

Yeni Nesil DNA Dizileme New Generation DNA Sequencing

Duran Üstek¹, Neslihan Abacı¹, Sema Sırma¹, Aris Çakiris¹

¹Genetik AD, Deneysel Tıp Araştırma Enstitüsü, İstanbul Üniversitesi, İstanbul, Türkiye

Yeni Nesil DNA Dizileme

Özet

Yeni nesil dizileme; genom, transkriptom, DNA-protein etkileşimlerinin geniş kapsamlı analizini ucuz, rutin ve yaygın hale dönüştürdüğünden biyolojik araştırmaları önemli ölçüde hızlandırma potansiyeline sahiptir. Yeni nesil dizileme teknolojisi; mRNA, küçük RNA profili, transkripsiyon faktör bağlanma bölgelerinin genom boyu karakterizasyonu, kromatin yapısı ve metilasyon paterni, atasal DNA mikrobiyolojisi ve metagenomik için yeni ve hızlı yollar sağlar. Bu makalede yeni nesil dizileme teknolojileri, yeni nesil dizileme sistemlerindeki yeni gelişmeler ve yeni nesil dizileme teknolojisinin uygulama alanları tartışılmıştır.

New Generation DNA Sequencing

Abstract

Next-generation sequencing has a potential to dramatically accelerate biological research, by enabling the comprehensive analysis of genomes, transcriptomes and DNA-protein interactions to become inexpensive, routine and widespread. The next-generation sequencing technologies offer novel and rapid ways for genome-wide characterisation and profiling of mRNAs, small RNAs, transcription factor binding regions, structure of chromatin and DNA methylation patterns, ancient DNA, microbiology and metagenomics. This review discusses next generation sequencing technologies, recent advances in next generation sequencing systems and applications for next generation technologies.

İletişim bilgileri:

Duran Üstek

Genetik AD, Deneysel

Tıp Araştırma Enstitüsü,

İstanbul Üniversitesi,

İstanbul, Türkiye

Telefon: (0212) 4142000 / 33317

Faks: (0212) 5324171

E-mail: dustek@istanbul.edu.tr

Giriş

İnsan genom projesinin tamamlanması postgenomik alanın başlangıcı olmuştur ve sistem biyolojisi yaklaşımı büyük önem kazanmıştır. İnsan genomundaki hücrel hemostaz, gelişim ve hastalık ilerleyişi ile ilişkili düzenleyici ağlar (regulator network) ve bu ağlarda rol oynayan fonksiyonel elementler tanımlanmaya başlanmıştır. Hücrel aktiviteleri düzenleyen genomdaki bu fonksiyonel elementlerin belirlenmesi ve bunların popülasyonlar arasındaki varyasyonları kişiye özgü tedavinin geliştirilmesine de temel oluşturur.

Son yıllarda geliştirilen yeni nesil DNA dizileme teknolojisi, genetik/epigenetik düzenleyici ağların, kromatin

yapısı, nükleer yapılanma ve genom varyasyonları (çeşitliliği) hakkında bilgi üretimini sağlayan önemli bir araç olmuştur.

Yeni nesil DNA dizileme sistemleriyle yüksek doğrulukla, ultra hızlı olarak, dizileme yapılabilmektedir. Bu yöntemle elde edilen bir mikrobiyal genom dizisi araştırmacılara başka hiçbir deneysel yöntem ile elde edilemeyecek kadar zengin ve özgün bilgi sağlamaktadır. Örneğin 4.6 Mb'lık E. coli genomu tek bir okuma ile tamamlanabilmektedir. Yapılan bir çalışmada E. coli genomu dört kere, her bir koşmada 400.000 okuma yapılarak de novo dizilenmiş ve sekanslamaların %99.997 ila %99.999 arasında doğrulukla yapıldığı tespit edilmiştir (1, 7).

Yeni nesil DNA dizileme teknolojisi kullanılarak, de novo dizileme yöntemi ile bitki, bakteri, maya, mantar, virus gibi farklı organizmaların genomlarının ve bakteri yapay

kromozomlarının (Bacterial Artificial Chromosomes (BACs)) ultra hız olarak, yüksek doğrulukla dizilenmesi mümkündür. De novo dizileme, shotgun ve paired-end dizileme yöntemlerini kapsamaktadır. Bütün Genom Dizileme Sistemi ile dizileme küçük DNA parçacıkları ve adaptör dizilerinin pikolitre hacimlerde büyük çaplı paralel dizileme (massive paralel sequencing) ile dizilenmesi prensibine dayanmaktadır (1).

Shotgun Dizileme metodu uzun DNA parçalarını dizilemek için kullanılır. Bu metotta büyük boyutlardaki genomik DNA veya bakteri yapay kromozomları fiziksel olarak, rastgele küçük (300-800 bp) boyutlara parçalanır. Bu DNA parçacıkları, uçlarına eklenen adaptörlerden DNA yakalama boncukları tarafından tutularak bir DNA kütüphanesi oluşturulur. Daha sonra her bir boncuk tarafından yakalanan DNA parçacığı ayrı ayrı dizilenir ve okunan bütün parçalar Biyoinformatik yazılımlarla birleştirilir. Shotgun Dizileme tekniği birçok önemli grup tarafından kullanılmıştır. Bu metod ile insan gibi karmaşık genomlardan (2) Drosophila melanogaster genomuna (3) ve Haemophilus influenzae (4) gibi bakteri ve virus genomlarına uzanan birçok türün tüm genom haritası ortaya çıkarılmıştır.

De novo dizi okumaları kontigler haline getirildikten sonra bu parçalar paired-end (küt uç) okumaları ile doğru yönde dizilerek birleştirilir. Bu paired-end okumalarının herbir tarafında birbirlerinden yaklaşık olarak 2.5 kb uzaklıkta iki adet 20-mer'lik DNA parçacığı bulunur. Biyoinformatik yazılımlar sayesinde 20-merlik bu parçacıklar oluşturulan kontig'lere haritalanır ve böylece bu kontigler doğru yönde dizilerek birleştirilir. Bu birleştirilmiş diziler sayesinde yüksek kalitede ve doğrulukta genom dizisi elde edilir.

Günümüzde kullanılan yeni nesil dizileme sistemleri, Roche 454 genome analyzer, Illumina Genome Analyzer, Applied BioSystem SOLiD, Complete Genomics, Helios, Pacific Biosciences ve IonTorrent'dir.

Yeni nesil DNA dizileme teknolojisi hayal bile edilemeyecek kadar çok bilgi ve yeni yaklaşımlar sağlar. Fakat bu kadar çok bilginin depolanması, analizi ve değerlendirmesinde büyük güçlükler vardır. Yeni nesil DNA dizileme teknolojisinin başarılı bir şekilde kullanılması için gelişmiş biyoinformatik analiz araçlarına ihtiyaç duyulmaktadır. Ayrıca çok sayıda kısa okuma elde edilmeside önemli bir

problemdir. Okuma uzunluğu ve hata oranı konusu hala geliştirilmektedir.

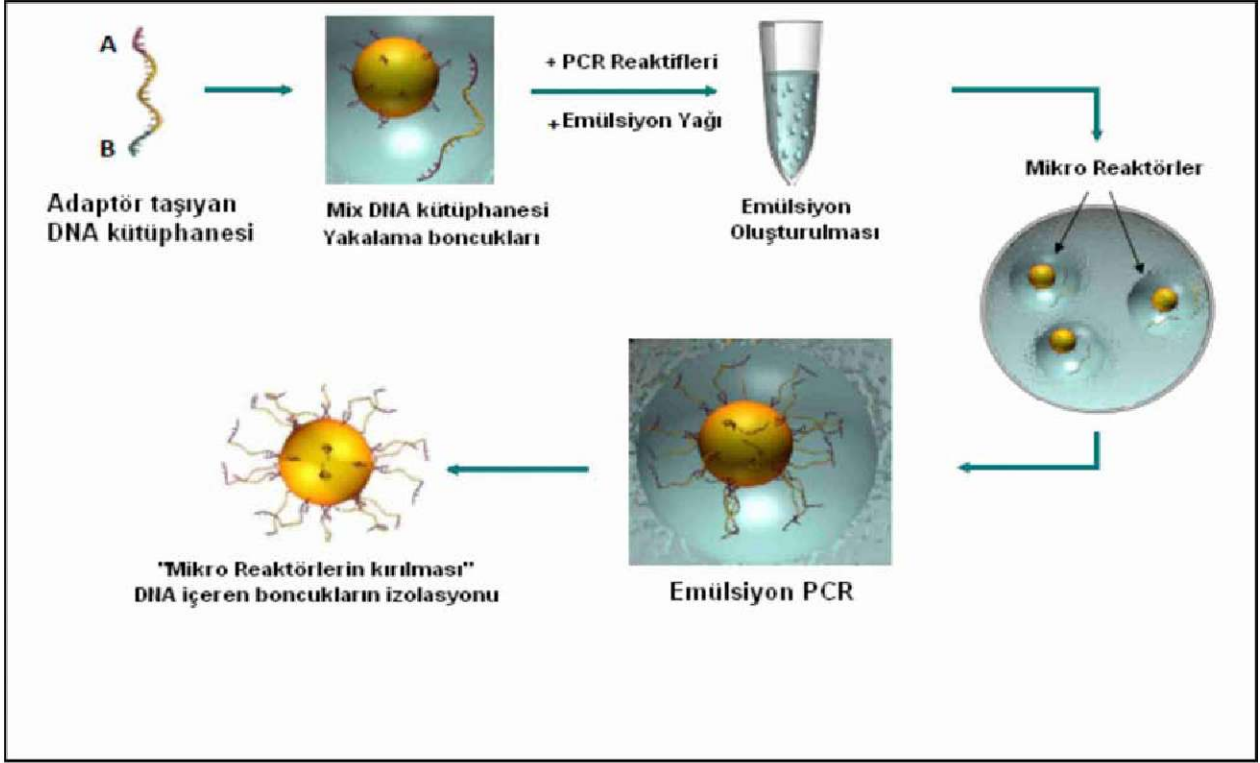
Sanger (Dideoksi) ve Yeni nesil dizileme metodları

Sanger Metodu

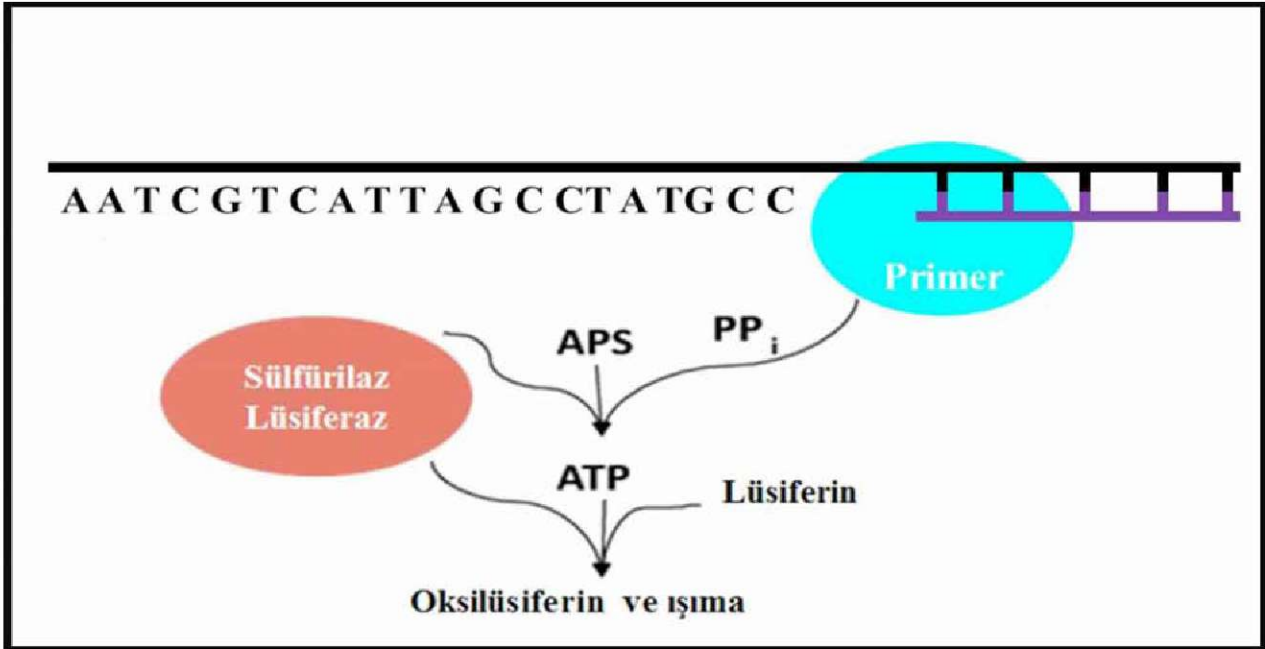
Genom sekansının belirlenmesinde en yaygın olarak kullanılan yöntem, shotgun tekniği ile Sanger (dideoksi) metodudur. Sanger metodu zincir sonlandırma metodu olarak da adlandırılır. Shotgun Dizileme metodu uzun DNA parçalarını dizilemek için kullanılır. Sanger yöntemi ile bir kerede dizi analizi yapılamayacak kadar uzun olan DNA'lar önce küçük parçalara bölünür. Elde edilen her bir parça bir plazmit klonlanır. Klonlanan plazmitler tek tek dizilenir. Bu dizilerin bioinformatik analizlerle bir araya getirilmesi ile uzun DNA parçasının dizisi elde edilir. Dizilemesi yapılacak olan DNA dizilerinin birçok kopyası primer normal deoksi nükleotidler (dNTP) dATP, dCTP, dGTP, dTTP, dideoksi nükleotidler: ddNTP; ddATP, ddCTP, ddGTP, ddTTP ve DNA Polimeraz I enzimi ile karıştırılır. DNA dört deoksiribonükleotid trifosfat (dATP, dCTP, dGTP, dTTP) molekülünün birbirine bağlanmasıyla oluşan bir polimerdir. Her yeni nükleotid bir önceki nükleotidin 3' OH ucuna eklenir. Dideoksinükleotid trifosfat (ddNTP) ise 3' OH ucu içermez. Dolayısıyla böyle bir molekül zincir uzamasını durdurur ve zincirin son nükleotidini oluşturur (5). Bu DNA dizileri kapiller jel elektroforezi kullanılarak okunur.

Piroidizileme Metodu

Sanger metodu ile uzun DNA bölgelerinin dizilenmesi için dizileme öncesi uygulanması gereken klonlama ve koloni seçimi işlemlerinin çok uzun süre alan zahmetli işlemler olması bilim adamlarını alternatif dizileme yöntemleri arayışına sokmuştur. Son yıllardaki kayda değer teknolojik yenilikler karmaşık örneklerin daha önce görülmemiş hızda ve ölçekte dizilenmesine olanak sağlamıştır. Bu doğrultuda, 1996'da Ronaghi ve arkadaşlarının "Sentez Yoluyla Dizileme" (sequencing by synthesis) prensibine dayalı "Pyrosequencing" metodunu geliştirmesi bu arayışların sonucu olarak ortaya çıkmıştır (6). Piroidizileme (pyrosequencing) yüksek işlem hacmi ve düşük maliyeti sayesinde geleneksel Sanger dizilemenin yerini almıştır.



Şekil 1: Adaptör taşıyan DNA fragmanlarının Emülsiyon bazlı PCR ile klonal amplifikasyonu (7).



Şekil 2: Pirodizileme metodu (7).

Pirodizilemenin temeli DNA polimeraz aktivitesinin kemilüminesan bir enzim aracılığıyla tespitine dayanmaktadır. Dizileme reaksiyonu, dizilenecek DNA'nın tek sarmalı (ssDNA) üzerinde tamamlayıcı sarmalının (complementary DNA / cDNA) sentezlenmesi şeklindedir. Öncelikle çift zincirli DNA'nın nebulizasyonu (400-800 bç'lik küçük fragmanlara ayrılması) gerçekleştirilir. Bu fragmanlara dizileme primerlerini taşıyan adaptörler eklenir. Adaptörleri taşıyan DNA fragmanlarının mikoreaktörler içerisinde emülsiyon bazlı klonal amplifikasyonu (em-PCR) gerçekleştirilir (Şekil 1). Bu aşama sonucunda yaklaşık olarak 10 milyon kopyaya çıkarılmış DNA fragmanları sentez yoluyla dizileme yöntemiyle yaklaşık 3.6 milyon kuyu içeren bir plate'de dizilenir. Dizileme primeriyle sabitlenen tek sarmallı kalıp DNA üzerine ardışık olarak akan A, C, G, T nükleotitlerinin kalıp DNA dizisine tamamlayıcı olması durumunda ortamda bir ışık oluşur. Bu ışığı yaratan kemilüminesan sinyalin hangi nükleotidin bağlanması sırasında oluştuğu tespit edilir. Sabitlenen ssDNA, DNA polimeraz, ATP sülfirilaz, lüsiferaz, apiraz, APS ve lusiferin ile inkübe edilir. Nükleotid akışı sırasında tamamlayıcı nükleotidin gelmesi durumunda DNA polimeraz pirofosfat (PPi) açığa çıkmasını sağlar. ATP sülfirilaz enzimi bu PPi'ı kantitatif olarak ATP'ye dönüştürür (Şekil 2). ATP, lüsiferaz enzimi aracılığıyla lusiferinin oksilüsiferine dönüşmesini sağlar. Oksilüsiferin görülebilir bir ışık yaratır. Bu ışığın şiddeti ATP miktarıyla doğru orantılıdır ve bu sayede aynı dizi üzerindeki tek nükleotid tekrarları (homopolimer) tespit edilir. Ortaya çıkan bu ışık CCD kamera tarafından kaydedilir ve bilgisayar program yardımıyla dizi verilerine (flogram) dönüştürülür (8).

Pirodizileme yöntemi kullanılarak, klonlama yapmaksızın, 7.5 saatte 100-500 milyon bazlık dizileme yapmak mümkündür. Bu sistem daha önce yıllar alan dizileme projelerini haftalar içinde sonuçlandırmayı mümkün kılmıştır. DNA'nın çift sarmallı yapısını bulan Nobel ödüllü araştırmacı Dr. James Watson'ın genomu bu yöntemle dizilenmiştir (9). Yine Neandertal genomu (10) ve nesli tükenen mamut türünün genomu (11) bu yöntemlerle kısa sürelerde bütünüyle dizilenebilmiştir. Bu yöntem bu tip karmaşık genomların haricinde, bakteri ve virus gibi daha basit genomların da bir gün içinde tüm genomunun dizilenebilmesini sağlar.

Siklik geriye dönüşümlü sonlandırma (cyclic reversible termination) ile dizileme

DNA molekülleri öncelikle slide üzerindeki primerlere bağlanıp amplifiye edilir (bridge amplification). Dört tip ddNTP eklenir ve yeni sentezlenen ipliğin yapısına girmeyen ddNTP'ler yıkanarak uzaklaştırılır. Pirodizilemeden farklı olarak, DNA, her yıkamadan sonra, bir nükleotid uzar. Floresan işaretli nükleotidler kamera aracılığı ile tespit edildikten sonra, 3' ucundaki sonlandırıcı kimyasal olarak çıkartılır ve sonraki siklusa geçilir (12).

Ligasyonu yoluyla dizileme

Bu metotta da pirodizilemedeki gibi emülsiyon PCR yapılır. Pirodizilemede olduğu gibi primerler adaptör diziler yardımıyla kalıp ipliğe hibridize olur. Bu metotta diğerlerinden farklı olarak DNA polimeraz yerine DNA ligaz enzimi kullanılır. Floresan işaretli oligonükleotid problemleri kullanılır, problemler kalıp DNA'ya hibridize olur ve primerle ligasyona girer. Floresan deteksiyonundan sonra oligonükleotid problemlerin işaretleri 5' fosfat ucu kesilerek yeni bir ligasyon siklusuna geçilir. Ligasyon, deteksiyon ve ayrılma siklusları tekrarlanarak DNA dizisi belirlenir (13).

İyon semikonduktör dizileme

Bu metod DNA'nın polimerizasyonu sırasında ortaya çıkan Hidrojen iyonlarının saptanmasına dayanır. Tek bir Kalıp DNA dizisini içeren mikrokuyuya tek tip nükleotid akışı uygulanır. Verilen nükleotid kalıp ipliğe komplementer olduğu durumda yeni oluşan ipliğin yapısına katılır ve hidrojen iyonu salınımı olur. Hidrojen iyonları hipersensitif iyon sensorleri tarafından algılanır. Kalıp DNA'da homopolimer tekrarların varlığı durumunda tek siklusta birden fazla nükleotid yeni ipliğin yapısına girecek ve salınan hidrojen iyonlarının sayısına göre daha fazla elektronik sinyal elde edilmesiyle DNA dizisi belirlenir (14).

Nano dizileme

Bu yöntem de sentez yoluyla dizileme esasına dayanır. Direkt moleküler sensor olarak, biyomühendislikle üretilmiş farklı bir polimeraz ve dNTP'ler kullanılır. Her bir nanomakinede tek bir DNA molekülünün sentezi gerçek zamanlı olarak izlenebilir (15).

Yeni Nesil Dizileme Teknolojisinin Kullanım Alanları:

1. Metagenomik

Herhangi bir çevresel örnekte var olan tüm mikrobiyolojik genomların (microbiome) rastgele dizilenmesi sonucunda ortamın mikrobiyolojik çeşitliliğini ortaya çıkarmak mümkündür. Bu çalışmalar çevresel ortamda hangi mikropların ne miktarda olduğu hakkında da fikir verdiği gibi, daha önce bilinmeyen veya çok az bilinen, sadece kendi doğal ortamında büyüyen ama kültür ortamında büyümeyen bakteriler hakkında da bilgi verir. Kullanım alanları: Su ürünleri araştırmaları, Çevre Mühendisliği, Orman Araştırmaları, Diş Hekimliği, Temel Tıp ve Klinik Araştırmalar'dır.

Metagenomik yaklaşımı, deniz suyu, atık sular, maden yatakları gibi sonuçların ekonomik değer yarattığı alanlarda daha sık kullanılmakla birlikte canlılarda kaynağı bilinmeyen enfeksiyonları tanımlamakta da kullanılır. Buna örnek olarak; A.B.D.'de milyonlarca bal arısının toplu ölümüyle arıcılık sektörünü tehdit eden sorunun kaynağında IAPV adlı bir virüsün olduğunun Cox-Foster ve arkadaşları tarafından 2007 yılında ortaya konulması gösterilebilir (16). Shotgun dizileme tekniklerinin gelişmesiyle insanlarla bakteriler arasındaki ilişkilerin açığa çıkarılmasında da Metagenomik yaklaşım daha sık kullanılmaya başlanmıştır. Turnbaugh ve arkadaşlarının yaptığı obezite ile mide florasındaki bakteriler arasındaki ilişkiyi ortaya koyan çalışma bunların ilkidir (17). İnsan sağlığı ve tıp alanlarında, hastane enfeksiyonları ve ağız ve diş sağlığı metagenomik yaklaşım için uygun alanlar olarak gelecek vaat etmektedir.

2. Bitkiler ve Tarımsal Biyoteknoloji

Bitki genomunun karmaşık yapısı ve büyüklüğü nedeniyle Sanger Dizileme teknolojisi kullanıldığında, maliyetin yüksek olmasının yanında yöntem teknolojik olarak da kısıtlı olduğundan bitki projeleri sadece büyük çaplı dizileme merkezleriyle sınırlı kalmaktadır. Yeni nesil dizileme, uzun ve yüksek doğruluk oranlı okumalarla yüksek veri kalitesi sunmaktadır ve uygun hız ve maliyet nedeniyle bitki çalışmalarında oldukça kullanışlıdır (18, 19). Soya fasulyelerinde büyük zarara yol açan *Heterodera glycines* (Soya fasulyesi kist nematodu) 400 milyon bazlık genomu Sanger dizileme yöntemi ve mikro-bead yöntemi ile dizilerek her iki yöntem karşılaştırılmıştır. Nematodun genomdaki SNP'lerin nematodun virulansı ve konak seçimini etkilediği

bilinmektedir. Bu nedenle SNP'lerinin tanımlanması tarımsal mücadelede önemlidir. Nematodun iki alt tipinin genomik DNA'sı kullanılarak her iki yöntemin dizi analizi sonuçları karşılaştırıldığında mikro-bead yöntemi ile genomdaki SNP'lerin ve "low and high copy number regions" çok daha çabuk ve doğrulukta tespit edilebildiği gösterilmiştir (20).

3. Atasal DNA

Yeni nesil dizileme sistemleri ile fosilden, saç, kemik gibi donmuş veya korunmuş örneklerden elde edilen DNA'ların geniş kapsamlı analizini yapmak mümkündür. Dizileme sonucunda, atasal hedef diziler çevresel kontaminasyondan ayırt edilebilir.

Yapılan bir çalışmada yeni nesil dizileme ile 38000 yaşında Neanderthal fosilinden elde edilen atasal DNA dizilenerek günümüz insan ve şempanze DNA'ları ile karşılaştırılmış ve Neanderthal DNA dizisinin modern insan DNA dizilerinden 500 000 yıl kadar önce birbirinden ayrıldığı tespit edilmiş (9). Bu çalışmada Neanderthal kemiğinden (0.3 gr) elde edilen parçalanmış 8341 mtDNA dizisi yazılım programı ile birleştirilerek 4.8 Gb'lık tam mtDNA dizisi elde edilmiştir. Neanderthal mtDNA'sı ile modern insan mtDNA'sı dizilenerek karşılaştırılmış ve Neanderthal mtDNA'sında kodlanan 13 protein ile elektron transport zincirinde sitokrom-c oksidaz'ın 2 alt biriminde çok sayıda amino asit değişimleri tespit edilmiştir (21).

4. Transkriptom

Transkriptom dizileme, mRNA transkripsiyon-ekspresyon analizi, yeni genlerin keşfi, henüz genomu tamamıyla tanımlanmamış organizmalarda gen bölgelerinin tanımlanması, tüm gen dizisinin oluşturulması, tek nükleotid polimorfizmlerin (SNP) belirlenmesi, insersiyon-delesyonlar ve splicing varyantların tespiti, allel spesifik ekspresyonların analizi ve kromozomal yeni düzenlenmelerin belirlenmesini kapsar.

Transkriptom analizinde başlangıçta 3-5 ug çift iplikli cDNA gerekmektedir. Bu yöntemde RNA'nın Uzun Poly-A kuyrukları dizileme aşamasından problem oluşturabileceğinden önce bunlar cDNA sentez aşamasında iki yol ile uzaklaştırılırlar. İlki, bilinen restriksiyon enzim kesim noktası içeren Poly-T primerleri ile cDNA'lar sentezletilir ve enzim kesimi ile Poly-A kuyrukları uzaklaştırır (22, 23). İkincisi ise, modifiye edilmiş Timin nükleotidlerini içeren kısa primerler ile cDNA sentezi

yapılarak Poly-A kuyrukları içerenler seçilip uzaklaştırılır (24). Daha sonra kütüphane hazırlanma aşamasına geçilir. cDNA, 800 baz çifti civarında ise adaptörler eklenir. cDNA fragmentlerinin uzunluğu 1- 2 kb arasında ise bu fragmentler adaptör eklemek için çok uzun, nebulize etmek için ise çok kısadır. Bu nedenle, bu cDNA fragmanları restriksiyon enzimi ile kesilip Random Transcriptome Sequencing (TSEQ) Protokol'ü uygulanır (25, 26).

5. Küçük RNA'lar (Small RNA'lar)

Küçük RNA'lar; MicroRNA'lar, endojen siRNA'lar ve Piwi-interacting RNA'ların dahil olduğu gruptur. Bu sistem sayesinde bakteri içerisinde klonlanma yapılmadan, yüzlerce hatta binlercesi tanımlanabilir ve miktarı belirlenebilir. Daha önce tanımlanmamış küçük RNA'lar tespit edilebilir ve farklı ekspresyon profili çıkartılabilir veya küçük RNA transkriptomları ile karşılaştırılabilir (27, 28). Küçük RNA'ların analizi için öncelikle total RNA (~12.5%) denatüre poliakrilamid jelde yürütülerek fragmanların büyüklüklerine göre ayrılması sağlanır. Small RNA büyüklüğü 15-30 baz arasında olduğu için bu büyüklüğe denk gelen bölge jelden alınarak RNA saflaştırılır. Önce 3' ucuna, sonra 5' ucuna T4 RNA ligaz ile adaptör eklenir ve oligo (dT)-adaptör primerleri kullanılarak tek iplik cDNA sentezi yapılır. Daha sonra 22 döngülük PCR yapılarak cDNA'lar çoğaltılır. Bu aşamada iki farklı yol izlenebilir. İlki adaptör ligasyonu yapıp Shotgun yöntemi ile dizileme yapmaktır. İkincisi ise PCR yöntemi ile 19-mer kuyruklu adaptörler eklenip, amplikon paired-end ile dizileme yapmaktır.

6. Metilasyon Analizi

Gen aktivasyonunun epigenetik düzenlenmesi, gelişimsel süreçte son derece önemlidir. Son yıllarda yapılan çalışmalar, DNA metilasyonundaki değişikliklerin kanser, nörolojik hastalıklar (Şizofreni, Alzheimer) ve kardiyovasküler hastalıklar (Ateroskleroz, Homosisteinemi) ile ilişki olduğunu göstermiştir. Genom boyu metilasyon kaybı ya da bazı genlerin promotor bölgelerindeki CpG adacıklarının hipermetilasyonu, bu hastalıkların gelişiminde son derece önemlidir. Kanserlerin ailevi formlarıyla ilişkili genlerin yaklaşık % 50'sinde promotor hipermetilasyonu görülmüştür. Bu sistem ile hedef genomik dizideki metilasyon belirlenir. Ayrıca hipermetile olan CpG dinükleotidlerinde birbirleri arasında metilasyon farkları tespit edilebilir.

Örnekler Bisülfid ile muamale edilir. Metile olmayan

sitozinler urasile dönüştürerek deamine olurken, metile sitozinler değişmeden kalır. PCR Amplifikasyonu ile urasile dönüştürülen sitozinlerin karşısına timinler gelir. Bu çalışmada primer dizayn önemlidir. Bilinen dizi için, transkripsiyon başlangıç noktasından 1000 baz gerisi ile birinci ekzondan 500 baz somaki bölge arasında dizayn edilmelidir. Literatür bilgisi olmayan dizilerde ise CpG adalarının bulunduğu SP1 proteini bağlanma bölgesine yakın alanlar tercih edilmelidir (29).

7. Kromatin İmmunopresipitasyon Dizileme (ChIP-Seq)

Kromatin İmmunopresipitasyon (ChIP) sonucu elde edilen kromatin materyalinin high-throughput paralel dizi analizi ChIP-Seq olarak adlandırılmış olup ilk kez 2007 yılında uygulanmıştır (30). Bu yöntemde DNA'ya bağlanan proteinlerin (transkripsiyon faktörleri, histon proteinleri vb.) global olarak, genom üzerindeki tüm bağlanma bölgelerinin dizilenmesi mümkündür. Genom boyunca bağlanma bölgelerinin analizi gen ekspresyonunu düzenleyici mekanizmaların anlaşılması için gereklidir.

Transkripsiyon faktörlerinin bağlanma bölgeleri bir kaç yıl öncesine kadar ChIP yöntemi ile araştırılmaktaydı. ChIP yöntemi yalnız aday genlerin taranmasına olanak sağladığından oldukça sınırlıdır. Microarray teknolojisinin geliştirilmesiyle ChIP on chip yöntemi kullanılmaya başlanmıştır (31). Fakat ChIP on chip yöntemi de kullanılan oligonükleotid problemleri sınırlı olup tüm genom analizi mümkün değildir. Son yıllarda geliştirilen ChIP-PET (pair-end ditags) yöntemiyle transkripsiyon faktörlerinin bağlanma bölgelerini genom boyunca analiz etmek mümkündür. Fakat ChIP-PET yöntemi bazı klonlama basamakları içermektedir ve klonlama sonucu elde edilen DNA'lar sanger yöntemine dayalı dizileme sistemleriyle analiz edildiğinden çok sayıda dizi analizi gerektirmektedir (32). Bundan dolayı pratik bir yöntem olmadığı gibi yüksek maliyetlidir. Yeni nesil dizileme ile tüm bu zorlukların aşılması mümkün olmaktadır.

ChIP dizileme için öncelikle protein-DNA bağlantıları formaldehit ile sabitlenir. Ardından DNA sonikasyonla veya enzimatik yollarla 100-500bp'lik fragmentlere ayrılır. Araştırılan proteine özgün antikorla immunopresipitasyon yapılarak, proteinin bağlandığı DNA bölgeleri elde edilir. Ardından DNA saflaştırılarak yeni nesil dizileme ile dizilenir (33).

Sonuç

Yeni nesil DNA dizileme teknolojisi bütün dünyada yeni olduğu gibi ülkemizde de oldukça yenidir. İstanbul üniversitesi, Deneysel Tıp Araştırma Enstitüsü, bu teknolojiye zamanında doğru yatırım yaparak bu alanda ülkemizde öncü olmuştur.

Kazanmış olduğumuz yeni nesil dizileme ve biyoinformatik deneyim, araştırmalarımızda bize güç ve avantaj sağladığı gibi, üniversitemiz de güçlü bir genom merkezine kavuşmuştur.

Referanslar

1. Margulies M, Egholm M, Aitman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005 Sep 15;437(7057):376-80
2. Venter JC, Adams MD, Sutton GG, Kerlavage AR, Smith HO, Hunkapiller M. Shotgun sequencing of the human genome. *Science* 1998, 280, 1540-1542
3. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, The genome sequence of *Drosophila melanogaster*. *Science*. 2000 Mar 24;287(5461):2185-95
4. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*. 1995 Jul 28;269(5223):496-512
5. Sanger F, Nicklen S, and Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc. Natl.Acad. Sci.*1977; 74:5463-5467
6. Ronaghi M, Nygren M, Lundeberg J, Nyrén P. Analyses of secondary structures in DNA by pyrosequencing. *Anal Biochem*. 1999; Feb 1;267(1):65-71
7. [http://454.com/downloads/my454/documentation/gs-flx/method-manuals/GS-FLX-Titanium-emPCR-Method-Manual-\(Apr2009\)](http://454.com/downloads/my454/documentation/gs-flx/method-manuals/GS-FLX-Titanium-emPCR-Method-Manual-(Apr2009))
8. Ronaghi M. Pyrosequencing sheds light on DNA sequencing. *Genome Res*. 2001; Jan;11(1):3-11
9. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008; 452: 872-877
10. N4 Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, et al. Analysis of one million base pairs of Neanderthal DNA. *Nature* 2006; 444: 330-336
11. Gilbert MT, Binladen J, Miller W, Wiuf C, Willerslev E, Poinar H, et al. Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. *Nucleic Acids Res*. 2007;35(1):1-10
12. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 2008; 9: 387-402
13. Schuster SC. Next-generation sequencing transforms today's biology. *Nat. Methods* 2008. 5 (1): 16-8
14. Rusk N. Torrents of sequence. *Nature Methods* 2011; 8(1)44
15. Treffer R, Deckert V. Recent advances in single-molecule sequencing. *Current Opinion in Biotechnology* 2010; 21(1): 4-11
16. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, Quan PL, Briese T, Hornig M, Geiser DM, Martinson V, vanEngelsdorp D, Kalkstein AL, Drysdale A, Hui J, Zhai J, Cui L, Hutchison SK, Simons JF, Egholm M, Pettis JS, Lipkin WI. A metagenomic survey of microbes

in honey bee colony collapse disorder. *Science*. 2007 Oct 12;318(5848):283-7.

17. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Center for Genome Sciences, Washington University, St. Louis, Missouri 63108, USA. *Nature*. 2006 Dec 21;444(7122):1027-31.

18. Weber AP, Weber KL, Carr K, Wilkerson C, Ohlrogge JB. Sampling the Arabidopsis transcriptome with massively parallel pyrosequencing. *Plant Physiology* 2007; 144: 32-42

19. Novaes E, Drost DR, Farmerie WG, Pappas GJ Jr, Grattapaglia D, Sederoff RR, Kirst M. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* 2008; 9: 312

20. Bekal S, Craig JP, Hudson ME, Niblack TL, Domier LL, Lambert KN. Genomic DNA sequence comparison between two inbred soybean cyst nematode biotypes facilitated by massively parallel 454 micro-bead sequencing. *Molecular Genetic and Genomics* 2008; 279: 535-543

21. Green RE, Malaspina AS, Krause J, Briggs AW, Johnson PL, Uhler C, Meyer M, Good JM. A complete Neanderthal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* 2008; 134: 416-426

22. Ng P, Tan JJ, Ooi HS, Lee YL, Chiu KP, Fullwood MJ, Srinivasan KG, Perbost C, Du L, Sung WK, Wei CL, Ruan Y. Multiplex sequencing of paired-end ditags (MS-PET) : a strategy for the ultra-high-throughput analysis of transcriptomes and genomes. *Nucleic acids res.* 2006; 13;34(12):e84

23. Ng P, Wei CL, Ruan Y. Paired-end ditagging for transcriptome and genome analysis. *Curr. Protoc. Mol. Biol.* 2007; 21,12

24. Beldade P, Rudd S, Gruber JD, Long AD. A wing expressed sequence tag resource for *Glycyphana aculeata* butterflies, an evo-devo model. *BMC Genomics* 2006; 7,130

25. Sugarbaker DJ, Richards WG, Gordon GJ, Dong L,

De Rienzo A, Maulik G, Glickman JN, Chirieac LR, Hartman ML, Taillon BE, Du L, Transcriptome sequencing of malignant pleural mesothelioma tumors. *Proc. Natl. Acad. Sci. USA* 2008; 105, 3521-3526

26. Toth AL, Varala K, Newman TC, Miguez FE, Hutchison SK, Willoughby DA, Simons JF, Egholm M, Hunt JH, Hudson ME, Robinson GE. Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* 2007; 318,441-444

27. Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Research* 2007; 35: e130

28. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 2008; 453:539-43

29. Taylor KH, Kramer RS, Davis JW, Guo J, Duff DJ, Xu D, Caldwell CW, Shi H. Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. *Cancer Res.* 2007 Sep; 15;67(18):8511-8

30. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods.* 2007 Aug; 4(8):651-7

31. Wang JC. Finding primary targets of transcriptional regulators. *Cell Cycle.* 2005 Mar; 4(3):356-8

32. Buisine N, Sachs L. Impact of genome assembly status on ChIP-Seq and ChIP-PET data mapping. *BMC Res Notes.* 2009 Dec; 16;2:257

33. Farnham PJ. Insights from genomic profiling of transcription factors. *Nat Rev Genet.* 2009 Sep; 10(9):605-16