

Válaszok Dr. Szüts Dávid bírálataira

Nagyon szépen köszönöm Dr. Szüts Dávidnak az MTA doktori dolgozatomra adott pozitív és támogató bírálatát és a bírálatra fordított energiát.

A feltett kérdésekre az alábbiakban válaszolok:

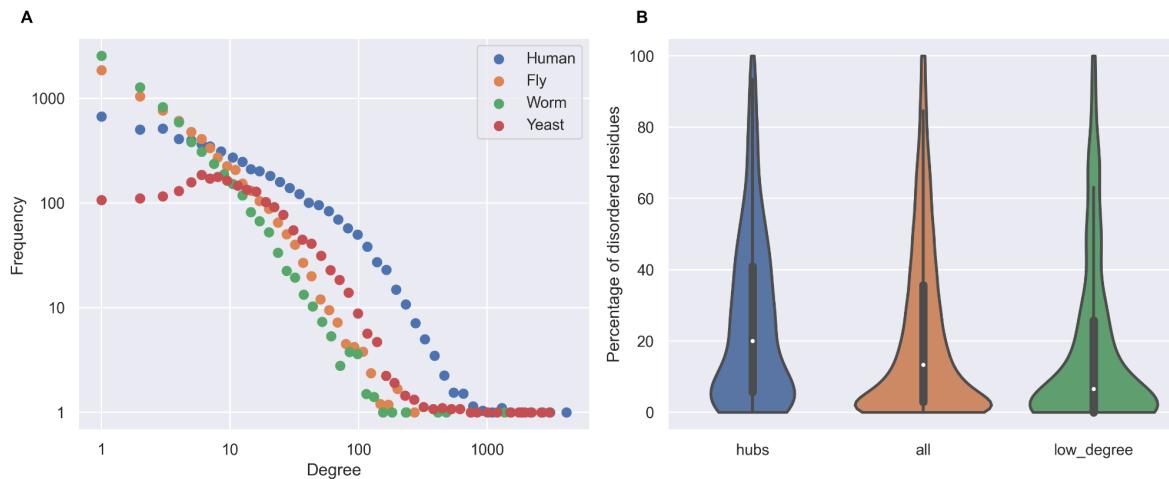
1. A dolgozat főleg 'rendezetlen fehérjék' néven említi a kutatások tárgyát, holott gyakran egy fehérjén belül rendezett és rendezetlen régiók is megtalálhatók. Mégis, például a 4. ábrán is külön szerepelnek a globuláris és a rendezetlen fehérjék, akár 800 aminosav hosszúságig. Kérdésem, hogy mi a tipikus doménszerkezete a rendezetlen fehérjéknek? Vannak-e tisztán rendezetlen fehérjék, hány és mekkora rendezetlen régió jellemző egy-egy fehérjére, lehet-e a fehérjéket eszerint csoportosítani, osztályozni?

A dolgozatomban a 4. ábrán a PDB szerkezetekre illetve az akkor ismert rendezetlen régiókra vonatkozó adatokat ábrázoltam, ezek tehát legtöbbször nem teljes fehérjék. Valóban, gyakran egy fehérjén belül rendezett és rendezetlen régiók is találhatóak. A legalább 30 aminosanyi rendezetlen régiót tartalmazó fehérjéket gyakran nevezik az irodalomban egyszerűen rendezetlen fehérjének (IDP - intrinsically disordered protein), és én is ezt a konvenciót követtem. Vannak tisztán rendezetlen fehérjék is, ezek azonban viszonylag ritkák. Általánosságban nem igazán beszélhetünk tipikus doménszerkezetéről, ez már a rendezetlen fehérjék jelentős számából is adódik. Például a humán fehérjék körülbelül 40-50%-a tartalmaz legalább egy 30 aminosavnyi rendezetlen régiót, ezek hossza pedig nagyon széles skálán mozog, elérheti a több ezer aminosavat is. A domének és a rendezetlenség kapcsolatát vizsgálta részletesen Pentony and Jones 2010-ben [Pentony, M.M. and Jones, D.T. (2010) *Proteins*, 78, 212]. Ők a Disopred2 módszer vizsgálatával arra jutottak, hogy a rendezetlen régiók száma egy fehérjén belül 1 és 25 között változik, és a legtöbb fehérje kevesebb mint 5 rendezetlen régiót tartalmaz. Az általános tendenciákon túl, tehát például, hogy a prokarióta fehérjékben jóval kevesebb a rendezetlenség mint a magasabb rendű eukariótákban, vagy hogy a szignalizációs és szabályozási folyamatokban is gyakoribb a rendezetlenség, másról nem igazán tudunk beszélni.

2. A fehérje interakciós hálózatok csomóponti (hub) fehérjéi mennyivel tartalmaznak több rendezetlen régiót az átlagnál? Csak a rendezetlen régiókban több a kölcsönhatás, vagy a rendezett doménjeikben is, az utóbbiakhoz viszonyítva is dúsulnak-e a kölcsönhatások a rendezetlen régiókban? A releváns 2006-os közlemény óta keletkezett adatok fényében is fennáll az interakciós hálózatok skálafüggetlen volta, és a csomóponti fehérjék kiemelkedő rendezetlensége?

A rendezetlen régiók számát nem szoktuk vizsgálni, mivel ez nagyon érzékeny a predikciók zajosságára, nagyban függ a fehérjék hosszától és egy hosszabb rendezetlen szakasz több funkcionális egységet is tartalmazhat. Azt sajnos általánosságban nem tudjuk megmondani, hogy mely régiók felelősek a kölcsönhatásokért. Néhány jól ismert fehérje esetén, például p53, egyértelmű, hogy a rendezetlen régiók jóval több kölcsönhatás kialakításában vesznek részt [Gibson, T.J. (2009) *Trends Biochem. Sci.*, **34**, 471]. Az interakciós hálózatokra vonatkozó legújabb adatok fényében, szigorú értelemben véve eltérést látunk a skálafüggetlen viselkedéstől (1A. ábra). Azonban a központi fehérjék megléte, ami leginkább

megkülönbözteti a skálafüggetlen hálózatokat a véletlenszerűektől, továbbra is fennáll. A csomóponti fehérjék magasabb rendezetlensége szintén továbbra is megfigyelhető (1B. ábra).



1. Ábra (A): A kölcsönhatási hálózatban az adott fokszámmal rendelkező fehérjék száma 4 faj (Humán, muslica, fonalféreg, élesztő) esetén. (B) A rendezetlen aminosavak arányának megoszlása a humán proteom kölcsönhatási hálózatában a hub fehérjékre (a partnerek számát tekintve a legfelső 1000), az összes fehérjére, és az alacsony fokszámú fehérjékre (a partnerek számát tekintve a legalsó 1000).

A számításokhoz a BioGrid adatbázis [Stark,C., *et al.* (2006). *Nucleic Acids Res.*, 34, D535] használtam (version 4.4.228). A rendezetlen aminosavak arányát az IUPred módszerünkkel határoztam meg.

3. Milyen, hányféle funkciója ismert a rendezetlen fehérjéknek/régióknak? Mennyiben jósolható meg a szekvencia (beleértve a méret) alapján a rendezetlen régiók funkciója?

A rendezetlen fehérjék funkcióinak Keith Dunker és munkatársai illetve Tompa Péter vezetett be klasszifikációt, melyek 28 illetve 6 főbb bővebb kategóriát tartalmaztak [Dunker,A.K., *et al.* (2002). *Biochemistry*, 41, 6573.; Tompa,P. (2005) *FEBS Lett.*, 579, 3346]. Jelenleg a DisProt adatbázis annotációjakor a Tompa Péter által javasolt funkcionális kategóriák továbbfejlesztett változatát használjuk. A főbb kategóriák: entrópikus láncok, illetve különböző molekuláris felismeréshez kapcsolódó funkciók, pl. összeszerelő, begyűjtő, effektor, bemutatási hely és chaperon funkciók. Ezeknek a kategóriáknak egy része direkt módon áttérképezhető a Gene Ontology (GO) funkcióira [Gene Ontology Consortium, Aleksander,S.A., *et al.* (2023) *Genetics*, 224]. Fontos ugyanakkor figyelembe venni, hogy a GO kategóriák alapvetően a teljes fehérje funkcióit írják le. Egyes, a rendezetlen régiókhoz szorosan kapcsolódó funkcionális kategóriák, például entrópikus funkciók, illetve az autoregulációs és módosítási helyek, nem értelmezhetők fehérje szinten. Ezért ezeket egy

külön, a rendezetlen fehérjékre használt, IDPO ontológia írja le [Aspromonte, M.C., *et al.* (2023) *Nucleic Acids Res.*, 10.1093/nar/gkad928].

A rendezetlen fehérjék funkciójának predikciójára viszonylag kevés módszer áll rendelkezésre, a rendezetlen kötőhely predikciós módszereken túl vannak módszerek rendezetlen DNS és RNS kötő fehérje, illetve lipid kötő régiók jóslására, illetve a linkerek predikciójára [Kurgan, L., *et al.* (2023). *Nat. Protoc.*, 18, 3157]. Összességében az ismert példák viszonylag kis száma erősen limitálja ezekben a módszereknek a teljesítőképeségét.

Van egy olyan törekvés, hogy a rendezetlen fehérjék funkciójának predikciója beépüljön a CAFA kihívásba (Critical Assessment of protein Function Annotation algorithms) [Zhou, N., *et al.* (2019). *Genome Biol.*, 20, 244.], aminek a célja a fehérje funkciók számítógépes annotációjának értékelése és javítása. A közelmúltban a DisProt-ban lévő rendezetlen fehérjék funkciójának annotációját felhasználták a CAFA egyik részfeladatában, azonban az eredmények eddig még nem kerültek közzétételre.

4. Létezik-e szekvenenciaalapú megközelítés az olyan fehérjék felismerésére, amelyekben a rendezetlen régió fázisszeparáció célját szolgálja, és nem pl. Közvetlen fehérjekölcsönhatást?

Igen, több ilyen megközelítés is létezik. Egy nemrégiben megjelent publikációban [Kuechler, E.R., *et al.* (2023) *Biomolecules*, **13.**] 11 módszert hasonlítottak össze, míg egy másik publikációban [Vendruscolo, M. and Fuxreiter, M. (2023). *Curr. Opin. Chem. Biol.*, 75, 102317.] majdnem 20 különböző eljárását sorolnak fel.

5. Az evolúció során megjelenő rendezetlen fehérjerégiók kialakulásában mekkora szerepet játszanak rögzült intronretenciós események?

Ez egy nagyon érdekes kérdés, de mi ezt nem vizsgáltuk. Mivel a rendezetlen régiókat tartalmazó fehérjékben nagyon gyakori az alternatív splicing [Romero, P.R., *et al.* (2006). *PNAS*, 103, 8390.], feltételezhető, hogy az intronretenciós fontos szerepet játszott a rendezetlen fehérjék kialakulásában.

6. Megfigyeléseik szerint a rendezett régiók gyakrabban mutálódnak a rák kialakulásában érintett fehérjékben. Feltehetőleg ez csak missense mutációkra igaz, nonsense mutációkra nem?

A TCGA rák genom projektek adatai alapján a nonsense mutációk eloszlása nem mutat szignifikáns eltérést rendezett és rendezetlen részek között.

7. A módszerek részben említik, hogy a predikciós módszereket vagy független adatokon tesztelik, vagy több program eredményének konszenzusához viszonyítva. Az utóbbi eset mit jelent? Nem fordulhat-e elő, hogy esetlegesen hasonló megközelítést alkalmazó predikciós programok hasonló módon tévednek?

Egészen pontosan azt írtam: "A predikciós módszerek alkalmazásával kapott eredmények megerősítése érdekében gyakran alkalmaztam több különböző programot, vagy ezek

konszenzusát.” Erre egy példa, hogy annak vizsgálatakor, hogy egy mutáció rendezetlen részre esik-e, nemcsak az IUPred módszert vettem figyelembe, hanem esetenként más, eltérő megközelítésen alapuló módszereket is. A predikciós módszerek tesztelésénél nem gondolom jó megoldásnak több program eredményének konszenzusához viszonyítani, természetesen előfordulhat, hogy a predikciós programok hasonló módon tévednek.

Ismételten köszönöm Dr. Szüts Dávidnak a bírálatot és kérdéseket. Kérem fenti válaszaim szíves elfogadását, és az MTA doktora cím odaítélésének támogatását!

Budapest, 2023. December 5.

Dr. Dosztányi Zsuzsanna