Melanogaster* (meyve sineği) ve *Caenorhabditis Elegans*'dir (nematod).

Diğer canlı genomlarının bilinmesi insan genomuyla karşılaştırma yapılmasını ve insan genomunun daha iyi anlaşılabilmesini sağlayacaktır. Laboratuvar faresi üzerinde yapılan çalışmalar devam etmektedir. İnsan ve fare genomlarında yaklaşık olarak 3 milyon baz çifti bulunur ve fare genlerinin büyük bir çoğunluğunun insan genlerinde karşılığı vardır. İki tür arasında farklı olan genler sadece %1 – 5 arasındadır ve genelde daha önce var olan genlerin duplikasyonu veya değişmesi sonucu yeni işlev kazanmış genlerdir. Genomlar arasındaki bu benzerlik farenin popüler bir model organizma olarak kullanılmasında en önemli etkindir.

Farenin model organizma olarak kullanılması ve transgenik fare modellerinin geliştirilmesi bize genlerin memeli organizmalarda nasıl işlem görebilecekleri hakkında fikir vermesi açısından çok önemlidir. İnsanda hastalıklara neden olan genlerin bulunması ve bu genlerin hücrede ne gibi işlevler gördüğünün anlaşılması uzun ve zahmetli araştırmalar gerektirmektedir.

Tablo 2 insan ve diğer organizmaların genom büyüklüklerini karşılaştırmaktadır.

**Tablo 2: İnsan ve incelenen diğer organizmaların genomlarının tahmini büyüklükleri:**

İnsan	3000 milyon baz (~100,000 gen)
Fare	3000 milyon baz (50,000 - 100,000 gen)
Drosophila (meyve sineği)	165 milyon baz (15,000 - 25,000 gen)
C. Elegans (nematod)	100 milyon baz (11,800 - 13,800 gen)
Maya (mantar)	14 milyon baz (8355 - 8947 gen)
E. Coli (bakteri)	4.67 milyon baz (3237 gen)
H. Influenzae (bakteri)	1.8 milyon baz
M. Genitalium (bakteri)	0.58 milyon baz

### **ETİK, HUKUK VE SOSYAL SORUNLAR**

İnsan Genom Projesi, kısa zamanda gen yapısı ve organizasyonu, genetik çeşitlilik, karşılaştırmalı genom ve insan dizi verilerinin bilinmesiyle birlikte oluşacak etik, adli ve sosyal konular üzerinde belirgin gelişmeler kaydetmiştir.

Etik, hukuk ve sosyal konular İnsan Genom Projesi tarafından oluşturulmuş bir program çerçevesinde ele alınmaktadır (ELSI programı).

İnsanlar arasında DNA dizisi %99'dan fazla benzerlik gösterse de kişiler arasındaki küçük farklılıklar bireylerin hastalıklara, hastalık etkenlerine ve ilaçlara nasıl cevap verdiklerini belirleyebilmektedir. En sık görülen değişkenlik tipi olan ve 100 - 300 bazda bir görülen tek-nükleotid polimorfizmlerinin saptanması kanser, diyabet gibi hastalıklardan sorumlu genlerin saptanmasında önemli rol oynamaktadır. Ayrıca data toplandıktan sonra da gen ifadesinin ve kontrolünün incelenmesi, mutasyonların oluşturdukları problemlerin belirlenmesi için daha hızlı analiz metodlarının geliştirilmesi gerektirecektir. Elde edilecek bilginin kötü amaçlarla kullanılmaması için korunması çok önemli olacaktır.

Genom haritasının elde edilmesinin beraberinde getireceği birçok adli ve etik sorun arasında; kişisel bilgilerin güvenli bir şekilde saklan-

ması, genetik bilginin işe girerken engel oluşturması, kişilerin üreme hakkına karşılanması sayılabilir. Hastanın genetik profilinin bilinmesi büyük bir sorumluluktur. Bu bilgiler nasıl kullanılmalı veya kullanılmamalıdır? Birçok kişi bazı hastalıklar için yüksek risk taşıdıklarını öğrenmeyi tercih etmektedirler. Ayrıca kişilerin genetik olarak bazı hastalıklara yatkın oldukları için sağlık sigortasından mahrum bırakılmaları da söz konusu olabilir. BRCA1 genini taşıyan bir kadının meme kanserine yakalanma riski yüksek olsa da o kişide hastalığın nasıl seyredeceği ve seçenekler tam olarak belirlenmeden bu durumun hastaya açıklanması doğru mudur? Genetik bilgimize başkasının sahip olmaya hakkı var mıdır? DNA dizilerinin patentlenmesi doğru mudur?

ABD’de genler ve gen parçalarının patentlenmesi için yapılan başvurular çok artmıştır. Patentler içinde 3 milyon kadar EST dizisi ve binlerce genom parçası veya tam gen mevcuttur ve bu durum ciddi bir tartışmaya yol açmıştır. Genler ve diğer biyolojik yapılar 1980 yılından beri patentlenebilmekte ve patent başvurusu yapılmasından sonraki 20 yıl için patent sahibine kullanım hakkı tanınmaktadır. Biyolojik buluşların patentlenebilmesi için yeni, yararlı ve doğru şekilde tanımlanmış olmaları gerekmektedir. Ancak patentler başka araştırmacıların çalışmalarını engelleyebilmekte, ayrıca patent başvuruları gizli kaldığından firmalar uzun çalışmalar sonucunda buldukları genlerin başka patentler tarafından korunabildiğini çok geç anlayabilmektedir.

### **SON GELİŞMELER**

Genom datası her gün hızla büyümektedir ve bu datanın incelenmesi için çeşitli çalışmalar yapılmaktadır. Araştırmalar, “Transcriptomics” haberci RNAları incelemekte, “proteomics” proteinleri incelemekte, “structural genomics” proteinlerin üç boyutlu yapısını incelemekte, “knockout” çalışmaları canlı organizmalarda genleri inaktive ederek gen işlevini belirlemeye çalışmakta, “comparative genomics” ise insan ve başka organizmaların DNA’sını karşılaştırmaktadır. 2001’de başlatılan Mikrobiyel Hücre Projesi yaşamın değişik parçalarının nasıl birbiriyile uyum içinde çalıştığını ortaya çıkarmayı planlamaktadır. Parçaların canlı hücrelerde beraber nasıl işlev gördükleri ve çevresel değişikliklere nasıl uyum sağladıkları bilinmemektedir. Bu nedenle elde edilen verilerin en iyi şekilde yorumlanabilmesi için biyoloji, mühendislik, kimya, fizik, bilgisayar, görüntüleme ve işletme gibi disiplinler arası dalların beraber çalışmaları çok büyük önem kazanmaya başlamıştır.

İnsan genomunun belirlenmesinden sonra önemli olacak konular şunlardır:

- Gen sayısının, kromozomlar üzerindeki tam yerlerinin ve işlevlerinin belirlenmesi,
- Genlerin ifadelerinin nasıl kontrol edildiğinin belirlenmesi,
- DNA dizisinin organizasyonu,
- Kromozomal düzeyde yapı ve organizasyon,
- Kodlamayan DNA tipleri, miktarı, dağılımı, bilgi kapsamı ve işlevleri,
- Gen ifadesi, protein sentezi ve translasyon sonrası olayların koordinasyonu,
- Kompleks moleküler yapılarda proteinlerin birbirleri ile ilişkisi,
- Tahmin edilen ve deneylerle saptanan gen işlevlerinin karşılaştırılması,
- Organizmalar arasında evrimsel korunma,
- Protein yapısı ve işlevi,
- Organizmalarda proteomlar (total protein miktarı) ve işlevleri,
- SNP'lerin (bireyler arasında tek baz DNA değişiklikleri) sağlık ve hastalık ile ilişkisi,
- Gen dizilerindeki çeşitliliklere bakılarak hastalığa yatkınlığın tahmin edilmesi,
- Karmaşık karakterler ve çoklu gen hastalıklarında etkili genler,
- Çevresel düzenleme için yararlı mikrobiyel konsorsiyumlar dahil kompleks sistemler biyolojisi,
- Gelişim genetiği.

### **Özel sektör ve genom projesi**

İlaç şirketleri son yıllarda yeni ilaçlar geliştirmek için kısıtlı sayıda ilaç hedefi üzerinde çalışmaktadır ve piyasadaki tüm ilaçlar yaklaşık 500 hedefi etkilemektedir. İnsan gen ve proteinlerinin tam olarak bilinmesi hedef sayısını çok artıracaktır.

İnsan genom projesinin ilerlemesinde özel firmaların büyük katkısı olmuştur. IBM, Compaq, DuPont ve büyük ilaç firmaları çalışmalara yardımcı olmaktadır ve sonuçların doğru şekilde uygulanmasıyla büyük kazançlar beklemektedir.

İnsan Genom Projesi'nin amacı başından beri tam bir kaynak genom belirlemek olmuştur. Özel sektör firmaları ise sadece müşterilerinin



işine yarayacak kısmı yani genden zengin bölgeleri sekanslamaya çalışmaktadır. Firmaların beklentileri insanlar için daha çok tanı ve tedavi yöntemi, hayvanlar için daha sağlam ve besleyici gıdalar, yeni enerji kaynakları, çevreyi temizlemek için kullanılacak yeni mikroorganizmalar, zararlı ajanların daha kolay tanımlanıp direnç sağlanması ve insan, hayvan ve bitkilerin daha kolay tanımlanabilmesi gibi konulardır.

ABD Hükümetinin, İnsan Genom Projesi için ayırdığı bütçe 2001 yılı için 400 milyon dolardır. Özel firmaların bir konuyla ilgilenmesi için genellikle önce devlet sektörünün yatırım yapıp konunun ne kadar önemli olduğunu göstermesi gerekmiştir. Örneğin; Celera Genomic firmasının dizi analizini belirlemek için kullandığı Perkin-Elmer 3700 makinesinin optik sistemi devlet desteğiyle geliştirilmiştir. Buna karşılık bu firmalar devletin yapamadıklarını yapmakta ve yarattıkları vergiler de yine devlet yatırımlarına dönüşmektedir.

## SONUÇ

İnsan Genom Projesi, insanın karmaşık yapısının anlaşılmasında sadece bir basamaktır. Elde edilen bilginin değerlendirilmesi için on binlerce bilim insanının yıllarca çalışması gerekecektir.

Önümüzdeki yıllarda projenin tamamlanması ve bulguların değerlendirilmesiyle insanlık kendisini belirleyen genetik alfabeyi ilk kez tümüyle anlama olanağına kavuşacaktır. Projeden elde edilen bilgiler şimdiden yaşamımızı değiştirmektedir.

Genetik alfabenin tamamlanması ve genlerin işlevlerinin tanımlanması insanlık için çok önemli bir adım olacaktır. Genetik alfabenin sağladığı bilgiler genlerin kodladığı proteinlerin hücredeki işlevlerini açıklamak için gerekli olan analizleri yapmayı ve genin istenen bölgelerinde seçici olarak mutasyonlar oluşturmayı mümkün kılacaktır.

İlk dizi verileri, insan genomunda 3 milyondan fazla tek nükleotid polimorfizmi tanımlamıştır. Genomdaki tüm varyasyonların tanımlanması ile sık görülen hastalıklara yakalanma riskini artıran genler tanımlanabilirse, insanlık hastalığın önlenmesi yolunda önemli bir adım atmış olacaktır.

**KAYNAKLAR**

1. International Human Genome Sequencing Consortium, "Initial Sequencing and Analysis of The Human Genome," *Nature*, Sayı 409, 2001, ss. 860-921,
2. Venter, Craig J. et al, "The Sequence of the Human Genome," *Science*, Sayı, 291, 2001, ss. 1304-1351
3. Collins, F.S. and McKusick, V.A. "Implications of the Human Genome Project for Medical Science," *J. Am. Med. Assoc*, Sayı. 285, 2001, ss. 540-544.
4. Watson, J.D. & Crick, F.H.C., "Molecular Structure of Nucleic Acids: a Structure for Deoxyribose Nucleic Acid," *Nature*, Sayı 171, 1953, 737-738.
5. Green, E.D. "Strategies for The Systematic Sequencing of Complex Genomes," *Nature Rev. Genet.* Sayı 2, 2001, ss. 573-583.
6. Hubbard, T et al, "The Ensembl Genome Database Project," *Nucleic Acids Res.* Sayı 30, 2002, ss. 38-41.
7. Stein, L, "Genome Annotation: From Sequence To Biology," *Nature Rev. Genet.* Sayı 2, 2001, ss. 493-503.
8. Sherry, S.T. et al. "dbSNP: the NCBI Database Of Genetic Variation," *Nucleic Acids Res*, Sayı 29, 2001, ss. 308-311.
9. Apweiler, R. et al, « Interpro--An Integrated Documentation Resource for Protein Families, Domains And Functional Sites," *Bioinformatics*, Sayı 16, 2001, 1145-1150.
10. Chothia, C. Proteins , "One Thousand Families for The Molecular Biologist," *Nature*, Sayı 357, 1992, ss. 543-544.
11. Smith, R.F, "Perspectives: Sequence Data Base Searching in The Era of Large-Scale Genomic Sequencing," *Genome Res.* 6, 1996, ss. 653-660.

**YARARLI WEB SAYFALARI**

- \* Genome Database  
<http://www.gdb.org/>
- \* UK Human Genome Mapping Project Resource Center  
<http://www.hgmp.mrc.ac.uk/>
- \* The National Center for Biotechnology Information  
<http://www.ncbi.nlm.nih.gov/>
- \* US Human Genome Mapping Project  
<http://www.ornl.gov/hgmis/>