

BIO 775 PROTEOMİK ve GENOMİK



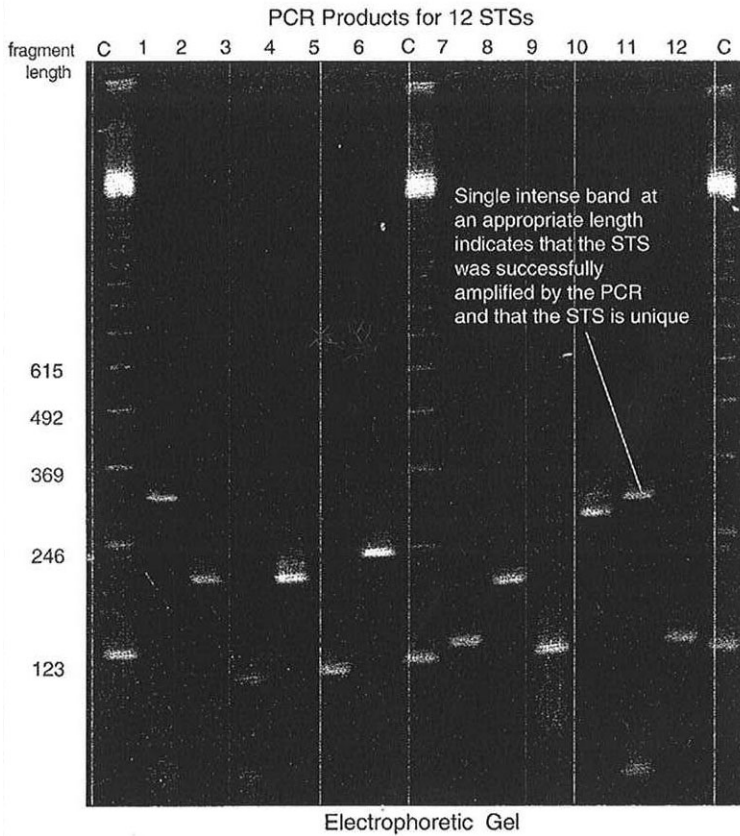
**SEQUENCE TAGGED SITES (STSs)
EXPRESSED SEQUENCE TAG (ESTs)**

S. Esin SELÇUK

SEQUENCE TAGGED SITES (STSs)

STS; genomda 200-300 baz uzunluğundaki kısa bir bölgedir ve bu sekans genomun başka hiçbir yerinde bulunmaz. Bu tek sekans PCR ile amplifiye edilebilir. STS sekansı tekrarlı elementler içerebilir fakat iki ucunun sekansının eşi yoktur ve eşsiz DNA primerleri bu uçlara göre sentezlenerek bu bölge PCR ile amplifiye edilir. Amplifiye olan ürünlerin jel elektroforezi spesifiklik gösterir. (şekil 1)

Figure 2. STSs from Chromosome 16



Şekil 1: STS'ler genomdaki eşsiz sekanslardır. Bu bölgeler PCR ile çağaltılır ve ürünler jel elektroforezinde birbirinden ayrılır. Eğer reaksiyon ürünleri aynı uzunlukta ise ve jelde tek belirgin bant gözüküyorsa bizim ilgilendiğimiz bölge çoğalmış demektir.

Los Alamos'ta 12 farklı STS markeri PCR'da çoğaltılmış ve ürünler 12 haneli jelde ayrılmıştır. Jelin her bir kuyucuğunda tek tip bir bandın varlığı STS'in gerçekten tek bölge olduğuna işaret eder. Amplifiye edilen ürünler büyüklükleri standart fragmentle kıyaslanır.

STS kavramı Olson ve ark. (1989) tarafından tanıtılmıştır. İnsan genom araştırmasında PCR'ın olası etkisinin saptanmasında, haritadaki lokasyonu bilinen tek kopya DNA sekansının kromozomlar arasındaki önemli genlerin fiziksel ve genetik haritasının yapılmasında bir marker gibi hizmet verdiğinin farkına varmışlardır.

Rough Sequence—347 Bases (lower case letters indicate uncertainty in the base call)

```

5 5'-GAATTCCTGA CCTCAGGTGA TCTGCCCGCC TCGGCCFCCC AAAGTGCTGG
51 GATTTACAGG CATGAGGCAC CACACCTGGC CAGTTGCTTA GCTCTCTAAG
101 TCTTATTTGC TTTACTTACA AAATGGAGAT ACAACCTTAT AGAACATTCG
151 ACATATACTA GGTTFCCATG AACAGCAGCC AGATCTCAAC TATATAGGGA
201 CCAGTGAGAA ACCAATGTCA GGTAGCTGAT GATGGGCAa GGgATGGGgA
251 CTGATATGCC cNNNNNGACG ATTCGAGTGA CAAGCTACTA TGTACCTCAG
301 CTTTtCATCT tGATCTTCAC CACCCATGGg TAGGTGTCAC TGAAaTT-3'
3'-CTAGAAGTG GTGGGTACCC AT-5' ——— Primer B
Primer A

```

| | | | | | | | | | Melting Temperature |
|----------|--------|-----|-----|-----|-----|-----|--------|--|------------------------|
| Primer A | 5'-GTT | TCC | ATG | AAC | AGC | AGI | CAG-3' | | 69.4°C |
| Primer B | 5'-TAC | CCA | TGG | GTG | GTG | AAG | ATC-3' | | 68.7°C |

Şekil 3: Yaklaşık olarak 171 bazın üstündeki sekanstan STS geliştirilir. Sekans 162. bazdan başlar ve 332. baza doğru koşar. Primer A, 21 baz uzunluğundadır ve sekans boyunca uzanır. Primer B de 21 baz uzunluğundadır ve sekansın 3' ipliğine doğru sekansı tamamlar. İki primerin erime sıcaklıkları yaklaşık olarak eşit olmalıdır. İki primer sekansını seçmek ve erime sıcaklıklarını hesaplamak için bir bilgisayar algoritması kullanılmaktadır.

Fiziksel Haritalama İçin STS Markerları

İnsan kromozomlarının kontig haritalarının yapılmasında overlap olmuş klon çiftlerini bulmak için STS'ler kullanılır. Her bir STS, genom üzerinde tektir. İki klon aynı STS bölgesine overlap olmalı ve overlap olmuş bölgeler STS içermelidir.

Overlaptan önce aynı STS'leri içeren klonlar DNA kütüphanesindeki klon koleksiyonundan tespit edilmelidir. Eğer bireysel olarak klonlanmışsa fragmentler nitroselüloz yada naylon membrana alınmalıdır. Daha sonra STS içerdiği belli olan klonlar STS markerlarından hibridizasyonla teşhis edilebilir. İlk önce STS kopyaları genomik DNA'dan PCR ile oluşturulur. Amplifiye edilen kopyalara radyoaktif P32 ile işaretlenerek denature edilir ve klon fragmenti membrana uygulanır. İşaretli markerler sadece klonları içeren DNA sekansına tamamen hibridize olacaktır. STS bakımından pozitif olan klonlar X-ray filminde koyu ışıkta membranla karşı karşıya bırakılır.

Daha hızlı bir inceleme metodunda klon kütüphaneleri havuzlara ayrılır ve PCR kullanılarak her bir havuzdaki STS'in varlığı sorgulanır. PCR tabanlı inceleme metodunda primerler her bir STS için sentezlenir ve çoğu havuzda paralellik görülür. Eğer klonlanmış fragmentlerin belli havuzu STS hedef sekansının PCR amplifikasyonu gerçekleşirse daha sonra en az bir klon hedef sekansı içermelidir. Çalışılan bir havuz kullanılarak proje altında tanımlanan amplifikasyonu desteklenmiş olan havuzlardaki klon yada klonların STS içerikleri incelenir.

16. Kromozomun Fiziksel Haritası İçin STS Markerları

Los Alamos'ta STS markerlarını içeren 16. kromozomun fiziksel haritasını yapmak için 16. kromozomda ortalama 100000 baz çifti denenmiştir. 16. kromozomun yaklaşık %60'ı kosmid klonlar tarafından örtülür. Ortalama olarak her bir kosmid kontig 100000 baz çifti aralığındadır. Yapılan çalışmada her bir kontigin sonuna uzanan klonlardan sekansı yapılan bölgeler tarafından STS'lerin geliştirilmesidir. Bu surette toplam 325 sekans bu klonlardan elde edilmiş ve yaklaşık 100'ünde STS gelişmiştir. STS markerları GenBank'ta stoklanır bu yüzden hiç kimse markerları tekrar oluşturmak istemez ve klonların STS içeriklerini tanımlamada kullanır.

Klonların sonlarındaki STS markerları çalışılmış olan kosmid kontiglerinde birkaç amaca hizmet eder. Önce YAC klonlarının kütüphanesini incelemek için kullanılır. Klonlar iki farklı kosmide overlap yapabilir. 16. kromozom için yapılan çalışmada kütüphanede 550 YAC bulunmaktadır. YAC'lar yalnızca 16. kromozoma insert olmuş DNA içermektedir. Ortalama uzunluk 215 kb'dır. Bunun gibi kromozoma spesifik YAC kütüphanelerinin yapılması fiziksel haritalama ve kromozom kütüphanelerinin tanımlanması için önemli bir buluştur.

Çalışmada YAC'lar havuzlara ayrılmış ve PCR tabanlı inceleme stratejisi her bir STS içeren YAC'ları tanımlamak için kullanılmıştır. Tasarlanmış havuz. Los Alamos'taki kuramsal biyoloji grubundaki David Torney tarafından bulunmuştur. STS içerdiği tespit edilmiş bir YAC klonu bir PCR tekniği kullanılarak (inter-ALU PCR olarak da bilinir) YAC'tan prob setleri oluşturulabilir. Problar kosmid klonlarla düzenlenen kütüphanede hibridize edilir. Eğer klonlar iki farklı kontigde pozitif hibridizasyon sinyali verirse daha sonra YAC iki kontig arasında köprü oluşturmaktadır. Kosmid klonlarından STS içeren 30 YAC identifiye edilmiştir.

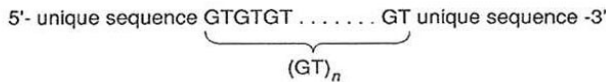
16. kromozomun çeşitli kısımlarını içeren fare/insan somatik hücre hibritleri tarafından belirlenen 16. kromozom üzerine lokalize olmuş her bir kosmid için aynı STS kullanılabilir. Colaborators David Callen ve Güney Avustralya'daki Grant Sutherland of Adelaide Çocuk Hastanesi 16. kromozom üzerinde ortalama 1.7 milyon baz uzunluğundaki 50 bölgenin somatik hücre hibritlerini biriktirerek bir panel oluşturmuşlardır. Hibridizasyon tabanlı metodu kullanarak, STS'ler ve bir PCR tabanlı strateji kullanılarak DNA her bir hibrit hücrede incelenmiştir. Bu yüzden lokalize olmuş her bir 70 kontig somatik hücre panelinde 50 bölgeden birinde belirtilir.

Genetik Linkaj Haritalama İçin STS Markerları

Ayrıca STS'ler genomda biri diğerinden uzunluk bakımından çeşitli eş olmayan bölgeler arasından geliştirilebilir. PCR ile bu çeşitli bölgeler amplifiye edilebilir. Farklı bölgeler genomun mevcut bölgesinin çeşitliliğine bağlıdır. Farklı bölgede belirlenen bir STS polimorfik bir DNA markerıdır. Bu DNA markerı diğer markerlarla birlikte ve genetik linkaj haritalarında lokalize olmuş şekilde izlenebilir.

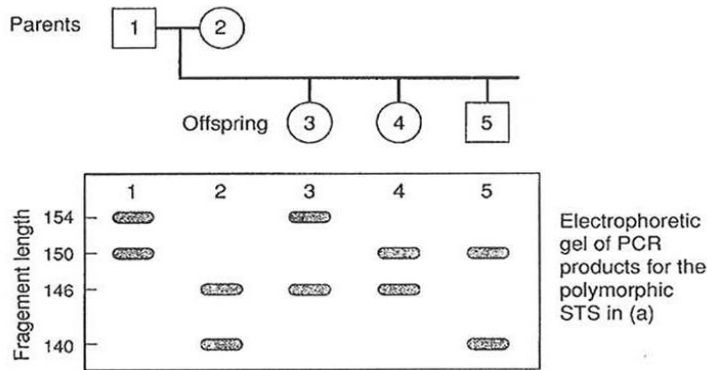
Şekil 4a çeşitli uzunluklardaki çeşitli bölgelerin bir örneğini göstermektedir, bu bölgelerde STS'ler gelişebilir. Bölgenin sonunda yaklaşık 20 nükleotitlik eşsiz bir sekans bulunmaktadır. Bu sekans PCR için primer gibi hizmet edebilir. Bu iki sekans arasında basit tandem tekrarlar bulunur (GT)_n veya (GT) n kez tekrar eder. Bunun gibi dinükleotid tekrarları insan genomu boyunca tri-, tetra-, penta- nükleotid tekrarları şeklinde dağılır. Ayrıca kromozomlarda lokuslar arasındaki bu n tandem tekrarlarının sayısı populasyonlar arasında geniş ölçüde farklılık gösterir.

(a) A Polymorphic STS



The number *n* of GT repeats varies among the population.

(b) Inheritance of the Polymorphic STS shown in (a)



Şekil 4: değişken bir lokus içeriği kısa tekrarlanan sekansa sahiptir, dinükleotid tekrarları gibi (GT)_n bir STS içinde olabilir. Bir örnek şekil 4a da gösterilmiştir. Amplifiye edilen ürünün büyüklüğü STS'in bulunduğu lokustaki n sayısına bağlıdır. Bu yüzden STS polimorfiktir. STS markerinin her bir bireyde iki kopyası bulunur ve her bir kopya n derecesine ve bu suretle farklı polimorfik STS allellere sahiptir.

Polimorfik STS'in kalıtımı 5 üyeli bir ailede şematik olarak açıklanmıştır (b). Her bir aile üyesinden STS'lerin PCR ürünleri elektroforetik jelde gösterilmiştir. İki allelden biri anneden diğeri babadan gelir. Çocuklar her bir bireyden bir STS allelini alırlar.

Çünkü markerler çok alleldeki tekrarlı sekanslar etrafından geliştirilirler. İki farklı marker için bir ebeveyn heterozigot olmalıdır. Bu markerların sahip olduğu çok sayıdaki allel linkaj analizi için büyük bilgi sağlar. Polimorfik STS'ler insan genomundaki her bir kromozomun 2-5 cM uzaklığındaki DNA hakkında bilgi vererek genetik linkaj haritasının yapılmasını sağlar. Ek olarak, STS'ler kolaylıkla fiziksel haritalara yerleştirilir ve böylece kromozomların fiziksel haritasıyla birlikte linkaj haritalarının sıraya dizilmesinde rahatlık sağlar.

STS'ler bu farklı bölgelerin geniş sınıfından geliştirilebilir. Çünkü bu markerların allellere karşılık gelen polimorfik STS'lerden çeşitli büyüklükteki PCR ürünleri elde edilip bu amplifiye edilen ürünlerin jel elektroforezi yapılarak bireyin marker allelleri tespit edilir.

Polimorfik STS'ler özellikle kullanışlıdır. Çünkü polimorfik STS'ler hem fiziksel haritaların her bir kromozomun genetik linkaj haritalarının yapılmasına hizmet eder. Bu suretle polimorfik STS'ler iki tip harita üzerinde farklı uzaklıklardaki sekanlar arasında bağlantıyı sağlar.

Los Alamos'ta (GT)n tekrarlarının lokasyonu, fingerprinting ve haritalama stratejisiyle identifiye edilmiştir. Bu çalışmada araştırmacılar şimdi linkaj haritalama içinde kullanmak için STS'ler içindeki bu bölgeleri çalışmaktadırlar.

EXPRESSED SEQUENCE TAGS (ESTs)

Bir EST, cDNA sekansının bir segmenti olup mRNA'ya benzer. EST'ler genomik sekansı açıklamada bir dayanaktır ve spesifik dokulardaki veya büyüme koşullarıyla ilişkili gen ekspresyonunu analiz etmeyi sağlar. İnsanda genlerin ekspresyonunun tanımlanması kodlanmayan DNA bölgelerinin (intronların)olmasından dolayı zordur. Bu nedenle mRNA izolasyonu ile ifadesi olan genleri tespit etmek mümkündür. Ancak mRNA hücre dışında stabil değildir. Bu nedenle mRNA'dan cDNA(komplementer DNA) sentezlenir. c DNA ifade edilen sekansı içerir. daha sonra molekülün iki ucundan EST oluşturmak üzere birkaç yüz nükleotid sekanslanır.

EST sekansları eksprese edilen genlerin kurulmasında hızlı bir metottur. Herhangi bir çeşit cDNA kütüphanesi sekanslama yapmak için kullanılabilir. Single-pass (tek geçişli) sekanslama cDNA klonunda sadece bir uçtan yapılır. Bu verilerin hızlı ve oldukça ucuz toplanmasını sağlar. Ancak single-pass sekanslama birçok hata içermektedir.

İlk olarak Bethesda'daki Ulusal Sağlık Enstitüsü'ndeki bir grup araştırmacı tarafından bilinmeyen cDNA klonlarının geniş ölçüde sekanslanmasıyla başlamıştır. Tüm genomik sekans hızlı bir şekilde ilerlerken bu çalışmaların ilk sonuçlarını 1990'lı yılların başlarında yayınlamışlardır. İnsan EST projelerini destekleyenler tüm genomu sekanslamının çok uzun zaman alacağını ve genlerin kodlanan sekanslarının genom içeriğinde büyük bilgi gösterdiğini fakat DNA'nın

yalnızca %3'ünü kapsadığı konusunda tartışmışlardır. cDNA'nın sekanslanması tamamlanmamış ve daha az hatasız olmasına rağmen genomun sekanslanmasından daha hızlı ve daha ucuz olduğunu ileri sürmüşlerdir. Ancak bazı araştırmacılar bu projeye karşı çıkmışlardır. Tüm genomun sekanslanmasına karşı bir değişiklik gibi görmüşlerdir. Araştırmacılar mRNA'nın ekspresyonunun bütün dokularda, hücre tipinde ve gelişimsel seviyede bulunmasının zor olduğunu konusunda tartışmışlardır. Hem EST'nin hem de genomik sekanslanmanın koordineli olarak yararlı olduğu konusunda uzlaşmışlardır.

Geniş ölçüde EST sekanslama, insan, pirinç, mısır, fare, HIV'yi kapsayan 300'ün üstünde türde ve *Arabidopsis elegans*, *Escherichia coli*, meyve sineği, *Caenorhabditis elegans* gibi çok sayıdaki model organizmada yapılmıştır. İnsan EST sekanslama çok dikkatle yürütülmüştür ve tüm genomun sekanslanmasıyla internet üzerinden ulaşılabilen 3,2 milyonun üzerinde EST klonu mevcuttur.

İnsan EST'lerine ek olarak diğer EST sekanslarına ulusal bir veri tabanı olan dbEST'ye internet aracılığıyla ulaşılabilir. Bu veri tabanlarına internet üzerinden ulaşabilmenin amacı internetteki güçlü programların bulunarak dünya çapında yapılan araştırmaların analiz edilebilmesine bir ölçüde yardımcı olabilmektir. Ayrıca sadece bir örneğe ait olan koleksiyonlar da bulunmaktadır.

Yaklaşık olarak sonsuz kaynaktan EST üretilir. Ticari olarak mevcut olan cDNA kütüphaneleri DNA sekanslaması için bir kaynak sağlar. Diğer EST'ler cDNA kütüphanelerinden sentezlenmişlerdir. Bu yüzden mRNA kaynakları spesifik hücre tipi, doku ve fizyolojik koşullarından gelebilir. İnsan EST koleksiyonundaki en erken kayıt beyin dokusundan izole edilen mRNA'ya ait cDNA kütüphanesidir.

EST'ler genomun fiziksel haritasının çıkarılmasında kullanılabilir. Fiziksel haritalar daha sonra insan hastalıklarıyla ilişkili genleri klonlamak için kullanılabilir. Ek olarak EST sekansları genomik sekansın anlaşılmasında pek bir açıklık getirmemektedir. Çünkü EST'ler cDNA'dan gelirken genomik sekans intergenik bölgeleri ve intronları içermektedir (Yani olgun RNA'nın sekansını yansıtmaktadır). Bilim adamları ve biyoinformatikçiler EST sekanslarını genler içinde uzanan intron-ekson bağlantılarının nerede olduğunu belirlemek ve aynı gendeki alternatif transkriptleri ortaya çıkarmak için kullanabilirler. Bu çalışmalar aynı mRNA'ların sergilediği overlapping (üst üste binmiş) sekanslara sahip EST kontiglerini meydana getirmeye yardım eder. Ancak EST koleksiyonu tarafından kaç tane farklı genin olduğuna karar vermek zordur. Çünkü

cDNA klonlarının sıklıkla 5' ucu kesilmiştir ve klonların sık olarak 5' ucunun üst üste binmeyen sekansların ürünlerine karşılık gelir.

Gen ürünlerinin fonksiyonu tanımlama EST koleksiyonları değerli bir çare olarak önerilmektedir. Mevcut veri tabanları sekansları karşılaştırmada ve sıraya dizmede kullanılabilir. Genin protein ürününün in vivo fonksiyonu aynı veya farklı türlerle ilişkili diğer genler test sekansında homoloji sağlayabilir/gösterebilir. Ek olarak, EST'ler klonların blotting analizi için bir kaynak sağlar. Çok yakın bir zamanda EST koleksiyonlarının kullanılması mikroarray'e göre ucuzdur. EST' den sonra mikro array tek deney içinde geniş ölçüde gen ekspresyonunu çalışılmasında kullanılır.

KAYNAKLAR:

- ❖ <http://www.usask.ca/biology/rank/316/genomics/genomics/html>
- ❖ <http://library.lanl.gov/cgi-bin/getfile?00326699.pdf>
- ❖ <http://www.bookrags.com/sciences/genetics/sequence-tagged-sites-wog.html>
- ❖ <http://www.bookrags.com/sciences/genetics/expressed-sequence-tags-ests-wog.html>