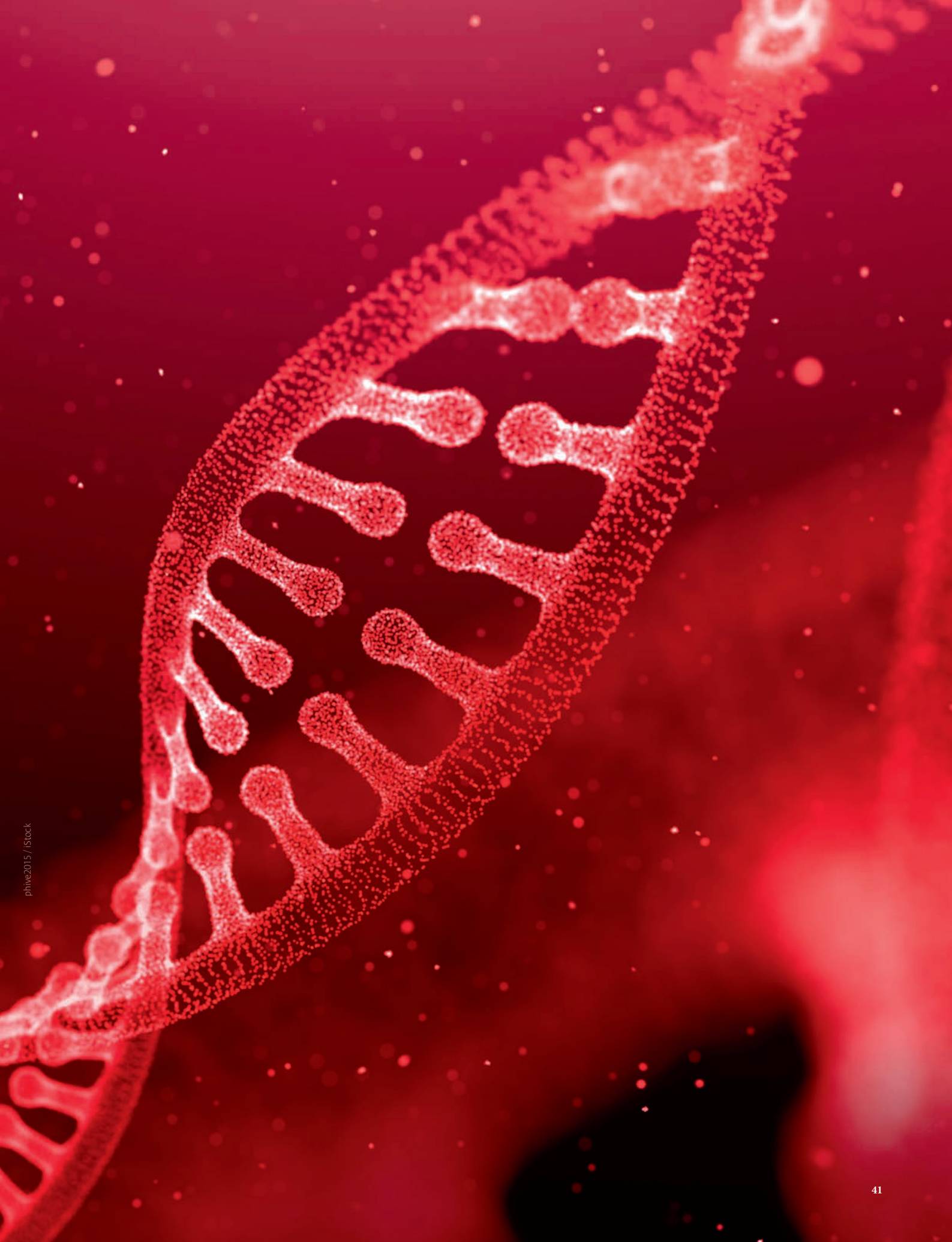
# DeepMind'dan Genomun "Karanlık Maddesi"ni Aydınlatmayı Amaçlayan Yapay Zekâ

# AlphaGenome

İlay Çelik Sezer | *TÜBİTAK Bilim ve Teknik Dergisi*

**Bir canlının sahip olduğu tüm genetik bilgiyi taşıyan DNA dizisi olan genom, canlıların vücudundaki tüm özellikler ve süreçler için temel belirleyicidir. Bu yüzden vücudumuzun işleyişinin ve hastalıkların nedeninin anlaşılabilmesi dolayısıyla da hastalıklar için tedaviler geliştirilebilmesi eninde sonunda genetik bilgiyi taşıyan genomun işleyişinin anlaşılmasına dayanır. Genetik bilgi denince her ne kadar aklımıza ilk olarak genler gelse de artık biliyoruz ki genler, genomun nasıl işlediğine dair ''hikâye''nin sadece bir bölümünü anlatıyor. İnsan genomunun, genleri kodlamayan, bir zamanlar işlevsiz sanılan ve hatta bu yüzden "çöp DNA" olarak adlandırılan %98'lik kısmının genlerin düzenlenmesinde hayati işlevler gördüğü uzun süredir biliniyor. Kodlamayan DNA'nın birçok önemli düzenleyici işlevi ortaya çıkarıldı ve hâlâ da yeni işlevler keşfedilmeye devam ediliyor. Ancak insan genomundaki genler kapsamlı şekilde haritalandığı yani konumları belirlendiği hâlde genomdaki kodlamayan DNA için bu tür bir haritalama henüz yapılamadı. Hatta henüz genler kadar anlaşılamadığı için kodlamayan DNA'ya genomun "karanlık maddesi" dendiği de oluyor. İşte DeepMind şirketinin geliştirdiği AlphaGenome adlı yeni yapay zekâ modelinin, genlerin yanı sıra kodlamayan DNA'daki gizemlerin de çözülmesine yardım etmesi umuluyor.**

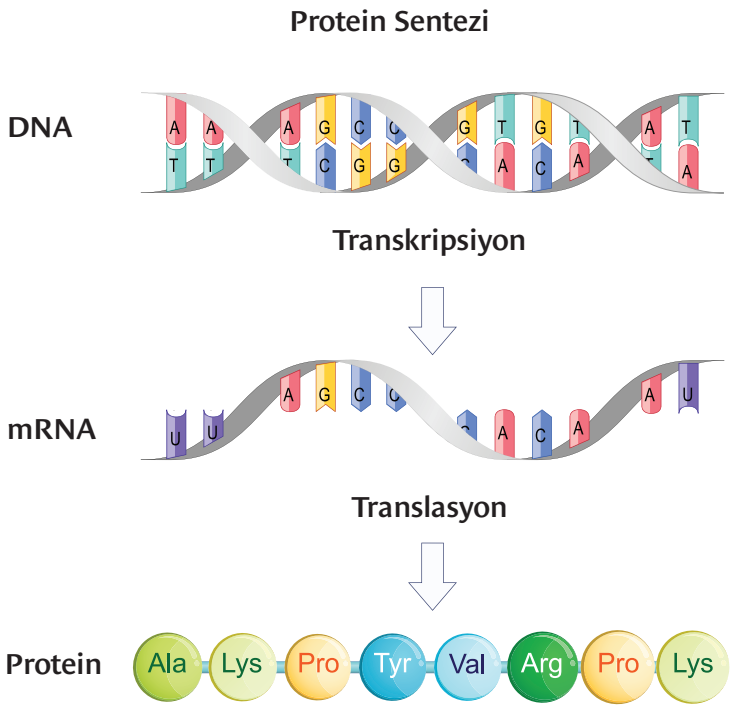


phive2015 / iStock

DNA'nın hücrelerdeki yaşamsal olayları kontrol etmesi temelde, taşıdığı genetik kodların hücrelerde üretilen proteinlerin yapılarını belirlemesi yoluyla gerçekleşir. Yaşamsal işlevlerin çoğu, proteinler tarafından gerçekleştirilir. Proteinlerin yapıları gen adı verilen genetik kodlarla belirlenir. Genler kromozomlar üzerindeki belirli DNA dizilerine karşılık gelen temel kalıtım birimleridir. Her bir gen, belirli bir proteinin yapısını belirleyen bir DNA kodu içerir. Bir genin taşıdığı genetik kod doğrultusunda protein sentezi gerçekleşirken önce transkripsiyon adı verilen aşamada ilgili genetik kodu taşıyan bir mRNA molekülü sentezlenir. Daha sonra da translasyon adı verilen aşamada mRNA'daki kod doğrultusunda protein sentezi yapılır. Bir genin kodladığı proteinin senteziyle sonuçlanan bu sürece o genin ifadesi ya da anlatımı denir. Bu adlandırmanın mantığı, her proteinin her zaman her yerde üretilmeyeceği gerçeğine dayanır. Yani her gen, her zaman, her yerde, her şartta ifade edilmez. Tüm hücrelerimiz aynı genoma yani aynı genetik bilgiye sahip olduğu hâlde vücudumuzda farklı hücre tipleri, farklı dokular ve farklı organlar bulunabilmesi dolayısıyla da vücudumuzun farklı işlevleri yerine getirebilmesi tam da bunun sayesinde olur.

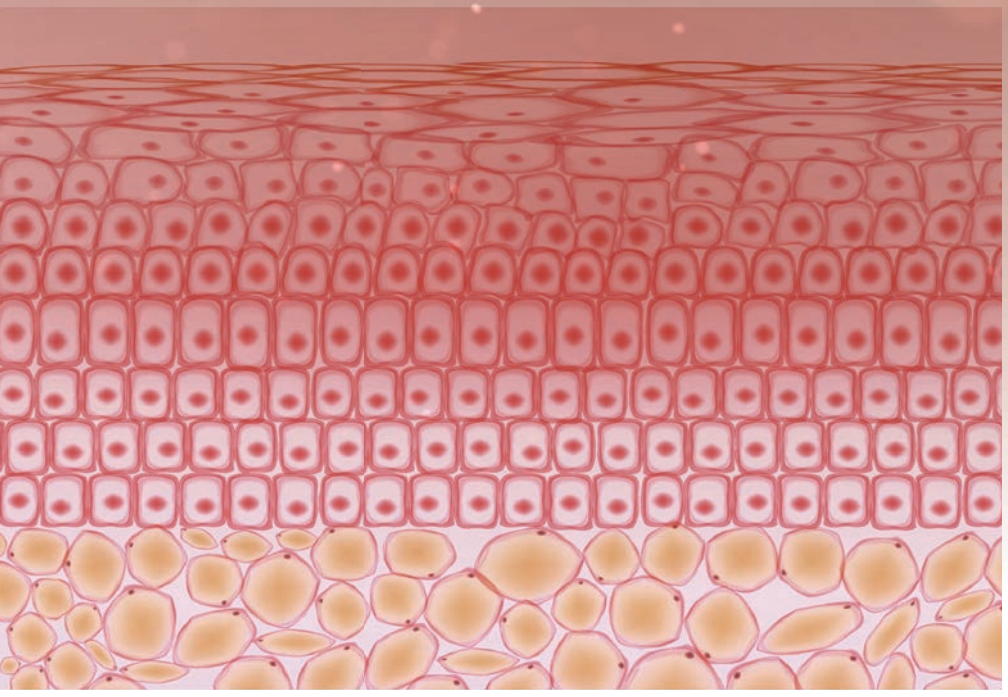
Hangi genin ne zaman, nerede ya da hangi şartlarda ifade edileceği ise topluca gen düzenlemesi adı verilen, çok çeşitli mekanizmalara dayanan bir süreçle kontrol edilir. Genlerin doğru şekilde düzenlenmesi, vücudun sağlıklı işleyişi için çok önemlidir ve pek çok hastalık genlerin düzenlenmesindeki aksaklıklardan kaynaklanır.

Moleküler biyoloji ve genetik alanındaki yöntemlerin ve tekniklerin gelişmesiyle genlerin düzenlenmesinde rol alan çeşitli mekanizmalar keşfedilmeye başlandı. Bu mekanizmalar aydınlatıldıkça da genleri kodlamayan ancak genomun %98 gibi büyük bir kısmını oluşturan "kodlamayan DNA"nın, genlerin düzenlenmesinde ne kadar önemli roller oynadığı daha iyi anlaşılabildi. Hatta bu mekanizmalardan biri olan, kodlamayan DNA dizilerinden sentezlenen küçük RNA moleküllerinin rol oynadığı RNA girişimi mekanizmasını keşfeden bilim insanları, bu keşifle 2006'da Nobel Fizyoloji veya Tıp Ödülü'nü kazandı. Yine kodlamayan DNA dizilerinden sentezlenen mikroRNA'lara dayalı düzenleme mekanizmasının keşfi ise 2024 Nobel Fizyoloji veya Tıp Ödülü'ne konu oldu. Kodlamayan DNA'nın rol oynadığı yeni mekanizmalar da keşfedilmeye devam ediliyor.




ttsz / iStock

Bir genin taşıdığı genetik kod doğrultusunda protein sentezi gerçekleşmesine o genin ifade edilmesi denir.



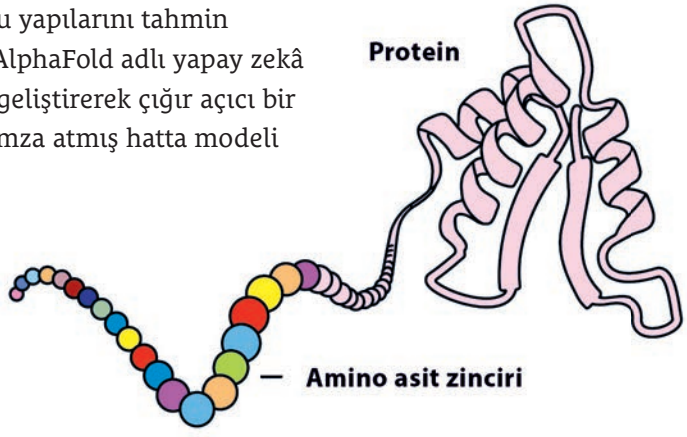
Benjamin Toth / iStock

Deri dokusu ile hemen altında yer alan yağ dokusunu gösteren bu çizimde deri dokusunun kendi içindeki farklı hücre tipleri ile yağ dokusundaki yağ hücreleri görülebiliyor. Bu hücre tipleri tıpatıp aynı genoma sahip olduğu hâlde gen düzenleme mekanizmaları sayesinde farklı genleri ifade ederek farklı yapılar kazanıp farklı işlevleri yerine getirebiliyor.

# DeepMind'dan Genom Kâşifi Yapay Zekâ Modeli AlphaGenom

Kodlamayan DNA'nın genlerin işleyişinde etkili olması sadece genlerdeki değil kodlamayan DNA'daki mutasyonların da vücutta bozukluklara ve hastalıklara neden olabileceği anlamına geliyor. Bu yüzden genler kadar kodlamayan DNA dizilerinin anlaşılması da önem taşıyor. Google'a ait Deepmind şirketi tarafından geliştirilen ve ilk versiyonu önceki ay akademik amaçlı kullanıma açılan AlphaGenome adlı yapay zekâ modeli, küçük genetik değişikliklerin genlerin ifadesinde nasıl değişikliklere neden olduğunu ortaya çıkararak kodlamayan DNA dizilerinin işlevlerinin de anlaşılmasına yardımcı olacak. Deepmind şirketi daha önce de proteinlerin amino asit dizilimlerine dayalı olarak üç boyutlu yapılarını tahmin edebilen AlphaFold adlı yapay zekâ modelini geliştirerek çığır açıcı bir başarıya imza atmış hatta modeli geliştiren bilim insanları 2024 yılında Nobel Kimya Ödülü'ne layık görülmüştü. DeepMind genom araştırmacılarından Natasha Latysheva, DNA'daki değişikliklerin bir organizmayı genel olarak nasıl etkilediğini anlamaya çalışmanın AlphaFold'un yaptığı işe göre daha karmaşık bir süreç olduğunu, tek bir genetik mutasyonun gen ifadesinde karmaşık ve zincirleme etkiler gösterebildiğini belirtiyor.

Yapay zekâ araştırmacıları şimdiye kadar bu karmaşık problemi parça parça ele almıştı. Öyle ki DNA'nın belirli bölümlerinde olası protein kodlayıcı bölgeler aramaya, hastalıklarla ilgili genetik örüntüleri tanımaya, genetik çeşitlilikleri anlamlandırmaya, hastalık riskini öngörmeye ve duruma özel klinik tedaviler geliştirmeye yarayan birbirinden bağımsız algoritmalar geliştirmişlerdi. AlphaGenome ise bu analizlerin pek çoğu ile bazı başka analizleri tek bir model ile tahmin edebiliyor.




nobelprize.org

Protein sentezinin ilk aşamasında amino asitler zincir şeklinde birbirine bağlanır. Ancak proteinler asıl işlevsel yapılarını, bu zincirlerin kıvrılıp her bir protein için özgün olan üç boyutlu şekillerini almasıyla kazanır. AlphaFold, proteinlerin üç boyutlu yapılarını amino asit dizilimlerine dayalı olarak tahmin edebilen çok güçlü bir yapay zekâ modelidir.

AlphaGenome, kamu kaynaklarıyla yapılan büyük bilimsel iş birlikleri kapsamında uzun yıllar içinde oluşturulan devasa moleküler biyoloji veri tabanlarına dayalı olarak geliştirildi. Bu veri tabanlarındaki verilerin bir kısmı, insan ve fare hücrelerinde belirli mutasyonların RNA üretimini ve gen düzenlenmesinde rol oynayan transkripsiyon faktörlerinin düzeylerini nasıl etkilediğine dair deney sonuçlarıyla ilgili. Bu veri kümeleri üzerinde eğitilen AlphaGenome, gen etkinliğinin kontrol edilmesinde rol oynayan gen dizilerini ve gen dışı DNA dizilerini, bunların yanında da önemli değişimlere yol açması en olası genetik değişiklikleri belirlemeyi öğrendi.

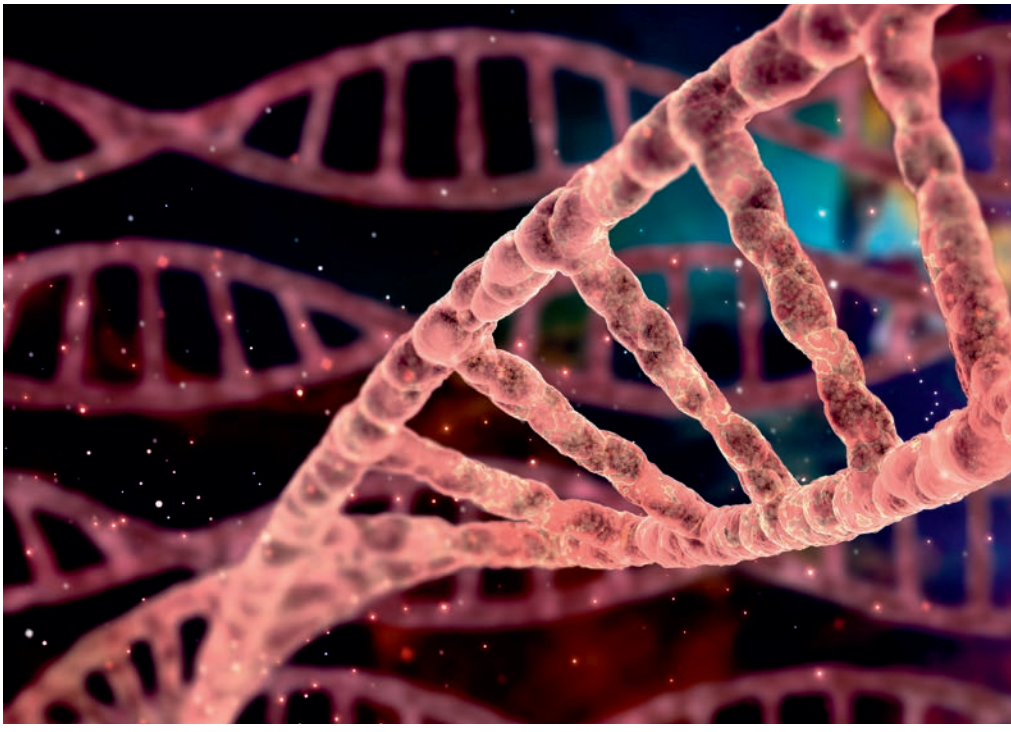
## AlphaGenome Nasıl Kullanılıyor?

Bir araştırmacı AlphaGenome'u kullanmak için önce en fazla 1 milyon baz çifti uzunluğundaki DNA dizilimi verisini programa yüklüyor. AlphaGenome da verilen DNA dizilimi üzerinde genlerin nerede başlayıp nerede bittiğini tahmin ediyor. Ayrıca RNA işleme sürecine yani genlerden hangi RNA formlarının üretileceğine ve ne kadar miktarda RNA üretileceğine ilişkin ayrıntıları da öngörüyor.

Daha sonra AlphaGenome DNA kodundaki tek harflik bir değişimin gen ifadesini nasıl etkilediğini, ayrıca genlerden üretilen RNA ve proteinleri nasıl değiştirdiğini tahmin edebiliyor. Her ne kadar bunu genomun yaklaşık %2'sini oluşturan protein kodlayıcı bölgeler için yapabilen yapay zekâ modelleri var idiyse de bunu genomun tamamı için yapabilen ilk model AlphaGenome.

DeepMind ekibinin yaptığı açıklamaya göre AlphaGenome, belirli bir DNA dizisinde gen kodlayan ve kodlamayan bölgeler ile transkripsiyon faktörlerinin bağlanma bölgeleri gibi spesifik özellikleri belirlemeye yönelik 24 testin 22'sinde başka bilgisayar modellerinden daha yüksek performans gösterdi. Ayrıca bir genetik değişikliğin gen düzenlenmesi üzerindeki etkisini öngörmeye yönelik 26 değerlendirme testinin 24'ünde başka bilgisayar modellerine üstün geldi.

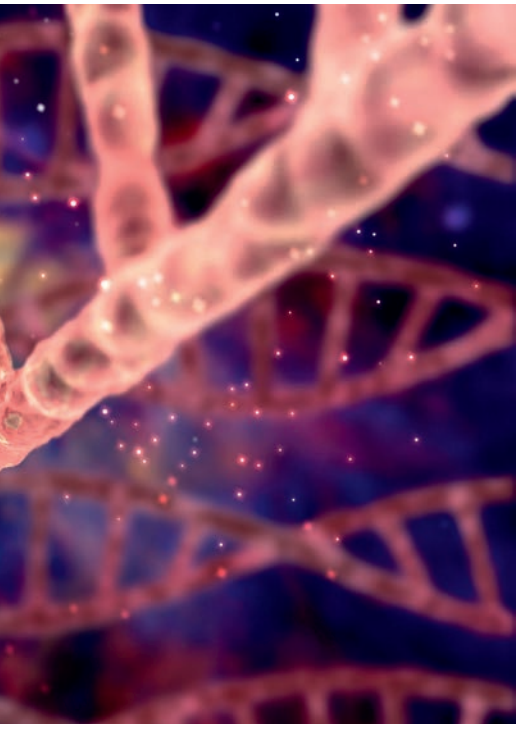
AlphaGenome'la yapılan denemelerden birinde model, kanserle ilişkili bir mutasyonun olası mekanizmasını araştırmak için kullanıldı. T-hücresi akut lenfoblastik lösemi (T-ALL) hastalarında yapılan bir çalışmada araştırmacılar genomdaki belirli konumlarda mutasyonlar gözlemledi. DeepMind ekibi AlphaGenome'u kullanarak bu mutasyonların yakınlarında bulunan *TAL1* adlı geni aktive edebileceğini öngördü. Hastalığın zaten bilinmekte olan mekanizmasıyla tutarlı olan bu bilginin elde edilmesi, AlphaGenome'un kodlamayan bölgedeki spesifik genetik değişikliklerle hastalık genleri arasında ilişki kurma becerisini doğruluyor.



## AlphaGenome Neler Vadediyor?

DeepMind ekibi AlphaGenome'un öngörü becerilerinin özellikle şu üç araştırma alanında faydalı olmasını bekliyor:

**1. Hastalıkların Anlaşılması**
AlphaGenome genetik aksaklıkların daha isabetli şekilde öngörülmesini sağlayarak araştırmacıların hastalıkların olası nedenlerini daha doğru şekilde belirlemesine yardımcı olabilir. Ayrıca belirli özelliklerle ilişkili genetik değişikliklerin, genetik işleyiş üzerindeki etkilerinin daha iyi anlamlandırılmasını sağlayarak yeni tedaviler için hedefler (örneğin hedef moleküller ya da mekanizmalar) belirlenmesine katkı sağlayabilir. Modelin, nadir görülmesine rağmen ciddi etkileri olabilen genetik değişiklikleri, örneğin tek bir gene bağlı nadir hastalıkları araştırmak için özellikle uygun olduğu düşünülüyor.



whitehoune / iStock

**2. Sentetik Biyoloji**
Modelin öngörüleri, spesifik düzenleyici işleve sahip (örneğin bir geni sadece sinir hücrelerinde aktive edip kas hücrelerinde aktive etmeyen) sentetik DNA dizilerinin tasarlanmasında kılavuzluk edebilir.

**3. Temel Araştırmalar**
AlphaGenome, genomdaki önemli işlevleri haritalayarak ve belirli bir hücre tipinin işlevlerini kontrol eden en önemli DNA kodlarını belirleyerek genoma ilişkin anlayışımızın hızlı bir şekilde gelişmesini sağlayabilir.

Her ne kadar AlphaGenome, DNA'nın gen kodlamayan kısımlarının işlevlerinin anlaşılmasında önemli bir ilerlemeyi temsil etse de hâlihazırda bazı kısıtlılıkları da bulunuyor. Genom diziliminde genin bulunduğu bölgenin çok uzağında (örneğin 100.000 nükleotid daha uzakta) yer alan ancak o genin aktivitesini kontrol eden DNA bölgelerinin etkilerinin tahmin edilmesi, DNA dizilimini temel alan başka modellerde olduğu gibi AlphaGenome için de önemli bir zorluk. Geleceğe yönelik olarak belirlenen önceliklerden biri modelin hücre ya da doku tiplerine özel örüntüleri yakalama becerisini geliştirmek. Benzer modeller gibi AlphaGenome da kişisel genom tahminleri için tasarlanmış ve doğrulanmış değil. AlphaGenome daha çok genetik değişikliklerin etkilerini ayrı ayrı tanımlamaya odaklanmış bir model. Ayrıca genetik değişikliklerin moleküler düzeydeki sonuçlarını öngörebilse de karmaşık özelliklerin ortaya çıkmasına ya da hastalıklara nasıl yol açtığına dair bütüncül bir anlayış sunamıyor. Çünkü çoğu özelliğin ya da hastalığın ortaya çıkışında genom dizilimindeki değişimlerin yanı sıra gelişimsel ve çevresel faktörler gibi çok daha kapsamlı biyolojik süreçler de etkili. Bunlar ise modelin doğrudan amacının kapsamı dışında kalıyor.

AlphaGenome'u geliştiren araştırmacılar, modeli tanıttıkları makalenin hakem denetiminden geçmemiş versiyonunu 25 Temmuz'da yayımladı. Makale hakem onayından geçince araştırmacılar, modelin kaynak kodunu ve model ağırlıklarını -ki bunlar bir modelin çıktıyı nasıl ürettiğini belirleyen unsurlardır- kamuoyuyla paylaşmayı planlıyor. Böylece araştırmacılar modeli kendi projelerine göre uyarlayabilecek. Büyük umutlarla tanıtılan ve karşılanan bu güçlü modelin bilim dünyasına neler kazandıracağını ise ilerleyen yıllarda bu modelin kullanılacağı araştırmaların sonuçları gösterecek. ■

**Kaynaklar**

https://deepmind.google/discover/blog/alphagenome-ai-for-better-understanding-the-genome/
https://www.science.org/content/article/deepmind-s-latest-ai-tool-makes-sense-changes-human-genome
https://www.nature.com/articles/d41586-025-01998-w