

Non-coding RNA Molekülleri

Dersin Sorumlusu: Dr. Hatice MERGEN

Dersin adı: Proteomik ve Genomik

Hazırlayan: Çiğdem KAPLAN

2004 – 2005 GÜZ

Non-Coding RNAs (Kodlama Yapmayan RNA'lar)

Genom projesinin ilk hedefi, sistematik şekilde genlerin tanımlanmasıydı. 2000 yılında, insan genom dizisinin taslağı yayınlandı, fakat tahmin edilen insan genlerinin sayısı sürekli azalıp arttı. En son tahminler ise 30.000-40.000 gen üzerinde durmaktayken, genel olarak 100.000 ve üzerinde olduğu sanılmaktaydı. Bu belirsizliğin sürmesinin nedeni genlerin çok iyi tanımlanabildiği veya kolayca belirlenebildiği bir yöntemin henüz geliştirilememiş olmasındandır. Genelde sayısal yöntemler 3 metod üzerine oluşturulmuştur;

1. cDNA cloning
2. Poliadenillenmiş mRNA'ların EST (expressed sequence tag) dizilenmesi
3. Karşılaştırılmalı genom analizi ile korunmuş kodlayıcı ekzonların tanımlanması

Bu metodlar geniş, oldukça fazla eksprese edilen, evrimsel olarak iyi korunmuş, protein-kodlayıcı genler için mükemmel yöntemlerdir ve gen sayısının daha az olduğunun tahmin edilmesinde neredeyse kesin sonuçlar verilmişlerdir. Bu metodlar tüm genler içinde yalnızca bir tanesi için çalışmaz, bu da yapısal, katalitik veya düzenleyici RNA'lar gibi doğrudan fonksiyonu olan transkriptleri oluşturan non-coding RNA genleridir.

İnsanlarda transkripsiyonel verimin %98'ini non-coding RNA'lar oluşturur. Protein olarak kodlanmayan bu RNA'ları oluşturan genler ise non-coding RNA (ncRNA) genleri olarak adlandırılmaktadır. RNA, diğer RNA'ların spesifik olarak tanınmasında oldukça işe yarayan bir moleküldür; çünkü diğer RNAlar ile tamamlayıcı baz eşleşmesi meydana getirebilir. Evrim bazı durumlarda, ncRNA'ları proteinlerin yerine tercih etmiş olabilir. Örneğin, spesifik mRNA'lar ile ilişkiye geçen post-transkripsiyonel düzenleyici moleküllerde RNA molekülleridir. ncRNA genleri proteinleri kodlamak yerine fonksiyonel RNA moleküllerini oluştururlar. Small nükleolar RNA'ların iki büyük ailesinin rRNA'ların site-spesifik modifikasyonlarında yer aldığı ve bu ailelere ait bilinen genlerin sayısının hızla arttığı bilinmektedir. 3 yeni yayında geniş bir RNA gen ailesinin tanımı yapılmıştır ve microRNA'lar – küçük 21-24 nükleotidlik RNA'lardır ve protein kodlayan mRNA'ların translosyonal regülatör olarak görev alma ihtimalleri yüksektir. Ayrıca dört yeni yayında ise, fonksiyonu henüz bilinmeyen 30'dan fazla yeni non-coding transkriptin deneysel olarak doğrulanması için *E.coli*'deki yeni ncRNA genlerinin taranma yolları tanımlanmıştır.

İnsan genom dizisinin tamamlanması ve protein kodlayıcı genlerin tanımlanmasının bitirilmesine rağmen, ncRNA genleri halen etkin şekilde ortaya çıkarılamamıştır. Son yıllarda pek çok sistematik tarama ile yeni ncRNA genlerinin büyük bir kısmı tanımlanmıştır. Yeni non-coding RNA genlerinin keşfedilmesinde;

1. bilgisayar bazlı karşılaştırmalı genom analizleri güçlü bir yaklaşım olarak kabul edilmektedir.
2. cDNA cloning stratejilerinin ncRNA çoğaltılması için özel olarak dizayn edilmesi de iyi sonuç verebilen bir yöntemdir.
3. Prensipte mümkün olabilen, yeni transkriptlerin (hem ncRNA hem de protein kodlayan RNA) tespit edilmesi için yüksek yoğunluklu oligonükleotid mikroarraylerinin kullanılması da verimli sonuçlar verebilir.

ncRNA'lar kompleks kataliz olmaksızın oldukça spesifik nükleik asit tanınmasında önemli bir rol üstlenirler. Örneğin, gen ekspresyonunun post-transkripsiyonel düzenlenmesini yönlendirir veya RNA modifikasyonlarının yönlenmesinde görev alırlar. Bunların yanı sıra, non-coding RNA'ların, translasyonel reseptör olarak (miRNA- lin4), mRNA degradasyonunun yönlendirilmesinde (siRNA) ve translasyonel düzenleyiciler olarak (miRNA) rol aldıkları belirlenmiştir.

Çizelge: Non-coding RNA'ların farklı sınıflarına ait kısaltmalar

frNA	Fonksiyonel RNA- non-coding RNA ile eş anlamlıdır
miRNA	MicroRNA – Translasyonel düzenleyici gen ailesi olarak kabul edilir
ncRNA	Non-coding RNA – mRNA hariç diğer tüm RNA'lar
rRNA	Ribozomal RNA
siRNA	Small interfering RNA – RNA interferencede aktif moleküller
snRNA	Small nuclear RNA – splisomal RNA'ları içerir
snmRNA	Small non mRNA – small ncRNA'larla eş anlamlıdır
snoRNA	Small nucleolar RNA – rRNA modifikasyonunda yer alan pek çoğu bilinen snoRNA'lar
stRNA	Small temporal RNA – lin4 ve let7 gibi <i>Caenorhabditis elegans</i> tRNA'sı

RNA aracılığıyla yapılan gen regülasyonu yüksek ökaryotlar ve kompleks genetik fenomende geniş bir yayılım gösterir. Bu kompleks genetik fenomen içinde, RNA interference, co-supression, transgene silecing, imprinting, methylation, RNA signalling'i

temel alan ya da buna bağılı kesişen yolların tamamını içerir. Non-coding RNA'ların ökaryotik organizmalarda yer aldıklarına dair bulunmuş kanıtların yanı sıra *E.coli* gibi en iyi çalışılmış prokaryotik bir organizmada da yer aldığı bulunmuştur. *E.coli*'de tespit edilen non-coding RNA'ların sayısı gün geçtikçe açığa çıkmaktadır, ayrıca “riboregülatörler” olarak işlev gören sabit küçük RNA'larında varlığı da rapor edilmiştir. Riboregülatörler ise translasyonel başlangıç bölgeleri ile özgül ilişkiyi baz eşleşmesiyle gerçekleştirerek ya translasyonu baskırlar ya da aktive edici olarak görev görürler. *E.coli*'de protein kodlayan 4200 gen vardır. Rivas ve arkadaşları, yanlış pozitif oranın simulasyon çalışmasıyla bu 4200 protein kodlayıcı genden 200-275 gen tahmininin gerçekte non-coding RNA'lar olabileceği ileri sürülmüştür.

Hücrel fonksiyonun temel efektörleri proteinler olmasına rağmen, ökaryotik kompleksliğin temeli ve fenotipik çeşitlilik başlıca bir kontrol mimarisine dayanıyor olabilir. Bu mimari ise gen ekspresyonunun koordinasyonu ve modülasyonunda gerekli bilgiye sahip trans-actin RNA'ların oldukça paralel bir sisteminin bileşimidir. Çünkü gen ekspresyonunun koordinasyonu ve modülasyonu kromozomların yeniden modellenmesinde, RNA-DNA, RNA-RNA ve RNA-protein etkileşimleri ile gerçekleştirilir.

Genom projesi sonucunda özellikle yüksek organizmalardaki gelişimsel kompleksliğin çözülmesi hedeflenen amaçlardan biriydi. Ancak görüldü ki, kompleks organizmalar beklenenden çok daha az sayıda protein kodlayıcı gene sahipti. Örneğin, *Drosophila melanogaster* ve bir nematod olan *Caenorhabditis elegans*'sın genomlarındaki protein kodlayan genlerin (~12-14.000), *Saccharomyces cerevisiae* gibi bir mikroorganizmanın (6200) ancak iki katı olduğu, insanların ise bunlardan sadece iki kat fazla gene (~ 30-40.000) sahip olduğu bulunmuştur. Yüksek organizmalardaki ekspresse edilen protein izoformlarının repertuarı alternatif splicing ile oldukça artarken, yüksek organizmaların evriminin diğer çarpıcı bir özelliği de non-protein-coding RNA miktarının çok yüksek olmasıdır.

Yüksek organizmalar ait proteomlar nispeten durağandır. Çünkü insan ve farede protein kodlayan genlerin %99'u ortaktır ve bu canlılardaki değişikliklerin, modüler tekrar kullanma ve proteomların farklı alt birimlerinin çoklu kullanımıyla başarıya ulaştığı sanılmaktadır. Bir bireyde diğer insanlara göre her haploid genomun ~3.000.000 dizisinin farklı olmasına rağmen, bu yalnızca ~10.000 (%0.3) protein kodlayan dizide yer alır ve

bunların çoğu da sessiz (3. baz) deęiřimidir. Bu nedenle, her bireye ait ve trler arasındaki fenotipik çeřitlilik byk oranda non-protein-coding dizilerdeki farklılıklardan ileri geliyor olabilir ve kontrol mimari sisteminin kontrol gibi gen ekspresyonundaki çeřitlilikteki bir problemten de kaynaklanıyor olabilir. Kontrol mimarisi ise cis-acting promotorlar ve enhancerlarda lokalize olmuř grnse de, bu cevabın yalnızca bir parçasıdır.

Non-coding RNA'lar incelendięinde, bunların geliřimsel olarak dzenlendikleri ve genetik etkilere sahip oldukları grlmřtir. karyotlarda genetik rnlerin byk bir kısmı olan ntronik RNA ve dięer non-coding RNA'ların yksek dizi komplekslilięine sahip olduęu, ayrıca bunlarda byk evrimsel uzaklıkların korunduęu gsterilmiřtir. Hem intronların hem de non-coding RNA'ların fonksiyonel oldukları ile bilgiler ortaya çıkmaya bařlamıřtır. İlginç bir rnek olarak, small temporal RNA'lar olan lin-4 ve let7 translasyon ve dięer transkriptlerin kalıcılıęında etkili olan RNA-RNA etkileřimleri aracılıęıyla *C.elegans*'da geliřimsel zamanlamayı (developmental timing) kontrol ederler. Ayrıca hem bitki hem de hayvanlardaki geliřim srecinde rol olan bařka non-coding RNA'lar da vardır; Xist, roX1/roX2... vs). Bu RNA'ların bazıları alternatif splicing'e uęrar veya alternatif poliadenilasyon blgelerine sahiptir ve muhtemelen protein kodlayıcı kapasitelerini kaybetmiř genlerden tremiřlerdir. Non-coding RNA'ların pek çok grevinin yanında kromatin mimarisinin dzenlenmesinde de yer aldıklarına dair kanıtlar bulunmaktadır. Kromozom remodeling olayını dzenleyen ve histon asetiltransferazlar iin gerekli bir steroid reseptrnn transkripsiyonel co-aktivatrnn bir non-coding RNA olduęu bulunmuřtur.

Non-coding RNA genlerinin genetik hastalıklarla da iliřkili olduęu bilinmektedir. Cartilage-hair hypoplasia (CHH), kısa kollu ccelik hastalıęının 267 nkleotidlik RMRP ncRNA genindeki mutasyonları ile iliřkili olduęu belirlenmiřtir. Otozomal-dominant disktratis congenita hastalarının ise telomeraz RNA'larında mutasyon tařıdıęı ve bununda gen bankasında yer alan aday bir gen olduęu bilinmektedir.

İnsan, fare ve dięer memeli genomları karřılařtırıldıęında, bu genomların protein-coding blgeler dıřındaki dizi homoloji derecesinin de olduka yksek olduęu gsterilmiřtir. Yksek organizmaların biyolojisinin anlařılmasının doęru řekilde gerekleřtirilebilmesi iin, yalnızca proteomların anlařılması yeterli olmayacaktır. Bunun yanında, non-coding RNA'ların tanımlanması, bunların ekspresyon paternlerinin, iřlenmelerinin ve sinyallenme yollarının da tanımlanması gereklidir.