

A cél alternatív splicing (AS) által módosított kötőhelyek feltárása volt. Ebből a célból két fehérje-fehérje kölcsönhatásokat tartalmazó adatbázist vizsgáltam eddig: az IntAct és a String adatbázisokat. Az IntAct mintegy 43 ezer kölcsönhatást (kh) tartalmaz, melyeknek nagyrésze páros kh, így céljainkra alkalmas volt. Ezekből kb. 1300 kh tartalmazott alternatív izoformát, melyek közül kb 900 volt minor izoforma. Ezeknek a fő izoformákkal és azok kötőpartnereinek az összehasonlításával betekintést nyerhetünk az AS által létrehozott változásokról. A legkézenfekvőbb azoknak a variánsoknak a vizsgálata, amelyeknek negatív kontrollja is megtalálható volt az adatbázisban (amikor kifejezetten páros kh hiányát rögzítették a vizsgálatok). Ilyen esemény, ill. izoformapár csak 8 volt az IntActban, melyeket részletes analízisnek alávetve azt találtuk, hogy hat esetben az AS által eliminált részben található volt olyan kötőhely, melyek hiánya okozhatta a minor izoformák eltérő kötési tulajdonságait. Egy másik megközelítés szerint a STRING adatbázisban lévő kh-okat vizsgáltuk meg, melyek elképzeléseink szerint az egyes fajokban lévő ortológok közötti különbségek alapján nyújt felvilágosítást a fehérje kötőhelyekről. Mint azonban kiderült, [1] a STRING inkomplett, a legtöbb páros kh-t az élesztő fehérjékről tartalmazza, nyilvánvalóan a kísérleti módszerek preferenciáiból adódóan (élesztő két-hibrid módszer) és az egyes fajokat (ember, egér vagy *Drosophila* között sem) nem lehet összevetni ilyen szempontból.

E sajnálatosan szűk adatmennyiség mellett a másik probléma az alternatíván összeszerkesztett fehérjék inkonzisztens viselkedése volt: azt találtuk a rendelkezésre álló Swissprot fehérjék vizsgálata során (melyeket az ANCHOR kötőhely predikáló szerver segítségével vizsgáltunk), hogy az alternatíván splice-olt régiókban nem feltétlenül nagyobb a kötőhelyek aránya, mint a nem AS régiókban ugyanezekben a fehérjékben. Ez utóbbi összefüggést találtuk a 2. és a 3. izoformák vizsgálata során, míg a magasabb számosságú izoformákban (tehát 4-es, 5-ös, etc) viszont az AS régiókban ténylegesen magasabb volt a kötőhelyek aránya, mint a konstitutív (tehát AS-t nem szenvedő) régiókban.

E nehézségek kiküszöbölése céljából további koncepciókat vezettünk be/adtunk a projekthez. Az egyik a kötőhelyek feltárása mellett az alternatív splicingnál található rendezetlenség és hidrofób felszín vizsgálata volt: az utóbbi meghatározó szerepet játszik az alternatíván szerkesztődő fehérjék túlélési esélyeit illetően: a túlzott mennyiségű hidrofób felszín kialakulása jelentősen lecsökkenti a fehérjék túlélési valószínűségét. Erről a *Nucleic Acids Research* c. folyóiratban publikáltunk egy elméleti cikket [2]. Itt felhasználtuk az ANCHOR kötőhelyekről szerzett negatív tapasztalatunkat is és ezt a paramétert kiiktattuk a figyelembe vett paraméterek, ill. változók sorából.

Egy másik koncepció a rendezetlenség, a fehérje-fehérje kölcsönhatások és az organizmus ill. a különböző szövetek komplexitása közötti összefüggések feltárása volt. Azt találtuk, hogy az alternatív splicing, a fehérje-kötőhelyek száma és a rendezetlenség is növekszik az organizmus komplexitásának növekedésével, bár az utóbbi csak a prokarióták és eukarióták közötti ugrás során volt számottevő. Ezzel szemben viszont a proteóm mérete egyértelmű korrelációt mutat a komplexitás mértékével, nagyrészt cáfolva ezzel az ú.n. G-paradoxont, mely azt posztulálta, hogy ilyen összefüggés nincs. Az erről szóló cikket a *Genome Biology*-ban közzeltük [1].

Egy harmadik irányvonal a nemi kromoszómákon kódolt fehérjék rendezetlenségét és a fehérje-fehérje kölcsönhatások eloszlását vizsgálta. Megállapítottuk, hogy az X, Y kromoszómákon kódolt fehérjékben mindig nagyobb a rendezetlenség, mint az autoszómákon kódolt fehérjékben, nemcsak emberben, hanem az összes vizsgált emlős fajban, madarakban ezzel szemben nincs ilyenfajta egyenlőtlenség. Ez a nagyobb mértékű rendezetlenség érdekes módon nem társult nagyobb mennyiségű fehérje-fehérje kölcsönhatással, éppen ellenkezőleg: az egy fehérjére jutó kölcsönhatások száma

az X, és különösen az Y kromoszómán kisebb volt, mint az autoszómákon kódolt fehérjék esetében. Ez utóbbi jelenség minden bizonnyal a fehérjék funkcionális specializációja miatt alakult ki és lényegesen több intra-kromoszomális kölcsönhatást találtunk az X és Y kromoszómán kódolt fehérjék esetében, mint az autoszómákon kódolt fehérjékre. Érdekes módon, a madarakban vagy egyéb gerincesekben nem volt ilyen összefüggés, míg Drosophilában szintén megtalálható volt. Az erről szóló cikket a Molecular Biosystems c. folyóiratban közöltük [3].

Utólag szeretném megjegyezni, hogy ez utóbbi publikáció esetében [3] az eredeti cikkben valamilyen okból kimaradt az Acknowledgement, és ily módon az OTKA támogatottság is, de ezt a repozitóriumban (<http://real.mtak.hu/>) elhelyezett változatban korigáltam, így az onnan letölthető file végén már korrekt módon fel van tüntetve az OTKA pályázataink száma (NK71582 és PD76286). Kértem a folyóirat szerkesztőit is, hogy lehetőleg javítsák ki ezt a kihagyást az elektronikus változatban.

Irodalomjegyzék

1. Schad E, Tompa P, Hegyi H The relationship between proteome size, structural disorder and organism complexity. *Genome Biol* 12: R120.
2. Hegyi H, Kalmar L, Horvath T, Tompa P Verification of alternative splicing variants based on domain integrity, truncation length and intrinsic protein disorder. *Nucleic Acids Res* 39: 1208-1219.
3. Hegyi H, Tompa P Increased structural disorder of proteins encoded on human sex chromosomes. *Mol Biosyst* 8: 229-236.