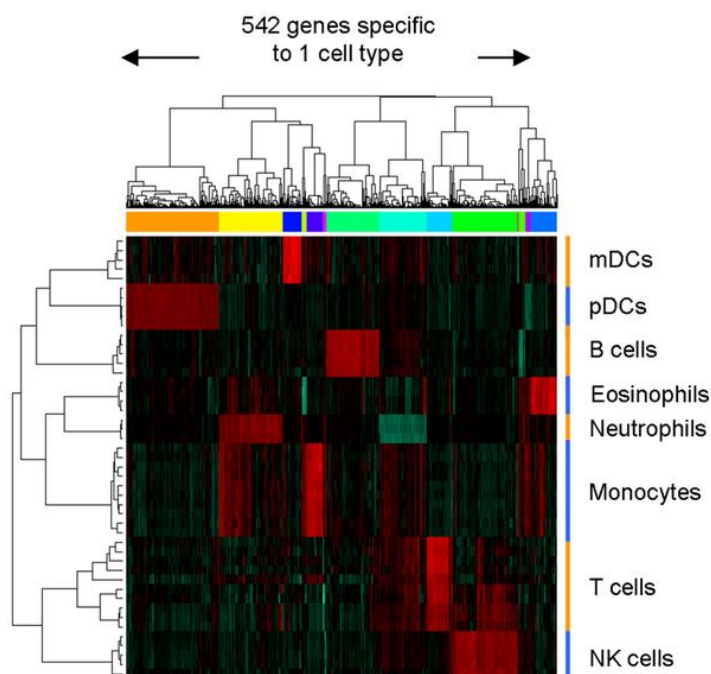


KIS RNS-MOLEKULÁK

NEMCSAK TRANSZKRIPCIÓS ZAJ – A GÉNSZABÁLYOZÁS ÚJ SZINTJE

Molnár Viktor, Falus András

A többsejtű szervezetekben az egyes szövetek közötti munkamegosztás meghatározott feladatokra specializálódott sejtek kialakulásával valósulhat meg. Ahhoz, hogy egy T-sejt vagy éppen egy idegsejt elláthassa a feladatát (antigének felismerése és aktiváció; illetve az ingerület érzékelése és továbbítása), az ehhez szükséges fehérjekészlettel kell rendelkeznie. Ez a fehérjemintázat szerepe persze túlmutat a kémiaiilag fehérjét tartalmazó rendszerek felépítésén (mint a strukturális fehérjék, jelátvivő ligandok és receptorok, transzkripciós faktorok stb.), hanem a megfelelő anyagcsere-enzimek előállításával a sejt lipid- és szénhidrát-összetételét is meghatározza. A terv, amely a specializált sejtekben megszabja a különböző fehérjéket kódoló DNS szakaszok „termelékenységét” (az egész genom szintjén: a génextpressziós mintázatot), az egyedfejlődés során alakul ki. A fejlődési folyamat végén aztán többé-kevésbé állandó marad a sejtre vagy szövetre jellemző génextpressziós profil. A sokoldalú igények és főként alkalmazkodási képesség miatt azonban szükséges lehet a terv átmeneti (esetleg tartós) módosítása (pl. a T-sejtek aktivációja során a jelátvivő lánc végén álló transzkripciós faktorok által indukált célgének bekapcsolása, illetve átírásának magasabb fokozatba történő kapcsolása: citokinek, adhéziós molekulák, perforin, granzym). Ennek lehetőségével számolva a sejtek képesek a kifejeződő fehérjék előállítási folyamatát több szinten biztosítva, igen pontosan szabályozni.



1. ábra. Emberi vérben található immunsejtek jellemző génextpressziós mintázata. A teljes genom szintű vizsgálatban 696 gén esetében teljesült a feltétel, hogy megjelenése egy, két, vagy három sejt típusra legyen jellemző, azaz jelentősen különbözzön az összes többi típustól. Ezek között 542 expressziója csak egyetlen típusra volt jellemző. A hőtérképen hierarchikusan klaszterzéssel rendezett expressziós értékek láthatóak, pirossal a viszonylag magasabb, zölddel az alacsonyabb szinten kifejeződők kerültek ábrázolásra. A színek intenzitása arányos a relatív különbségekkel.

Allantaz F, Cheng DT, et al. (2012) Expression Profiling of Human Immune Cell Subsets Identifies miRNA-mRNA Regulatory Relationships Correlated with Cell Type Specific Expression. *PLoS ONE* 7(1): e29979. doi:10.1371/journal.pone.0029979

A szabályozás célpontja lehet egyetlen fehérje, amikor például egy kiemelten fontos (energiaellátás) vagy éppen veszélyes (programozott sejthalál) folyamatban központi szerepet játszó

elemként erős és egyértelmű utasításokra van szükség. Egy magasabb szinten pedig lehetséges akár a programok közötti váltás is, számos célpont kifejeződésének egyidejű befolyásolásával.

A genetikai információ kifejeződését, illetve annak ütemét több szinten történő beavatkozási lehetőség alakítja. A kromatinstruktúra kondenzáltsága (azaz hisztonok által alakított hozzáférhetősége), a metilációs minta, a rendelkezésre álló transzkripciós faktorok, a DNS-ről átíródó hírvivő RNS-ek életideje, a transláció folyamatát módosító tényezők stb. egyaránt döntőek lehetnek a végeredmény, a funkcióképes fehérje megjelenésében.

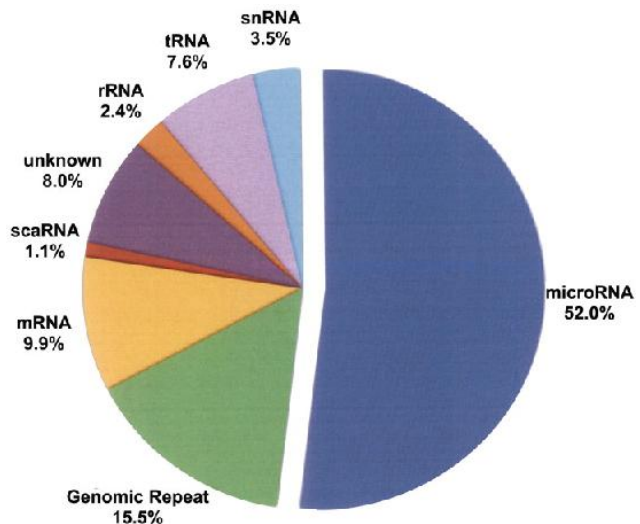
A nem-kódoló RNS-családok

A nem-kódoló RNS-ek (ncRNS; non-coding RNA) mind ezidáig kívül estek a vizsgálódások fókuszán, talán éppen annak a karakterisztikus tulajdonságuknak köszönhetően, hogy sosem fordítódnak le fehérjékké, megnehezítve a szerepük értelmezését. Ezen genetikai információt kódoló információkat a DNS junk, vagyis szemetet tartalmazó szegmenseihez sorolták, amelyet meg kellene a transzkripció zajától különböztetni. Ugyanakkor a becslések szerint ez a fehérjét nem kódoló RNS az emberi sejt transzkripciós kimenetének akár 97-98%-át is alkothatja. A DNS-ről átíródó összes RNS-szekvenciák csekély hányada játszik tehát szerepet a fehérjék aminosav-sorrendjének meghatározásában, amit a teljes transzkriptomot átfogó szekvenálással kapott eredmények is alátámasztanak. Az RNS-interferencia¹, és ennek kapcsán a mikroRNS-ek, felfedezését követően a genom fehérjét nem kódoló szakaszaihoz való hozzáállás alapvetően megváltozott, és az élettudományi kutatások megkezdték az újonnan definiált génexpressziós szabályozási rendszer töredék-információkból való összeállítását.

Vajon miért van szükség erre az iszonyú pazarlásra? Ha azonban jobban megfontoljuk, kiderül, hogy jóval olcsóbb a szabályozást RNS-szinten erősíteni, és ezen a szinten dönteni egy fehérjekezdemény sorsáról, amikor a sejt számára roppant költséges fehérjék még nem szerelődtek össze. A szabályozás összetettségével nőnek az élőlény rugalmas alkalmazkodása és fejlődési lehetőségei. A biológiai komplexitás (amelyben az alacsonyabb- és magasabbrendű fajok közötti különbséget kifejezhetjük) kétségtelenül független az élőlényt leíró adat mennyiségétől, mint a gének száma vagy a DNS-tartalom. A bonyolultság hátterében inkább a jóval összetettebb és kifinomultabb szabályozás sejthető, amelynek hátterében részben a kémiai RNS-nek megfelelő molekulák állhatnak. A kis nem-kódoló RNS-ek vizsgálata már eddig is számos, előzőleg meg sem jósolt szabályozási kapcsolatra derített fényt. Mára már ismertek a nem-kódoló RNS különböző fajtáiról, hogy azok a genetikai információ megnyilvánulásának több szintjén képesek hatásukat kifejteni, támadáspontjuk lehet a kromatinstruktúra megváltoztatása, a DNS imprinting, transzkripció, illetve a transláció szintjén jelentkező hatás. Ezzel párhuzamosan számos törekvés indult az új nem-kódoló családok azonosítására a genom ismert szekvenciáit tartalmazó adathalmazában.

¹ RNS-interferencia: RNS szakaszok közötti szekvencia-komplementaritáson alapuló kölcsönhatás, amellyel az élő sejtek géneinek aktivitását specifikusan szabályozni tudják poszttranszkripcionális szinten.

2. ábra. Emberi embrionális őssejt-tenyészetben előforduló RNS-családok csoportjai és azok reprezentációja. A kísérlet során készített „kis RNS”-könyvtár rövid szekvenciadarabjait olvasták le, majd azokat a referenciaszekvenciára illesztésükkel azonosították be. Az összesen több mint 1.6 millió egyedi darabot a fenti 8 osztály szerint sorolták be. (scaRNS=kis Cajal-test-specifikus RNS, az snoRNS-ek egy alcsoportja, amelyek az snRNS módosításáért felelősek)
Morin RD et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res. 2008. 18: 610-621



A nem kódoló RNS-ek gyűjtőfogalma (non-coding RNAs) alá azok a transzkriptumok tartoznak, amelyek nem kerülnek transzlációra (fehérjét nem kódoló RNS, non-protein-coding RNA). Feloszthatóak kis (snoRNS, small nucleolar RNA; piRNS, Piwi interacting RNA; stb.) és a közepes/nagy mérettartományba tartozó RNS-ek családjára, viszont a mikroRNS-ek, népszerűségüknek is köszönhetően, gyakran sorolják önálló kategóriába. Egy másik rendszerezés szerint az RNS-ek családjai a jól ismert riboszómális, transzfer és hírvivő RNS-eket a háztartási, míg az újonnan felismerteket tömörítő csoport a szabályozó kategóriákra oszthatóak.

<ul style="list-style-type: none"> • hírvivő RNS (mRNS) • transzfer RNS (tRNS) • riboszómális RNS (rRNS) • kis magi/nukleáris RNS (snRNS) • kis magvacskai/nukleoláris RNS (snoRNS) 	<p>háztartási RNS</p>
<ul style="list-style-type: none"> • mikroRNS (miRNS) • piwi RNS (piRNS) • ismétlődés-associált kis interferáló RNS (rasiRNS) • hosszú nem-kódoló RNS • [kis interferáló RNS (siRNS)], ...? 	<p>szabályozó RNS</p>

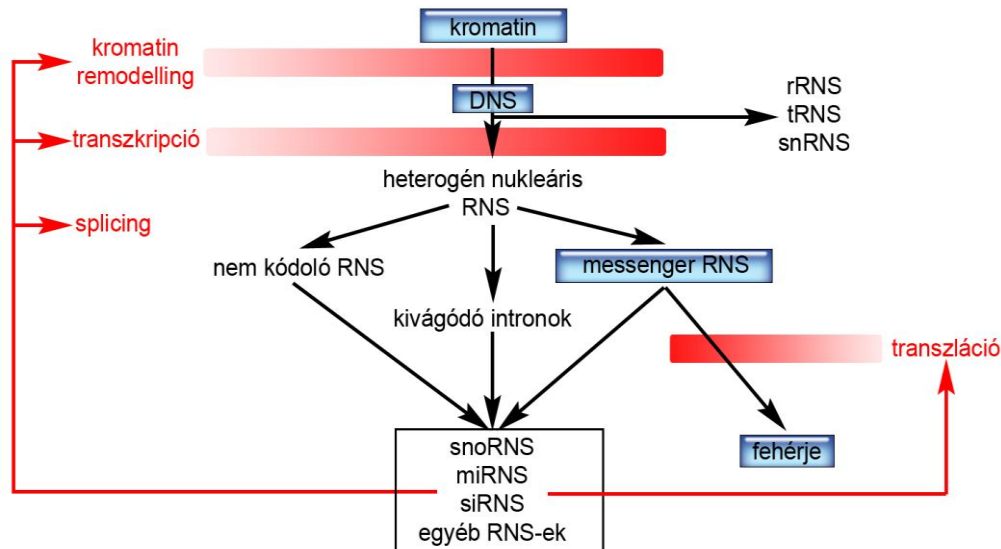
Az RNS típusok hagyományos felosztása szerint megkülönböztetünk tRNS-eket, amelyek a transzlációban kulcsszerepet töltenek be transzfer molekulaként, mRNS-eket, amelyek a genetikai információt közvetítik a génektől a riboszómaig, és a rRNS-eket, melyek a riboszómák funkcionális egységei, valamint a kis nukleáris RNS-eket (snRNS), amelyek a splicing folyamatában vesznek részt. Az utóbbi évek felfedezése alapján az RNS világ „háztartási” feladatokat ellátó családjait tovább bővíthetjük a „szabályozók” népes csoportjával, mint a mikroRNS-ek (miRNS).

piwiRNS: Feltehetően a csírasejtek fejlődéséhez elengedhetetlen 25-30nt hosszú, egyszálú RNS-ből képződő „kis RNS” típus. Nevüket a hozzájuk kapcsolódó az Argonauta fehérjecsaldba tartozó PIWI fehérjékről kapták.

rasiRNS: repeat-associated small interfering RNA, jellemzően az antiszenz szárlól keletkező, méretükben a miRNS-ektől és a piRNS-ektől némileg eltérő RNS-osztály. Újabban a piRNS-ek közé sorolják őket közös fehérjepartnereik (pl. PIWI) miatt.

siRNS: a miRNS-ekhez hasonló méretű, 20-25 nt hosszú, duplaszálú hosszabb RNS-ekből képződő „kis RNS”-osztály. Az ugráló genetikai elemek, a transzpozonok gátlásában és a vírusfertőzések során van jelentőségük.

[]: eukarióta kísérleti rendszerekben egy szintetikus siRNS segítségével mesterségesen, specifikusan csendesíthetők megcélzott mRNS-ek. Az siRNS teljesen tökéletesen illeszkedik a cél-mRNS-hez (általában annak kódoló régiójához), a kölcsönhatás eredménye így a cél-mRNS degradációja a Dicer enzim közreműködésével. A szögletes zárójel arra utal, hogy endogén termelésük az állatokra nem jellemző (annál inkább a növényekre!).



4. ábra. Az új RNS-családok beillesztésével a genetikai információ megnyilvánulási folyamatának újraértelmezésére lehet szükség (snRNS= small nuclear RNS, snoRNS= small nucleolar RNS)

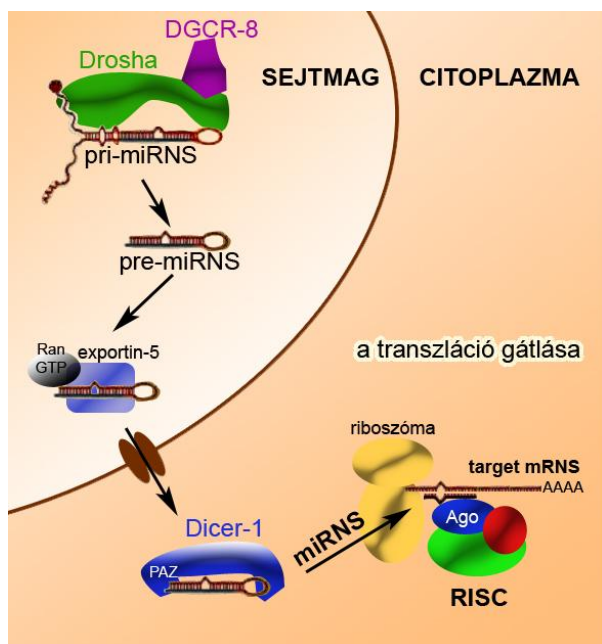
A kis nem-kódoló RNS-ek (small non-coding RNA), mint miRNS-ek is, uralják a tudományos irodalmat, egyre inkább fordulnak az ún. hosszú nem-kódoló RNS-ek kutatása felé (long non-coding RNAs, >200 nt). A hosszú nem-kódoló RNS-ek esetében is rendkívül széles spektrumon azonosítottak általuk befolyásolt folyamatokat. Hatásuk a kromatin (pl. a PCG2 kromatin remodelling komplex HOTAIR-indukálta rekrutálása), a transzkripció (pl. egy a ciklinD1 promoterével asszociált, stress-indukálta ncRNS represszív hatása) és poszttranszkripcionális szinten egyaránt (akár endogén siRNS-sé történő konverzióval, akár a splicing mintázatának befolyásolásával), sőt akár közvetlenül a fehérjékre hatva (pl. NRON szuppressálja az NFAT sejtmagi akkumulációját) is tetten érhetőek.

A fehérjét nem kódoló RNS-ek felfedezésével szükségessé vált olyan alapvető terminológiák átdefiniálása, mint például a gén fogalma. A klasszikus biológia centrális dogmája alapján a genetikai információ megnyilvánulása DNS→RNS→fehérje (→tulajdonság) irányba folyik, a gén a fehérjét kódoló régió (a kapcsolódó szabályozó elemekkel) szinonimájaként értelmezhető. A genomika korában a gén definíciót ki kell terjeszteni a „transzkripció egységre”, amely olyan komplett kromoszómális szegmens, amely funkcionális termék létrehozásáért felelős.

A mikroRNS-ek a hírvivő RNS-ek mennyiségét szabályozzák

A fehérjét nem kódoló RNS-ek között a mikroRNS-ek szerepe a leginkább ismert. A génexpresszió szabályozásában a megfelelő mRNS-ekhez kapcsolódva, képesek lehetnek a cél-mRNS-ek degradációját indukálni vagy azok translációját gátolni. Érett formájukban a miRNS-ek 21-25 nt hosszúságú egyszálú RNS-ek, amelyek processzálatlan előalakjai önálló transzkripcionális egységekként kódoltak a genom döntően intergenikus² szakaszain.

A jellegzetes érési folyamat végén álló érett miRNS kész a miRISC-komplexbe (miRNA-induced silencing complex) integrálódni, és azt a cél-mRNS 3'UTR régiójához³ vezetni, amely vagy a célmolekula degradációját vagy translációjának gátlását segíti elő. Azt, hogy a két kimenetel közül melyik következik be, döntően a két RNS molekula közötti szekvenciahasonlóság, a komplementaritás⁴ határozza meg.



A mikroRNS-ek transzkripciójában az RNS polimeráz II vesz részt, amely először egy hosszú prekuzort, a primer mikroRNS-t (pri-miRNS) hozza létre. A mikro RNS kialakulásában az első lépés ezen stem-loop (törzs-hurok) rész leválasztása egy heterodimer, a DROSHA (celluláris III-as típusú RNáz enzim) és kofaktora, a DGCR-8 által. Ez a hasítás egy kb. 60 nukleotid hosszú kétszálú köztesterméket eredményez, amit pre-miRNS-nek nevezünk. A következő lépés a pre-miRNS kijuttatása a sejtmagból, amelyben az Exportin-5 és kofaktora, a Ran GTP-kötő formája vesz részt.

Ezen prekursor miRNS-eken belül a 22 nt hosszúságú, érett mikroRNS tag a kb. 80 nt hosszú hajtókanyaros szekvencia egyik *karjának* része. Egy másik RNáz, a DICER és kofaktora a TRBP (Tar RNA binding protein) végzi a terminális hurok eltávolítását, és elősegíti a képződő RNS kettős szál RISC-hez való kapcsolódását. Ebben a komplexben a duplex szál szétválik és a 3'-5' komplementer szál degradálódik. Az így képződő érett mikroRNS-ek kötődése a cél mRNS-ekhez a miRISC segítségével történik meg.

A miRNS-ek által mediált endogén szabályozás bizonyítékait mind az állatok, mind a növények esetében azonosították. Tekintve, hogy a miRNS-ek számos egyedfejlődési folyamat

² intergenikus: A genom nagy részét kitevő, fehérjekódoló génektől mentes régió

³ 3'UTR: 3' untranslated region, az érett mRNS 3' végén található szakasz (a kódoló régió és a poly-A szakaszok között). Transzkripcióra igen, translációra nem kerül.

⁴ komplementaritás: szekvenciamegefelelőség, adeninnel szemben timin, citozinnal szemben guanin (és fordítva) bázisok állnak az egymással kapcsolódó (hibridizáló) szálakon.

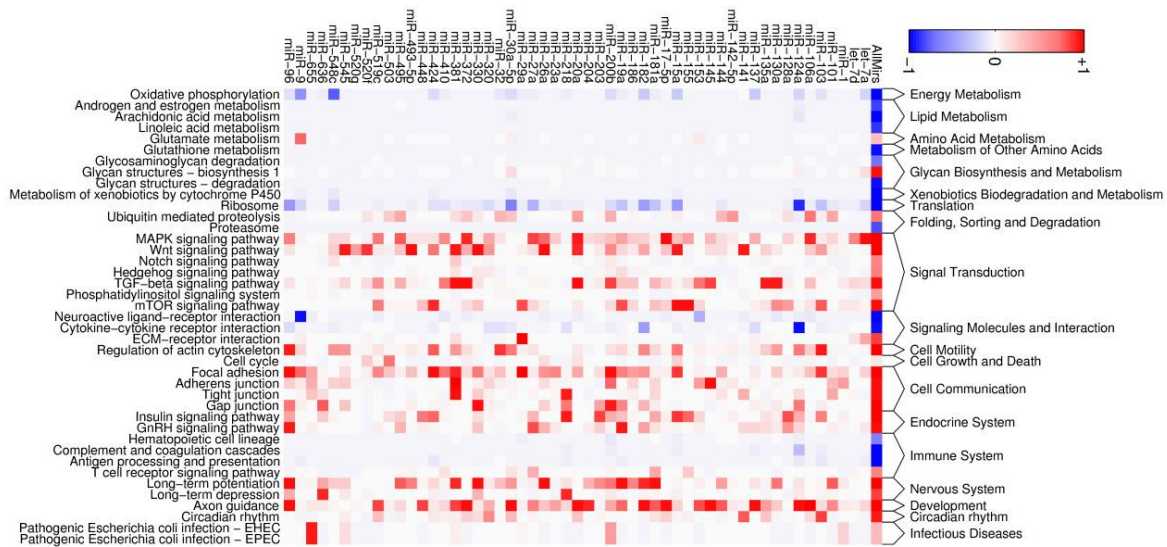
irányításában működnek közre, mint az egyes sejtsorsok molekuláris időzítése, az összejt-állapot fenntartása, egyes szervek morfogenezise, valószínűsítik, hogy a szövet- vagy sejtspecifikus expressziós mintázatok kialakításában, annak finom beállításában állhat a fejlődés során a jelentőségük. Nem korlátozódik azonban a miRNS-ek eddig megismert szerepe a fejlődési folyamatokban való közreműködésre, és az a lehetőség, hogy lehetséges számos fehérje gyors és szinkronizált módon történő koncentrációjának beállítása, a sejt válaszaiban és környezethez történő alkalmazkodás során bizonyul felhasználhatónak. Mindezeket túl a miRNS-ek a sejtek közötti kommunikáció eszközei is lehetnek. Az exoszómáisan szállított RNS (“exosomal shuttled RNA”) részeként azonosíthatók a hízósejtek által szekretált exoszómákban. A sejt-sejt kommunikáció ezen újonnan felismert formáján keresztül a donorként közreműködő sejt direkt módon képes lehet a célsejtek géneexpresszióját direkt módon poszttranszkripcionális szinten befolyásolni.

mikroRNS target predikciók: „in vitro” helyett „in silico”

Mivel a miRNS és a cél-mRNS-ek közötti komplementer szakasz viszonylag rövid, sőt a kapcsolat kifejezetten a tökéletlen illeszkedést követeli meg, egy miRNS több száz célponttal is rendelkezhet. A miRNS-közvetítette szabályozás (állatokban) főleg fehérjeszinten valósul meg, és ennek vizsgálatára nagy áteresztőképességű fehérjemeghatározási módszer lenne szükséges. Ennek a problémának a feloldására számos számítási algoritmust fejlesztettek ki, amelyek többé-kevésbé megbízhatóan határozzák meg a cél-mRNS-ek körét. Ezek az algoritmusok a miRNS-ek kapcsolódó szekvenciájának magjait („miRNA seed”) képesek megtalálni a 3'UTR szekvenciák gyűjteményében. A keresés során a különböző algoritmusok a következő paramétereket veszik figyelembe: a szekvenciakomplementaritás, a heteroduplex stabilitását leíró szabadenergia, és az evolúciós konzerváltság. Annak ellenére, hogy a legelterjedtebb algoritmusok alapvetően ezeket a paramétereket veszik számításba, mégis minimális eltérések is az eredménylistákat nagymértékben eltolhatják. Az egyes target predikciós programok által generált halmazok átfedése jellemzően alacsony mértékű, és a sokszor több száz jóslott célpont közül csak csekély hányaduk igazolt kísérletesen. Jóllehet az RNS molekulák közötti interakciók, pusztán a szekvencia-adatok felhasználásával, sokkal megbízhatóbban modellezhetők, mint a transzkripció faktorok-kötőhelyek azonosításban szükséges fehérje-DNS kapcsolatok esetén, mégis a nagy áteresztőképességű proteomikai kísérletek tanúsága szerint teljesítményük éppen, hogy csak elfogadható a prediktált miRNS-kötőhelyek 2/3-a fals-pozitív. A miRNS-mRNS kapcsolatok pontosabb modellezésével, a kötőhelyek megbízhatóbb predikciója a fals-pozitív arány csökkentése mellett a bioinformatika fontos nyitott kérdései közé tartozik.

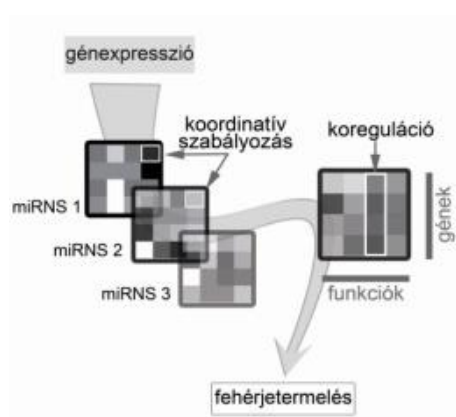
A kísérletes megfigyelések és a bioinformatikai predikciók mára kezdik körvonalazni a miRNS-ek által mediált szabályozás alapvető elveit. Érdekes módon az ubikviter, alapvető metabolikus folyamatok elemeit kódoló gének viszonylag alul-, míg a sejtsztódás, sejthalál, ontogenezis (különösen az idegrendszer fejlődés, mint az axonok növekedését irányító útvonallal), transzkripció szabályozás, és sejtek közötti kommunikáció viszonylag felülreprezentáltak a miRNS-

közvetített szabályozást tekintve. Feltehetően a miRNS-ek az alapvető regulációt kiegészítő közreműködése azokban az esetekben lehet fontos, amikor a sejtben az egyensúly finom megtalálása szükséges az adott külső és belső szignálokra.



6. ábra. A mikroRNS-ek által megcélzott folyamatok köre meghatározható, ha megvizsgáljuk, hogy a prediktált célgének jellemzően milyen életfolyamatban játszanak szerepet. A sorokban a tesztelt életfolyamatok (útvonalak, funkciók), az oszlopokban az egyes mikroRNS-ek láthatóak. A piros szín az adott mikroRNS kötőhelyeinek viszonylagos felül-, a kék az alulreprezentáltságot mutatja a hozzátartozó életfolyamathoz tartozó mRNS-ek körében. *Dimos Gaidatzis et al. Inference of miRNA targets using evolutionary conservation and pathway analysis. BMC Bioinformatics 2007, 8:69*

Egy adott miRNS kötőhelyei több esetben is több, funkcionálisan rokon transzkriptumban megtalálható és ez a koregulációs elv lehetővé teszi egész útvonalak vagy több együttműködő útvonal elemeinek harmonizációját. Bizonyítékok támasztják alá a több miRNS koordinatív együttműködését, amikor különböző miRNS-ek független kötőhellyel rendelkeznek ugyanabban a 3'UTR-ben, szinergista módon a cél-mRNS-ből keletkező különböző termelési szinteket határoznak meg. Ezzel

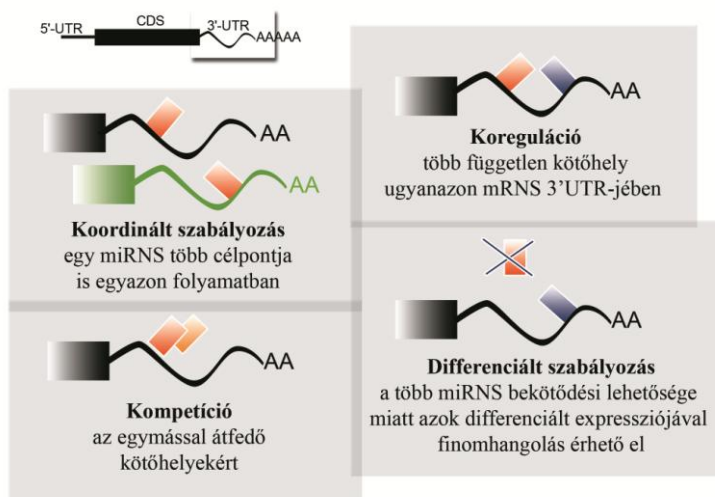


összhangban, ha egy cél-mRNS több kötőhelyet tartalmaz egy adott mikroRNS számára, azok száma jól korrelál a represszió mértékével.

A mikroRNS-ek által képzett célmolekula-mintázat, mint a génexpresszió szabályozásának új rétege. A mikroRNS-ek együttese által közvetített szabályozási héj sematikus ábrázolása. Az egyes mikroRNS-ek cél-mRNS-ek sejtfunkciók szerint rendezett mátrixai egymásra rakódnak, ezáltal egy eredő mintázatot alakítanak ki. A vizsgált sejtben expresszálódó miRNS-ek együttesen tehát egy szabályozási héjat képezve, finomhangolhatják a transzkripcióból eredő információt. Az engedékeny kapcsolódási szabályszerűségeknek köszönhetően a legtöbb mRNS számos miRNS kötőhelyet is tartalmazhat. Egy adott mRNS-t tehát több miRNS is szabályozhat (ebben az esetben, koordinatív módon együttműködnek). Ugyanakkor egy vagy több

miRNS regulálhat adott funkcióhoz tartozó több mRNS-t is (koreguláció). Mivel a miRNS-közvetített szabályozás negatív, gátló jellegű, nemcsak a gátolt funkciónak lehet jelentősége, hanem a nem-targetált, a miRNS-ek hatósugarán kívül eső géncsoportok esetében is, amelyek éppen ezáltal tudnak érvényre jutni.

Továbbá, a 3'UTR-ben fekvő, egymással átfedést mutató kötőhelyek versenyhelyzetet teremtenek az egyidőben expresszálandó, bekötődni képes miRNS-ek között. Végül, ha több független kötőhelyet találunk egy adott 3'UTR-ben, az nem jelenti automatikusan az összes miRNS együttes jelenlétét. Bizonyos körülmények között a különböző adott kombinációban megjelenő miRNS-ekkel tehát lehetőség van a célgén kifejeződéseinek adott szintre történő beállítására (mintázat-differenciált szabályozás; pattern differential regulation). A számos említett jelenség megfigyelése alapot szolgáltat a kísérleti adatok megfelelő értelmezéséhez és megengedi a génexpresszió ezen új szintjének integrációját az eddigi ismeretek közé.



8. ábra. A mikroRNS-ek interakciós lehetőségei a cél-mRNS-ben található kötőhelyek szerint.

A kísérleti módszerek folyamatos fejlődése (mint pl. az újgenerációs szekvenálás⁵, transzgenikus kísérleti állatok és nagy áteresztőképességű proteomika) a bioinformatikai eszköztár kibővülésével nyithatják meg a lehetőséget a genom nem-kódoló szakaszainak belefoglalását a szabályozási hálózatok felderített ismerethalmazába. Jóllehet a folyamatosan bővülő számú nem-kódoló RNS mellett, amelyre éppen a miRNS-ek osztálya szolgáltatja a legprominensebb példát, minden bizonnyal számíthatunk újak, akár új családok felfedezésére

⁵ újgenerációs szekvenálás (NGS=next generation sequencing): korunk molekuláris biológiáját forradalmasító eljárás. Igen nagy mennyiségű szekvencia leolvasása jelentősen kisebb laborköltséggel jár, de annál intenzívebb informatikai feldolgozást igényel. Az új megközelítés lényege, hogy a DNS (vagy RNS-ről készült cDNS másolat) kis darabjainak bázissorrendjét „válogatás nélkül” leolvassuk (párhuzamosan több milliót egyszerre!), majd a referenciaszekvenciához hasonlítva határozzuk meg, melyik (ismert funkciójú) genomi szakaszhoz tartozik a leolvasott szakasz (annotáció).