

Gen Haritalaması: Ne Demek, Haritalar Nasıl Oluşturuluyor, Neler İçeriyor, Nasıl Yorumlanıyor?

GENE MAPPING: HOW ARE GENES MAPPED, WHAT DO THESE MAPS CONTAIN, HOW ARE THEY INTERPRETED?

A. Nurten AKARSU*, Güven LÜLECI**

Hacettepe Üniversitesi, Çocuk Sağlığı ve Hastalıkları Araştırma Merkezi, Gen Haritalama Laboratuvarı*
Akdeniz Üniversitesi Tıbbi Biyoloji ve Genetik Anabilim Dalı**

ÖZET

Genlerin kromozomlar üzerinde bulunduğu yerlerin saptanması işlemine *gen haritalaması* adı verilir. Sitogenetik, genetik ve fiziksel haritalama olmak üzere başlıca üç tip gen haritası vardır. *Sitogenetik haritalar* genlerin lokalizasyonlarının sitogenetik yöntemler (başlıca *in situ* hibridizasyon) kullanarak bulunması işlemidir. Bu yöntemle genlerin hangi kromozom bandında yerleştiği kesin olarak söylenebilir. Somatik hücre hibridleri ve radyasyon hibritleri kullanılarak kromozomlardaki doğal ya da yapay olarak oluşturulmuş kırık noktaları üzerinden de genlerin haritalanması yapılabilir. Bunlar da sitogenetik haritalamanın özel metodlarını oluşturur. *Genetik haritalar*, genomun matematik metodlar kullanılarak analizi sonucu oluşturulmuştur ve özellikle genetik hastalıkların lokalizasyonlarının saptanmasında çok başarılı bir şekilde kullanılmıştır. Bu haritalar genlerin yerlerinin saptanmasında kullanılacak olan polimorfik nitelikteki genetik belirleyicilerin (marker) lokalizasyonlarını içerir ve bu marker bilgileri üzerinden hastalıkların haritalanması bir dizi karmaşık istatistik analiz kullanılarak tamamlanır. Genetik haritalama sonunda bulunan lokalizasyonlar olabilecek en olası lokalizasyonlardır ve fiziksel açıdan (baz çiftleri) ölçülebilir bir uzaklık göstermez. Genlerin baz çiftleri üzerinden fiziksel yapısını ve lokalizasyonunu gösteren haritalar ise *fiziksel haritalardır*. Fiziksel haritaların tamamlanması insan genom projesinin en önemli aşamalarından biridir ve tüm dünyadan çok sayıda bilim adamının yıllar süren çabaları sonucunda bugün insan DNA'sının nukleotid dizisi hemen hemen tamamlanmış gibidir. İyi bir gen haritalaması çalışması için her üç harita bilgilerinin birarada değerlendirilerek kullanılması gerekir. Bu yazıda haritalara genel bir bakış açısı getirilmiş ve farklı metodolojik yaklaşımları nedeniyle araştırmalarda bir kavram kargaşası yaratan genetik haritalama ve genetik haritaların kullanımı konusu detaylı olarak incelenmiştir.

Anahtar Sözcükler: Gen haritalama, sitogenetik harita, fiziksel harita

SUMMARY

Gene mapping is a general term to describe various methodologies used in locating chromosomal positions of genes. Cytogenetic, genetic and physical maps are the major categories. Chromosomal band location of a certain marker or a gene is clearly identified in *cytogenetic maps* which are constructed using cytogenetic methods such as *in situ* hybridisation. Somatic cell hybridisation technique and radiation hybrid maps are special sub-types of cytogenetic maps and detect gene localizations through natural and artificial chromosomal breakpoints respectively. Serious of mathematical and statistical analysis involve construction of *genetic maps* which are successfully used in identification of

Nurten AKARSU
Hacettepe Üniversitesi
Çocuk Sağlığı ve Hastalıkları
Araştırma Merkezi
Gen haritalama Laboratuvarı
e-mail: nakarsu@hacettepe.edu.tr

Güven LÜLECI
Akdeniz Üniversitesi
Tıbbi Biyoloji ve Genetik Anabilim
Dalı
Tel: 242-227-4482
Fax: 242-227- 4333
e-mail: luleci@med.akdeniz.edu.tr

disease causing genes. Genetic mapping is a probabilistic approach and identifies only the most likely localisation of the genes. It does not reflect physical distance and does not show base pair localisation of the genes. The development of *physical maps* is the final step of human genome project which targets to cover nucleotide base pair sizes of the entire human genome. Physical maps help to localize the exact base pair position of the genes. It is necessary to use informations obtaining from all three maps comperatively to be able to map a gene efficiently. In this paper, a general overview of different mapping informations will be given. Additionally, detailed information will be provided for genetic mapping strategies.

Key words: Gene mapping, cytogenetic map, physical map

Gen haritalaması, genlerin kromozomlar üzerinde bulunduğu yerlerin (lokus) gösterilmesidir. Böylece insan genomunun anatomisi ortaya çıkarılır. Pekçok genin ve diğer genetik marker'ların birbirlerine göre bir kromozom boyunca diziliş sırasının haritalanmasıyla bir kromozomun haritasını veya tüm genom haritasını çıkarmak mümkündür. Bu haritalama, insan vücudu fonksiyonlarının bilinmesi için gereklidir. Böylece insan genetik hastalıklarının heterojenite ve segregasyon analizleri bu bilgiler ışığında yapılabilecek ve gen tedavisi gerçekleştirilebilecektir.

Sitogenetik haritalama, genetik haritalama ve fiziksel haritalama olmak üzere temelde üç tip gen haritası vardır.

Sitogenetik haritalar, ilk olarak sitogenetik bantlama tekniklerinin oluşturulmasıyla farklı kromozomları tanımanın yanında, subkromozomal bölgelerin de ayırıldılmasıyla ortaya çıkmıştır. Bu tür haritalar düşük rezolüsyonlu (yaklaşık 6Mb) olmakla birlikte genlerin ve diğer DNA dizilerinin haritalanmasında çok basit bir yöntem olan kromozom in situ hibridizasyonu için zemin hazırlamıştır. Böylece uygun koşullarda yapılan hibridizasyondan sonra elde edilen sinyal proba, tanıyan DNA dizisinin bir harita lokasyonunun tanımlanması sağlanabilir. Günümüzde in situ hibridizasyon teknikleri floresan in situ hibridizasyon olarak geliştirilmiştir. Yöntem, belirli bir DNA bölgesi kullanarak tüm kromozomlar içinde eşdeğer bölgeyi bulmak şeklinde uygulanır. DNA bölgesi, radyoizotopla veya floresanla işaretlendikten sonra ısıtma ve soğutma işlemlerini takiben tek iplik haline getirilir. Boyanmamış ve lam üzerinde yayılmış metafazlara uygulanır. Hazırlanan prob, metafazda yer alan kromozomlardan kendi

ile eşdeğer bölgesi ile birleşir (1). Floresan kullanılarak geliştirilen yeni yöntemler, çok kısa sürede klonlanmış DNA dizilerinin haritalanmasında kullanılmaktadır. Son yıllarda birden fazla farklı renkli floresan işaretli problemlerle, kısa sürede daha fazla bölge incelenebilmektedir. Ayrıca fenotipin, gözlemlenebilir kromozomal yeniden düzenlenmeleri ya da kromozomal eksikliklerle ilişkilendirilmesiyle de sitogenetik haritalar oluşturulabilmiştir. Bunun yanında translokasyonlar ve kromozomal delesyonlardan türemiş parçaları içeren somatik hücre hibridleriyle doğal kromozom kırık noktaları ya da radyasyon hibridleri kullanarak yapay kırık noktaları haritalanabilir. Aynı zamanda düşük rezolüsyonlu fiziksel haritalar olarak da adlandırılan bu sitogenetik haritalarda çözünürlük, genellikle birkaç megabaz uzunluğundadır (2).

Somatik hücre hibrid çalışmaları ile de, kromozomlar ve onlar üzerindeki genlerin haritalaması çalışmaları yapılmıştır. En sıklıkla fare ve insan olmak üzere iki farklı tür somatik hücreler, birleştirici bir ajan veya özel inaktif viruslar aracılığı ile birleştirilir. Farklı farklı fare insan hibrid hücre klonları elde edilir. Başlangıçta bu klonlardaki hibrid hücrelerde hem insan hem fare kromozomları bulunurken, insan kromozomlarının bazıları her klonda farklı olmak üzere selektif olarak kaybolur. Özel bantlama yöntemleri ile morfolojik olarak, fare ve insan kromozomları kolayca ayırt edildiğinden hangi kromozom kaybı ile hangi özelliğin etkilendiği belirlenmeye çalışılmıştır. Örneğin bir enzim aktivitesi açısından fare somatik hücresinde eksiklik varsa, ama hibrid hücrede bu eksiklik gözlenmiyorsa hangi insan kromozomu üzerinde bu enzimin geninin bulunmakta olduğu belirlenmiştir. Somatik

Rekombinasyon birimi centimorgan (cM) ile ifade edilir. 1cM her 100 kişide 1 rekombinant birey olduğunun göstergesidir ($1/100=0.01$) ve Θ (theta) işareti ile gösterilir. Genetik uzaklıklar fiziksel anlamda ölçülebilir uzaklıklar değildir. Teorik olarak iki lokus arasında 1cM'lik bir uzaklıktan söz edildiği zaman, yaklaşık 1 milyon baz çiftinden söz ediliyor demektir. Ancak bu her zaman kesin değildir. Sık rekombinasyon yapan bölgelerde (rekombinasyon hot spot'ları) 2-3 cM'lik bölgeler fiziksel anlamda bazen beklendiğinden çok daha kısa olabilirler.

Bu metod yolu ile bir gen haritalama çalışmasını yapabilmek için kuşaklar arası kalıtımı izleyebileceğimiz ailelere ihtiyaç vardır. Yine hipotezimizi oluşturmak için kalıtım kalıbının belirlenmesi (tek gen, kompleks, mitokondrial vs.) ve marker kavramından ne anladığımızın iyice açıklanması gerekmektedir. Genetik haritalama işlemi, farklı moleküler biyolojik yöntemleri teknik olarak kullanmakla birlikte temel olarak istatistiksel bir analizdir ve tüm istatistik yöntemlerde olduğu gibi etkin bir genetik haritalama yapabilmek için hipoteze yönelik parametrelerin çok sağlam olarak belirlenmesi gerekir. Bu metod kullanılarak herhangi bir nitelik (gen, marker, hastalık gibi) haritalanmaktaysa da en geniş kullanım alanını, genetik hastalıkların haritalanması oluşturmaktadır. Bu noktada genetik etkenlere bağlı olduğunu düşündüğümüz bir hastalığın gen haritalamasını yapmak istediğimizi varsayalım ve bu işlem için gerekli aşamalar, nedenleri ve zorunlu parametreleri üzerinde tartışalım.

1. **Haritalanacak fenotipin özellikleri iyi tanımlanmalıdır:** *Fenotip:* Bir varlığın genler ve çevre etkileşimleri sonucu türlü metodlar ile ortaya konulabilen özelliklerinin tümü olarak tanımlanabilir (5). Gen haritalamasının ilk aşaması haritalanacak genin, hastalığın, ya da karakterin özelliklerinin standartlar oluşturularak saptanmasıdır. Gen haritalama çalışmalarında genellikle çok sayıda aileden toplanmış örnekler ihtiyacı vardır. Bu nedenle birden fazla merkez örnek toplama işlemine katılır. Merkezler arasında fenotip karmaşası yaşandığı du-

rumda gen haritalama çalışması yapmanın imkanı yoktur. Yine psikiyatrik hastalıklar gibi hastalık spektrumunun kesinlikle belirlenemediği durumlarda gen haritalama çalışmaları sıklıkla başarısızlıkla sonuçlanır.

2. **Haritalamada kullanılacak metoda karar verilmesi ve kalıtım kalıbı olabildiğince kesin olarak belirlenmelidir.** Kalıtım kalıbının saptanmasının gen haritalama için seçilecek metodun belirlenmesi ile çok yakın ilgisi vardır. Gen haritalama metodları başlıca iki ana gruba ayrılır.

a) *Parametrik metodlar:* Temel olarak **Linkage** (Bağlantı) analizleri olarak bilinir (Ülkemizde kullanılan terim **linkaj analizidir** ve metin içinde bu şekilde kullanılacaktır). Linkaj analizinin başarılı olabilmesi için değişmez parametrelere ihtiyaç vardır. Bunlar: 1- kalıtım kalıbı kesin olarak bilinmelidir; 2- üç kuşaklı geniş aileler tercih nedenidir 3- örnek toplama yaklaşımı kalıtım kalıbına göre olur. Örneğin otozomal dominant bir gen ile kalıtılan bir hastalık haritalanmak isteniyorsa örnek toplama hasta bireyler, varsa eşleri ve tüm çocukları yanısıra, normal bireylerin sadece kendileri şeklinde olmalıdır. Otozomal resesif bir hastalık söz konusu olduğu zaman ise, sıklıkla taşıyıcı olan anne-baba ve tüm çocuklarının toplanması yeterlidir. Pedigri (aile ağacı) analizleri yapılmaksızın sadece sporadik vakalar üzerinden linkaj analizini uygulayabilmenin imkanı yoktur (6). Şekil 2'de otozomal dominant bir genle kalıtılan bir hastalık üzerinden linkaj analizi görülmektedir.

b) *Non-parametrik metodlar:* Niteliklerin kalıtım kalıplarını belirlemek her zaman kolay değildir. Pek çok genetik nitelik çok faktörlü bir kalıtım kalıbı gösterir ve bunlar için hipotez oluşturmakta zorluklar yaşanır. Yine her zaman birkaç kuşağı bir arada bulmak ve örneklemek olanaksızdır. Özellikle geç yaşta başlayan hastalıklarda genellikle önceki kuşaklar ölmüştür, genç kuşaklarda ise henüz hastalık ortaya

çıkamıştır. Bu durumda, eksiklikler göz önüne alınarak parametrelerden bağımsız olan non-parametrik metodlar diğer bir deyişle assosiyasyon çalışmaları önerilmektedir. Assosiyasyon çalışmaları için farklı istatistik analizler önerilmişse de bunların hemen hepsinde kuşaklar arası segregasyonun test edilmesinden çok, vaka ve kontroller arasında anlamlı fark olup olmadığı test edilmektedir. Bu metodların linkaj analizine göre avantajları: 1- Kalıtım kalıbının bilinmesine gerek yoktur 2- Ailelerden çok vakaların ve kontrollerin toplanmasına gerek vardır. Dezavantajları ise 1- Örnek sayısı çok fazla olmalıdır 2- Özellikle kontrol grubunun oluşturulmasında çok dikkatli olunmalıdır 3- Sadece bu metoda dayalı gen bulunması maliyeti çok yüksek olan çalışmalardır ve genellikle başarısızlıkla sonlanmıştır. Şekil 3'te assosiyasyon çalışmalarından anlamlı bir sonuç bulabilmek için ideal olarak seçilmesi gerekli örnek sayıları verilmiştir.

Görüldüğü gibi her iki metodun da birbirine göre avantajları ve dezavantajları vardır. Eğer lokalizasyonu tek başına ispatlayacak büyüklükte aile paneli oluşturulabilmişse ve nitelik tek gen kalıtım modellerinden birini gösteriyorsa seçilecek metod tereddütsüz linkaj olmalıdır. Kompleks niteliklerde assosiyasyon metodlarından birisi seçilebilir. Bu amaçla günümüzde yaygın olarak önerilen iki aşamalı bir çalışma planı izlenmektedir ("two stage strategy") (7). Burada genellikle büyük aileler üzerinden tüm genom taranır. Potansiyel lokus'lar saptandıktan sonra sadece ilgili kromozom bölgelerine yönelik olarak assosiyasyon çalışmaları düzenlenir.

3. **Haritalamada kullanılacak marker'lar (genetik belirleyiciler) doğru olarak seçilmeli, mevcut gen haritaları etkin olarak kullanılmalıdır.** Bu kısımda gen haritalamasında kullanılan "polimorfik marker" terimini tam olarak anlamak gereklidir. Genetik yapıdaki değişiklikler en genel anlamı ile mutasyon tanımı ile belirlenir. Mutasyonlar toplumda görülme frekansları nadir olan (allel frekansı <0.0001) ve genetik hastalıklardan sorumlu olan

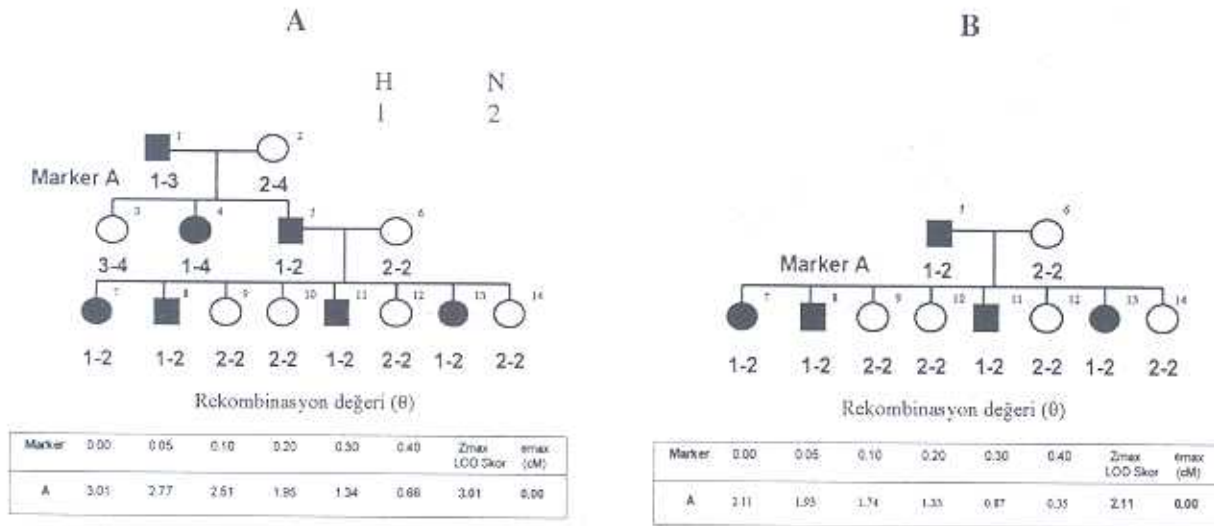
değişikliklerdir. Polimorfizm kavramı da yine mutasyon gibi genetik materyaldeki bir değişikliği gösterir ancak çoğu zaman mutasyonda olduğu gibi fenotipik bir değişiklikle sonlanmaz. Polimorfik değişikliklerde, nadir allel frekansı 0.01 dolayındadır yani toplumda yaygın olarak gördüğümüz değişikliklerdir (8). Örneğin genomda hastalıklardan sorumlu olmayan ancak kişiden kişiye farklılık gösteren kısa DNA tekrarlarından oluşan bölgeler vardır. Bunlar STRP (Short Tandem Repeat Polymorphism) olarak adlandırılırlar. Burada ikili, üçlü ya da dördü nükleotid tekrarları vardır. Bunlar içerdikleri kopya sayılarına göre DNA örneklerinin farklı büyüklükte olmaları ile sonuçlanır ve elektroforezde yürütüldükleri zaman farklı büyüklükte bantların görülmesine neden olurlar (Şekil 1). Eğer bir kişiye ait DNA örneği, genomda bu özelliği gösteren bölgelere özgü primerler kullanılarak "Polimeraz zincir reaksiyonu" (PCR, polymeraz chain reaction) yolu ile çoğaltılır ve elektroforezde yürütülecek olursa o kişiye ait polimorfik marker'lerden oluşan bir DNA profili ortaya çıkar. Bu tıpkı parmak izi gibidir, bireyler arasında farklılık gösterir. Birkaç kuşaklı aile bireylerinde böyle bir amplifikasyon yapılsa kromozomların kuşaklar arasında dağılımı, bu polimorfik marker'lar aracılığı ile belirlenir. Böylelikle baştan beri tartıştığımız kromozomlardaki mayozda oluşan parça değişimleri ve sonundaki olası rekombinant ürünler aile bireylerini birkaç kuşak izleme yolu ile saptanabilir.

İnsan genom projesinde genlerin haritalanması için ilk aşamada planlanan, tüm genomda böyle polimorfik özellikler gösteren markerların bulunması ve sadece bu bölgeleri çoğaltmaya yönelik primer dizinlerinin kromozom lokalizasyonları ve birbirlerine göre sıraları saptanarak bu bilgilerin kullanıma açılması idi. Bu amaçla üç kuşaklı geniş aileler bulunarak örneklendi (CEPH; Centre d'Etudes du Polymorphisme Humaine) (9). Bunlar genomda bulunan polimorfik markerlar için tiplendirildi. Birbirlerine göre olası lokalizasyonları linkaj analizi yapılarak hesaplandı ve haritalar oluşturulmaya başlandı. Eş zamanlı olarak sitogenetik

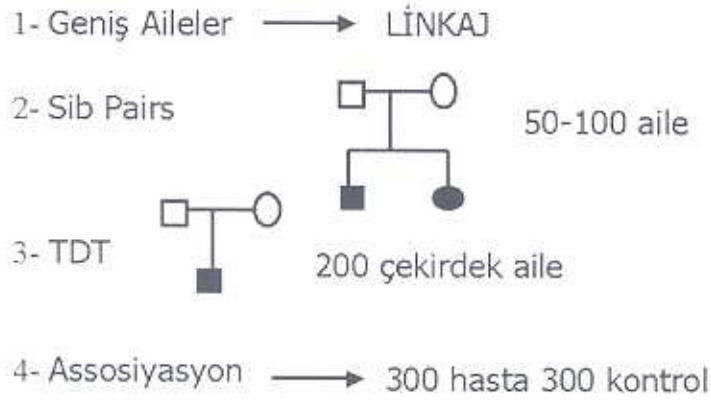
Gen haritalaması

haritalama da bir kısım markerlar için yapılarak sadece olasılık olarak lokalizasyonları değil kromozom bantlarına göre de lokalizasyonları bulundu. Sitogenetik olarak lokalizasyonları bilinen markerlar referans noktaları olarak kullanılarak genomu 1-2 cM aralıklarla tarayacak haritalar oluşturuldu. Genethon, Evry, Fransa laboratuvarlarında son derece başarılı olarak saptanan bu marker lokalizasyonları, araştırmacıların kullanımına açıldı (9-11). Bunu takiben diğer gen haritalama üniteleri de kendi bulgularını veri bankalarına koydular.

Tablo P'de böyle bir haritadaki birimlerin ne anlama geldiğini ve nasıl kullanılacağı gösterilmiştir. Bu haritalara ait bilgilere, The Center for Medical genetics, Marshfield (<http://research.marshfieldclinic.org/genetics>); Cooperative Human Linkage Center (CHLC) (<http://lpg.nci.nih.gov/CHLC>); UTAH Genome Center (<http://www.genome.utah.edu>); The Genome Database (<http://gdbwww.gdb.org>) internet adreslerinden ve Genethon haritası için 6 numaralı kaynakçadan ulaşmak mümkündür.



Şekil 2. Otozomal dominant bir genle kalıtılan bir hastalıkta linkaj analizi (Şekilleri değerlendirirken metin içindeki istatistik analizler kısmını okuyunuz) **A)** Üç kuşaktan toplam 14 birey kullanarak A markeri ile hastalık arasında bağlantı olup olmadığını test ettiğimizi varsayalım. Örnekte 5 numaralı birey babasından 1 numaralı marker allelini ve hastalığı almış görünmektedir. 2 numaralı marker alleli anneden gelmiştir ve anne normal olduğu için bu allelin hastalığın kalıtılması ile ilgisi yoktur. Otozomal kalıtım kalıbı varsayımı altında eğer tiplendirdiğimiz marker allel hastalık genine çok yakın ise yavru kuşaklarda 1 numaralı allelin bulunduğu kromozomu alan bireyler hasta olacak 2 numaralı allelin bulunduğu kromozomu alanlar ise normal olacaklardır (H=Hasta; N=Normal lokusu gösterdiği durumda H1 ve N2 birlikteliği non-rekombinant, H2 ve N1 birlikteliği ise rekombinant bireyleri gösterecektir). Eğer marker allel ve hastalık geni arasında bağlantı yok ise hastalar ve normaller arasında böyle bir seçim olmayacak hastalık ve normallik koşuluna bağlı olmaksızın yavru kuşaklar rastgele 1 ya da 2 numaralı alleli alacaklardır. Örneğimizde hasta olan herkes beklenildiği şekilde 1 numaralı allelin bulunduğu kromozomu almış, normaller ise karşı kromozomu seçmişlerdir yani hiç rekombinasyon gözlenmemiştir. Şeklin altında bu aileye ait LOD Skoru (Z) analizi verilmiştir. Analizde görüldüğü gibi bu aile en büyük olasılık değerini (Z_{max}= 3.01) rekombinasyon değeri 0.00 olduğu zaman (yani hiç rekombinasyon gözlenmediği zaman) almıştır. Hastalık geninden giderek uzaklaşıldığının varsayıldığı durumlarda (θ =0.05; 0.10; 0.20 ...cM), olasılık değerleri de giderek azalmaktadır (Z=2.77; 1.95; 1.34...). Z_{max}= 3.01 sonucu istatistik olarak anlamlıdır ve aradığımız hastalık geninin tiplendirdiğimiz A markerinin bulunduğu lokusta olması ihtimali bu bölgede bulunmaması ihtimalinden 10³ kez daha fazladır. **B)** Aynı aile bireylerinin üç kuşak yerine 2 kuşak olarak örneklendiğini varsaydıysaydı LOD skor değerlerinin düştüğünü görmekteyiz. Bunun nedeni önceki kuşağın genotipleri bilinmediği için 5 numaralı bireyin hastalık genini hangi allel ile aldığı bilinmemesidir. Hastalık geni 1 numaralı allel ile gelebileceği gibi 2 numaralı allel ile de kalıtılmış olabilir. Bu durum sonraki kuşaklardaki rekombinasyonun değerlendirilmesini değiştirmektedir. Hastalık geninin 1 numaralı allel ile aynı kromozomda olduğunu varsaydıysaydı bütün yavru kuşaklar non-rekombinant, 2 numaralı allel ile gittiğini varsaydıysaydı ise hepsi rekombinant olacaktır. LOD skor analizi yapılırken her bir birey için iki durumun ortak olasılığı alınır ve bu toplam LOD skor değerinin düşmesine neden olur (Şekilde en yüksek LOD değeri 2.11 olmaktadır). LOD skor analizinde 3 ve üstündeki değerlerin anlamlı olduğu bilindiğine göre istatistik olarak bağlantıyı ispatlayabileceğimiz bir ailede sırf üç kuşak toplanmadığı için bağlantı gösterilemeyecebilir. Linkaj analizinde ısrarla üstünde durulan üç kuşaklı geniş ailelerin önemi buradan kaynaklanmaktadır.



Şekil 3. Matematik metodlar kullanılarak hastalık geni haritalamasında oluşturulacak örnek paneli için gerekli olan bireyler ve ideal örnek sayıları. 1- Linkaj analizi için olabildiğince geniş ve çok sayıda aileler. Örnek toplama kalıtım kalıbına göre hastalık geni geçişinin takip edilebileceği bireyler üzerinden olur. 2- Kardeş çiftleri (sib pair) metodu için ana-baba ve kardeş çiftlerinden oluşan aile paneli. 3- TDT ve /veya HRR analizleri için ana-baba ve çocuk üçlülerinden oluşan çekirdek aileler. Bu tip çalışma için kontrol örneği toplanmasına gerek yoktur. 4- Vaka-kontrol çalışmalar için özellikle yaş, memleket, etnik köken, cinsiyet vs. gibi faktörlerin çok iyi belirlendiği vakalar ve kontrol örnekleri. Gerekli sayılar şekil içinde belirtilmiştir.

Bu haritalar kullanılırken dikkat edilecek başlıca iki nokta vardır. 1- Haritalar markerların kromozomlar üzerindeki sırasını ve birbirlerine genetik olarak uzaklıklarını vermektedir. Genotipleme yapılırken bu sıraların yanlış tutulması, yerlerinin karıştırılması veya rekombinasyon biriminden (cM), aralarındaki uzaklıklara dikkat edilmemesi gen haritalama çalışmalarının başarısızlıkla sonlanmasına neden olur. 2- Farklı merkezlerin marker haritaları birbirlerine göre ayarlanmamıştır. Başka bir deyişle bir haritadan elde edilen bilgiler ile yapılan bir haritalama çalışması diğer haritalarla uygunluk göstermeyebilir. Bu nedenle genelde tüm harita bilgilerine hakim olarak gerekli çalışmayı yapmak en doğrusudur. Bu noktada tüm veri bankalarındaki verileri karşılaştırmalı olarak bir oranda bünyesinde içeren The Center for Medical Genetics, Marshfield, USA (<http://research.marshfieldclinic.org/genetics>) veri bankasını referans olarak kullanmak yazarlardan birinin (A.N.A; nakarsu@hacettepe.edu.tr) önerisi olacaktır.

Marker haritalarındaki bilgileri kullanarak yapılacak bir gen haritalaması için başlıca 2 yol seçilebilir:

a- *Aday lokalizasyon yaklaşımı*: Burada haritalanacak olan niteliğe ait eldeki bütün veriler göz

önüne alınarak aday genler belirlenir. Örneğin baş-boyun bölgesine ait bir malformasyon haritalanacaksa baş ve boyunda ifade bulan ve gelişimsel olarak bu bölgelerin oluşumuna katkıda bulunan genlerden başlamak daha olasıdır. Aday genler belirlendikten sonra bu genlerin hangi kromozom bölgelerinde bulunduğu saptanır ve bu bölgelere sıkı bağlantı gösteren marker'lar öncelikli olarak test edilir. Bu tip marker bilgilerine ulaşmak için OMIM (On-Line Mendelian Inheritance in Man) ve LOCUSLINK internet adresleri en kullanışlı adreslerdir. Bu adreslere National Resource for Molecular Biology Information (NCBI) web sayfasından ulaşılabilir (<http://www.ncbi.nlm.nih.gov>). Bu veri bankaları birbirleri ile ilişkili olduğu için düşünülen aday gen saptandığı anda, diğer veri bankalarına ulaşarak ilişkili marker bilgilerini almak olanaklıdır. Aday bölgelerin taranması bitip de herhangi bir lokalizasyon saptanamazsa tüm genomun taranmasına geçilir.

b- *Genom taraması* (random genome-wide search): Burada hastalığın hangi kromozomda ya da hangi aday gen ile ilişkili olacağına dair bir ön-

Gen haritalaması

görü yoktur. Genomu 10-20 cM aralıklarla tarayan, kromozom lokalizasyonları belli marker panelleri kullanılarak tüm genom rastgele taranmaya başlanır. İstatistiksel olarak anlamlı sonuç bulunacak kromozom parçasına rastlanmaya çalışılır. Bazen tüm genom tarandıktan sonra bile lokus saptanmayabilir. O zaman genomu daha sık aralıklarla tarayan (5 cM) marker panellerine geçmek gerekli olabilir. Genomu 20cM aralıkla taramak için gerekli olan marker sayısı yaklaşık 200; 8-10 cM aralıkla taramak için gerekli olan 400 ve 5cM ara-

lıkla taramak için gerekli olan sayı ise 750'den fazladır (10-11). Bu sayıda marker ile çalışma maliyetinin, genotiplenecek örnek sayısı arttıkça artacağı açıktır. O nedenle genom taramaları için genellikle eldeki tüm örnekler kullanılmaz. İstatistiksel olarak en çok bilgi verecek bireylerden oluşan bir başlangıç paneli (initial screening panel) seçilir, genom bu bireyler üzerinden taranır ve potansiyel lokusların ispata ve bölgenin daraltılması için eldeki tüm bireyler kullanılır (saturation mapping).

Tablo I. Genetik haritaların yorumu. Farklı merkezlerin oluşturduğu polimorfik marker haritaları kullanım açısından ortak noktalar içerirler. Bir haritada önemli olan noktalar şunlardır: 1- Marker ismi: Her merkezin kendine göre verdiği bir isimlendirme vardır. Haritalarda "Afm" ile başlayan markerlar Genethon; "UT" ile başlayanlar UTAH; "Mfd" ile başlayanlar Marshfield ve "GATA", "ATA" gibi içerdiği nukleotid tekrarları ile başlayanlar ise CHLC tarafından saptanan markerları gösterir. Bu isim karmaşasını önlemek için ortak isimlendirmeye gidilmiştir ve bu haritalarda 'lokus sembolü' olarak gösterilir. Bu isimlendirme marker'ın yer aldığı kromozoma göre yapılır. Örnek olarak D1S2697 adı 1 numaralı kromozomda yer alan 2697 numaralı marker anlamına gelir. 2- Markerların kromozom üzerindeki sırası ve birbirlerine göre uzaklıkları: Markerlar arasındaki uzaklıklar genetik uzaklık birimi cM ile ifade edilirler. İki marker arasındaki uzaklığın değerlendirilmesi haritalarda yukarıdan aşağıya doğrudur. Bu kurala göre tablodaki örnek haritada D1S436 ile D1S2826 arasındaki uzaklık 3.9 cM ile gösterilmiştir ve bu iki marker arasında her 100 kişide yaklaşık 4 bireyde rekombinasyon gözlemlendiği gösterir. Buna karşılık D1S2697 ile D1S436 arasında 0.0 cM uzaklık vardır ki bu aralarında hiç rekombinasyon gözlemlenmediği anlamına taşır. Yani iki marker birbirine çok yakındır. Böylelikle birbirlerine yakın olan marker'lar bir arada gruplanır ve diğerleri uzaklıklarına göre sıralanır. Özellikle aralarında 0.0 cM olan marker pozisyonları tam belli değildir. Örnek haritada D1S2697 ve D1S436 arasındaki sıra D1S2697/D1S436 şeklinde olabileceği gibi D1S436/D1S2697 şeklinde de olabilir. Yine bu markerlar arasında hiç rekombinasyon gözlemlenmemiş olması başka çalışmalarda rekombinasyon gözlemleneceği anlamına gelmez. Bu nedenle haritaları etkin olarak kullanabilmek için farklı harita bilgilerini ve o ana kadar ilgili bölge ile yapılmış olan yayınları karşılaştırmalı olarak kullanmak gerekir. Markerlar arasındaki uzaklıklar her iki cins için ortak olarak verildiği gibi cinslere özel olarak da haritalara eklenmiştir. Bazı bölgeler için kadın ve erkek arasında rekombinasyon gözlenmesi sıklığı değişir (Örnek tabloda D1S436'da görüldüğü gibi). 3- Marker'ların içerdiği allel sayısı ve allellerin dağılımını gösteren değerler: Tabloda D1S2697 için 273-281 baz çifti büyüklüğünde bir amplifikasyon ürünü tanımlanmıştır. Bu aralıklar belirlenirken metin içinde anlatıldığı gibi CEPH ailelerinde yapılan genotip çalışmaları kullanılmıştır ve bu genotiplerde gözlenen en büyük ve en küçük allel büyüklüklerine göre bu aralıklar oluşturulmuştur. Bu allel büyüklükleri toplumlara göre farklılıklar gösterebilir ve o ana kadar gözlemlenmemiş olan bazı alleller farklı toplumlar ile yapılan çalışmalarda ortaya çıkabilir. Aynı nedenle markerların içerdiği beklenen allel sayısı ve allellerin frekanslarındaki dağılımlar da değişebilir. 4- Heterozigotluk: Genetik haritalama çalışmalarının yapılabilmesi için kullanılan markerların olabildiğince fazla allel çeşitliliği göstermesi gereklidir. Bu bir bireyin anne ve babasından kalıtığı bir lokustaki iki kromozomun farklı yapı göstermesi ile sonuçlanacak ve genotipleme yapıldığı zaman bireyin aldığı kromozomun anne ya da baba kaynaklı olduğunu ayırt etmemize yarayacaktır. Bu olay heterozigotluk olarak adlandırılır ve yine haritalara olasılık değeri olarak yansır. Örnek tabloda D1S2697 için heterozigotluk gözlenmesi olasılığı %70 olarak verilmiştir ki bu her 100 kişide 70 bireyin heterozigot bulunacağını 30 bireyin ise hem anne hem de babasından gelen kromozomları arasında allel farklılığı gözlemlenmediği (homozigot yapı) anlatır. Takdir edilecektir ki bu oran %80-90 arasında olduğu zaman gen haritalama çalışmaları için büyük avantaj oluşturacaktır.

Marker ismi	Lokus sembolü	Her iki cins için ortalama uzaklık(cm)	Kadınlar için ortalama uzaklık (cm)	Erkekler için ortalama uzaklık(cm)	Allellerin dağılım gösterdiği aralık (baz çifti)	İçerdiği allel sayısı	Heterozigotluk (%)
Afm298yc5	D1S2697	0.0	0.0	0.0	273-281	4	70
Afm217zc3	D1S435	3.9	5.6	0.1	200-240	12	75
Afm296te9	D1S2826	2.4	1.7	3.5	123-141	7	64

Buraya kadar anlatılanlardan anlaşılacağı gibi genetik haritalama için kullanılan laboratuvar deneyleri teknik olarak karmaşık bir yapı içermez. Kromozom lokalizasyonu bilinen primerler kullanılarak PCR amplifikasyonu, poliakrilamid jel elektroforezi, boyama ve genotipleme işlemleri yüzlerce kez tekrarlanır. Bu nedenle floresan markerların kullanıldığı ve tek bir çoğaltma işleminde çok sayıda marker kullanılan, otomatik genotipleme yöntemleri geliştirilmiştir. Ancak bu yöntemler zamandan tasarruf sağlarken maliyeti de çok yükseltmektedirler.

Genetik haritalama projelerinde kabul edilmesi gerekli bir diğer önemli nokta da diğer araştırma disiplinlerinden farklı olarak bu tip projelerin hazırlanmasında kesin sınırlarla belirlenen kuralların olmayışıdır. Örnek sayısı, seçilecek metodoloji, marker sayısı ve maliyet, niteliklerin kalıtım kalıplarına, fenotipe, saptanabilen aile sayısına ve en önemlisi şans faktörüne bağlı olarak değişebilir. Tamamen şans eseri olarak taramaya başlanılan ilk lokalizasyonda bağlantının saptanması ile genomu iki kez taradıktan sonra lokalizasyonu bulabilmenin maliyetinin değişeceği açıktır. Kısaca denilebilir ki yapılacak teknik işlem her çalışmada aynı olmasına karşın yorum ve planlama her bir çalışmada farklı özellikler ve tuzaklar içerir.

4. **İstatistik değerlendirmeler metoda göre seçilmedir: Linkaj analizi için;** LOD Skor (Logarithm of Odds Ratio) analizi uygulanır. LOD Skor: Linkaj saptanması olasılığının linkaj gözlenmemesi olasılığına oranının logaritmik değerinde ifade biçimidir. Başka bir deyişle aranılan genin, test edilen kromozom lokusunda olması olasılığının ilgili lokusta bulunmaması olasılığına oranıdır. Sonuçta çıkan değer arttıkça lokalizasyonun saptanması olasılığı da artacaktır. Örneğin bu değer 5 gibi bir sayı çıkarsa aranılan genin test edilen kromozom lokusundan seçilen marker'a bağlantı göstermesi olasılığı bağlantı olmaması olasılığından 10^5 kez (100.000 kez) daha fazla olacaktır. Negatif değerler de lokustan uzaklaşıldığının göstergesidir. LOD Skor = 3 ve üstü olan değerler bağlantıyı desteklemesi açısından anlamlı kabul edilirken 2 ve giderek negatifleşen değerler ise kesin olarak

bağlantı yokluğunu destekler. Aradaki değerler yoruma açıktır. Lokusun ispatlanabilmesi için bir dizi farklı işlem yapılması gerekebilir (aile ve örnek sayılarının artırılması, genomun diğer bölgelerinin araştırılması vs. gibi). Burada önemle belirtilmesi gereken nokta bu analizde sonuçta bulunan bir olasılık değeridir ve saptanan lokus hakiki lokus olmayabilir. Nitelikten sorumlu ilgili gen ve gen içi mutasyon gösterilinceye kadar lokus bilgisi yanlıtır olabilir.

LOD Skor analizleri için yaygın olarak kullanılan program J. Ott tarafından geliştirilmiş olan LINKAGE paket programıdır (12). Bu program amaca göre kullanılacak 4 alt programdan oluşur (MLINK, ILINK, LODSCORE ve LINKMAP programları). İki lokusun birbirine göre test edilmesi (örneğin, hastalık lokusunun marker allele'e bağlantısı) için en çok MLINK kullanılırken, polimorfik nitelikteki markerların birbirlerine göre lokalizasyonlarının ve sıralarının saptanmasında ILINK programı kullanılmaktadır. LODSCORE tüm olasılıkların hesaplanması yerine sadece maksimum olasılıkların hesaplanmasında, LINKMAP ise birden fazla marker içinde haritalanmak istenen genin göreceli pozisyonunun bulunmasında ("multipoint linkaj analizi") kullanılan alt programlardır. Programın kullanımı sırasında özellikle parametrelerin belirlendiği dosyaların oluşturulması özellik arz eder. Bu amaçla farklı gen haritalama laboratuvarları LINKAGE programı ile ilişkili ek programlar kullanılmaktadır. Bu konuda kullanılan tüm programlara <http://linkage.rockefeller.edu/soft/list.html> adresinden ulaşmak mümkündür.

Assosiyasyon metodları için; Bu alanda en sık kullanılan metodlar şunlardır. 1- Vaka ve kontrol çalışmaları, 2- Identity By Descent (IBD) (13) 3- Sib pair (14), 4- Affected Pedigree Member (APM) (15), 5- Transmission Disequilibrium Test (TDT) (16) ve 6- Haplotype Relative Risk (HRR) (17).

Fiziksel ya da moleküler haritalar, genomik DNA'nın klonlanmış parçalarının düzenlenmesiyle oluşturulurlar ve baz çifti sayılarına göre ayarlanmışlardır. Genetik haritadan farkı burada direkt olarak

DNA'yı oluşturan bazların sırası belirlenmiştir. Böylelikle genlerin fiziksel yapıları kesin olarak ortaya konabilmektedir. Haritalama projelerinin en son aşamadaki amacı olan fiziksel harita en yüksek rezolüsyona sahiptir. Diğer yandan, ökaryotik genomun çok küçük bir yüzdesi ifade olduğu için, bazı fiziksel haritalama yöntemleri, sadece transkripsiyonun gerçekleştiği dizilerin tanımlanmasına yönelmiştir. Genetik haritada olduğu gibi insan genomunun fiziksel haritası da 24 farklı kromozom için oluşturulur. Farklı kromozomlar için bağlantılı polimorfik marker gruplarını gösteren genetik haritaların tersine, farklı tiplerde fiziksel haritalar oluşturmak mümkündür.

Fiziksel haritalamanın en son hedefi DNA baz dizisinin çıkarılmasıdır. Ancak geniş DNA bölgelelerinin dizi analizinin yapılması teknik açıdan güç olduğu için haritalamada çözünürlüğü 1 Mb'dan daha aza indiren iki ana yöntem geliştirilmiştir. Bu yöntemler kullanılarak "Restriksiyon Haritalama" ve doğal olarak uzatılmış ya da yapay olarak uzatılmış kromatin ya da DNA fiberlerinde yüksek çözünürlüklü FISH'dir.

Rekombinant gen teknolojisindeki son gelişmelerle birlikte genomik DNA parçalarının klonlanması ve karakterizasyonu ile detaylı bir gen haritası elde etmek mümkün hale gelmiştir. Sonuçta da sadece gen yapısı hakkında bir kaynak olmakla kalmayan aynı zamanda genomda dizi organizasyonu, gen ve genomların evrimi hakkında da son derece değerli bir kaynak olan tüm bir genomun DNA dizisinin elde edilmesi mümkündür. Günümüzde, pekçok model organizmanın örneğin; bakteri, *S.cerevisiae*, *C.elegans*, *D.melanogaster*, *F.rubripes*, *D.erio*, insan gibi organizmaların genetik ve fiziksel haritalarını çıkararak sonuçta, tüm genom dizilerinin saptanmasını amaçlayan çok yoğun bir şekilde devam eden bir uluslararası işbirliği sürmektedir ve bu model organizmaların bazılarının tüm genom dizi analizi tamamlanmıştır (Tablo II) (18). Genomların haritalanması, analizi ve dizi analizi ile ilgili disipline "genomiks", adı verilmiştir.

Tüm dünyada binlerce bilim adamının yaklaşık 15 yıl süren çabaları sonucunda insan DNA'sının hemen hemen tamamlanmış nükleotid dizisi yani haritası or-

taya çıkmıştır. Gen haritaları sayesinde bitki ve hayvan üreticileri genetik olarak iyileştirilmiş kalitede üretim yapabilirken, insan genom haritasının çıkarılması açısından bakıldığında, biyokimyasal temeli bilinmeyen genetik hastalıklara neden olan genlerin özgül bir kromozom ya da kromozom bantı üzerine haritalanması, soyağaçları üzerinde bir marker ile olan bağlantısına dayanarak bu genin izinin sürülmesi (gene tracking), hastalığın tedavi edilmesi, insan ırklarının evrimleri ve ırkların dünya üzerinde göç yollarının çıkarılması mümkün olacaktır.

Tablo II. Dizi analizi yapılmış canlılardan örnekler

Organizma	Genom büyüklüğü	Tamamlanma tarihi	Tahmin edilen gen sayısı
H.influenza	1.8 Mb	1995	1740
<i>S.cerevisiae</i>	12.1 Mb	1996	6034
<i>C.elegans</i>	97 Mb	1998	19099
<i>D.melanogaster</i>	180 Mb	2000	13061
<i>H.sapiens</i>	3000 Mb	-	35000-45000

KAYNAKLAR

1. Eberhard Passarge. Çevirenler, Lülecı G, Sakızlı M, Alper Ö. Renkli Genetik Atlası. Nobel Tıp Kitapevleri Ltd.Şti ve Yüce Reklam Yayın A.Ş. İstanbul 2000;174, 192.
2. Strachan T, Read AP. Human Molecular Genetics, Second Edition. Newyork. John Wiley and Sons Ltd. 1999;241-250.
3. Hoffee PA. Medical Molecular Genetics. First Edition. Connecticut. Fence Creek Publishing 1998;135-140.
4. Haines JL. Genetic mapping. In: Current Protocols in Human Genetics, first edition. Eds. Dracopoli NC, Haines JL, Kotf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR. John Wiley and Sons Inc 1994;1: 1.01-1.8.30.
5. Başaran N. Tıbbi Genetik, Yedinci baskı, 1999.
6. Akarsu AN. Genetik hastalıklara neden olan genlerin saptanmasında aile ağacı analizlerinin önemi. Hacettepe Tıp Dergisi 1999;1: 85-91.
7. Davies JL, Kawaguchi Y, Bennt ST, et al. A genome-wide search for human type 1 diabetes susceptibility genes. Nature 1994;371:130-136.

8. Thompson & Thompson: Genetics in Medicine. Eds. Thompson MW, McInnes RR, Willard HF. Fifth edition 1991, WB Saunders Company.
9. Dib C, Faure S, Fizames C, et al. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 1996;380:1-138.
10. Dubovsky J, Scheffield VC, Duyk GM, et al. Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. *Hum Mol Genet* 1995;4: 449-452.
11. Gyapay G, Morissette J, Vignal A, et al. The 1993-94 Genethon human genetic linkage map. *Nature Genet.* 1994;7: 246-339.
12. Lathrop GM, Lalouel JM. Easy calculation of lod scores and genetic risks on small computers. *Am J Hum Genet* 1994;460-465.
13. Amos CI, Dawson DV, Elston RC. The probabilistic determination of identity-by-descent sharing for pairs of relatives from pedigrees. *Am J Hum Genet* 1990;47:842-853.
14. Blackwelder WC, Elston RC. A comparison of sib-pair linkage tests for disease susceptibility loci. *Genet. Epidemiol* 1985;2:5-97.
15. Weeks DE, Lange K. The affected-pedigree-member method of linkage analysis. *Am J Hum Genet* 1988;42:315-326.
16. Ewens WJ, Spielman RS. The transmission/disequilibrium test: History, subdivision and admixture. *Am J Hum Genet* 1995;57:455-464.
17. Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 1987;51:227-233.
18. Pennisi E. *The Human Genome Science* 2001; 291:1178.