Válasz DOSZTÁNYI ZSUZSANNÁNAK

„Átfedő modulok molekuláris biológiai kölcsönhatási hálózatokban"

című MTA doktori értekezésem bírálatára

Köszönöm Dosztányi Zsuzsannának a dolgozat alapos átolvasását, és a dolgozat kapcsán írt pozitív konklúziót. A bírálatban feltett kérdésekre a következő válaszokat adom.

(1) „A jelenlegi biológiai hálózatokban nagyon sok fals pozitív, tehát hibás, illetve fals negatív, tehát hiányzó adatot találunk. Hogyan befolyásolja a kapcsolatok hiánya, illetve a hamis kapcsolatok megléte a hálózatok általános jellemzőit, illetve a modulok megjelenését és eloszlását?"

Referenciaként érdemes először megvizsgálni két hálózat modellt. Az **Erdős-Rényi** (ER) hálózatban az összes tulajdonság két paramétertől függ: az élek számától (E) és a csúcsok számától (N). Ezek hányadosának kétszerese a <k> átlagos fokszám. Ha az ER hálózatban $N \gg \langle k \rangle \gg 1$, akkor véletlenszerű él eltávolítás vagy hozzáadás után a hálózat fő jellemzői mennyiségileg változnak, de a típusuk nem változik. Az **általános jellemzők** közül például a fokszámeloszlás marad <k> paraméterű Poisson és a klaszterezettség marad <k>/N. A **skálafüggetlen** (SF) hálózatban az élek eltávolítása vagy hozzáadása a csúcsok véletlenszerű eltávolításához/hozzáadásához hasonló típusú, de annál kisebb hatást jelent. Nagy SF hálózatban a kevés random él eltávolítása vagy hozzáadása esetén is megmarad a skálafüggetlen fokszámeloszlás és az alacsony klaszterezettség. Az ER és SF hálózat között hasonlóság, hogy jelentős **modulok** egyik hálózatban sem találhatóak. Az ER és az SF hálózat között eltérés, hogy nagy számú él (vagy csúcs) eltávolításakor az SF hálózatban jóval később „esik szét" az óriás komponens, mint az ER hálózatban [1, 2].

Valós hálózatok esetén a **korrelálatlan fals pozitív** kapcsolatok (tehát a tényleges hálózathoz véletlenszerűen hozzáadott élek) egy alacsony átlagos fokszámú ER hálózatot „helyeznek rá" (szuperponálnak) a tényleges hálózatra. Emiatt a fals pozitívok esetén az **általános jellemzők** az ER hálózat felé tolódnak, például a legkisebb fokszámú csúcsok fokszáma növekszik és a klaszterezettségi együttható csökken. Korrelálatan fals pozitív kapcsolatok esetén a **modulok** (és emiatt a méret eloszlásuk) nem változik jelentősen. Valós hálózatokban a random **fals negatív** kapcsolatok egy olyan zajt („hibát") jelentenek, amelyek a fokszámeloszlást kevéssé befolyásolják, továbbá a klaszterzettséget és a modul méreteket enyhén csökkentik.

Érdemes megjegyezni, hogy a valós (főként szociális és molekuláris biológiai) hálózatok fals pozitív és fals negatív kapcsolatainak azonosítására létezik már számos, részletesen **kidolgozott numerikus módszer** [3, 4, 5]. Elsőként tekintsük a fals pozitív élek eltávolítását. A valós hálózatok fals pozitív kapcsolatainak egyik fő oka az, hogy számos mérési módszer nem csak közvetlen kapcsolatokat mér, hanem közvetett kapcsolatokat is, és így az egyes mérési módszerek eredményeinek összegzése után a közvetlen és a közvetett kapcsolatokat már nem lehet szétválasztani. Erre a biológiában egy jól ismert példa a koexpresszió (mRNA coexpression). Ha az f(x) és g(x) függvény korrelációja magas, és ugyanez teljesül a g(x) és h(x) függvényre, akkor az f(x) és a h(x) függvények korrelációja is jelentős lesz. Ennek alapján ha két génnek sok kísérleti körülményre számított (mRNS) expressziós vektora között a korreláció abszolút értéke magas, akkor ennek az oka a közvetlen kapcsolaton (például fizikai kölcsönhatáson) túl lehet a közvetett kapcsolat is.

Megfelelő mennyiségű adat esetén a **közvetett kapcsolatok eltávolításának** egyik lehetséges módszere az, hogy a mért kapcsolatokat tartalmazó G szomszédsági mátrix-ot (a közvetlen és közvetett kapcsolatok összegét) összehasonlítjuk a közvetlen kapcsolatokat tartalmazó – de kezdetben ismeretlen – S mátrix-szal, és az így S-re felírt implicit egyenletből kiszámítjuk az S mátrix elemeit (a közvetlen kapcsolatokat) [4]. Ezután hiányzó (az az **fals negatív**) élek azonosításának az egyik módja az, hogy a korrelációs tulajdonságok (például a szomszédok kapcsolat listáinak korrelációja, a „topological correlation") átlagos értékétől való nagy helyi eltéréseinek a valószínűsége kicsi.

(2) „A fehérje kölcsönhatási hálózatok esetén meg lehet különböztetni asszociációt illetve direkt kölcsönhatást. Hogyan változnak a modulok, ha csak direkt kölcsönhatásokat vizsgálunk? Így is leírhatóak-e az azonos funkcionális csoportban található fehérjék, mint átfedő klikkek?"

Röviden összefoglalva: igen, a dolgozatban kapottaknál kisebb, egymással átfedő modulokat kapunk, amik funkcionális csoportokat jelölnek. A pontos név „átfedő klikkek" helyett „átfedő k-klikk perkolációs klaszterek".

A közvetett kölcsönhatások törlésével a kölcsönhatási hálózat éleinek száma erősen csökken, ezzel szemben a csúcsok (a legalább 1 kölcsönhatással rendelkező fehérjék) száma kevésbé csökken. Emiatt **kisebb modulok** várhatóak. A törlés után megmaradó közvetlen (direkt) fehérje-fehérje kölcsönhatások jelentős részét fehérje komplex-ek alapján ismerjük. Egy fehérje komplex a fehérjék egy olyan csoportja, amelynek minden tagja azonos időben összekapcsolódik egy egységbe. Ezzel szemben a funkcionális modul/csoport egy olyan fehérje csoport, amelynek a tagjai egyszerre nem kapcsolódnak össze mindannyian. A dolgozatban felhasznált [T2] publikációnk kísérleti adatokra [6] hivatkozva írja, hogy élesztőben (*S. cerevisiae*) a mért fehérje komplex-eket 2750 fehérje alkotja, és ugyanezeknek a komplex-eknek a méret (fehérje szám) összege 8932. Tehát egy fehérje átlagosan 3-nál több komplex-ben vesz részt. Ezért a k-klikk perkolációs módszerrel kiszámított hálózati modulok között is **jelentős átfedések** várhatóak.

(3) „A dolgozatban négyfajta biológiai hálózatot elemzett. Érdemes lett volna ezeknek a hálózatoknak az összehasonlítását elvégezni. Vannak-e jelentős eltérések például a hálózatok általános paramétereinek, illetve a modulok méretének és eloszlásának tekintetében? Tudná-e értelmezni az eltérések biológiai következményeit?"

A dolgozatban szereplő 4 molekuláris biológiai hálózat: fehérje-fehérje kölcsönhatások és asszociációk (PPI), transzkripció szabályozás (TR), transzláció szabályozás (TL), és a jelátvitel (ST). A 4 hálózat között **két fő típus**t találhatunk. A PPI hálózat **irányítatlan**, a többi három hálózat **irányított**. Az irányított jelátviteli hálózat jelentős részben a PPI hálózat irányítatlan kapcsolatait tartalmazza jelátviteli irányokkal ellátva.

Biológiai szempontból az irányítottság legfontosabb következménye az információ terjedésének egyértelmű iránya, például a jelátviteli hálózatban az aktivált állapot továbbadásával. Az irányítottsággal kapcsolatos egyik legfontoosabb tulajdonság az, hogy a TR és az ST irányított hálózatokban gyakoriak a **motívumok** [7, 8] (az irányított fokszámokat megtartó élkeverés eredményéhez képest szignifikánsan felülreprezentált kis irányított részgráfok), amelyeknek a dinamikai viselkedése is ismert. Például az **FFL** (pontosabban: C1-AND-feed-forward-loop) motívum képes két bemenő jelet összesíteni, vagy 1 bemenő jelben gyengíteni a gyors változásokat (aluláteresztő szűrő). Biológiai szempontból az FFL mellett szintén központi szerepű a visszacsatolás, amely a TR és az ST hálózatban sok helyen megtalálható. Ezen felül ismert kombinált transzkripciós-transzlációs „feedback" is [9]. A pozitív **visszacsatolás** az adott jelet telítődésig erősíti, és ezt a telítő-

dést (szaturációt) általában egy negatív visszacsatolás valósítja meg. Biológiai szempontból szintén fontosak az irányított hálózatok **jellemzői úthossz**ai, mert az információk ezeken az irányított útvonalakon (kaszkádokon) futhatnak végig. A TR és ST hálózatokban gyakoriak a 4-5 hosszúságú útvonalak, ellenben a transzláció szabályozás egy bipartit hálózat, ezért nincsenek benne 1 lépésnél hosszabb kaszkádok. Tehát a **transzláció szabályozási hálózat önmagában csak** a jelek „**finomhangolás**át" végzi. A transzláció szabályozásnak ez a szerepe a transzkripció szabályozással közös evolúciója során fejlődött ki [10].

A fent leírt szerkezeti eltérések miatt a **négy hálózat típus moduljai** matematikai és biológiai értelemben is erősen **eltérőek**. A fehérje-fehérje kölcsönhatási hálózatokban irányítatlan modulokat szokás azonosítani. Ismereteim szerint emögött **két fő biológiai meggondolás** áll. Az első az, hogy egy általános fehérje-fehérje kölcsönhatás mindkét irányban képes információt továbbítani, ami természetesen csupán egy egyszerűsítő közelítése a kölcsönhatások során lezajló összetett biokémiai folyamatoknak. A második biológiai meggondolás az, hogy az élő sejtek folyamataiban erős a kompartmentalizáció, és ez a térbeli kötöttség együttműködő fehérjék csoportjait alakítja ki, amelyekben nincsen alá- és fölérendeltség, tehát nincsen irányítottság. Ehhez az irányítatlan működéshez képest lényegesen eltér a transzkripció szabályozása (TR) és a jelátvitel (ST). Mindkét esetben a sejt egy konkrét helyéről egy másik helyre kell eljuttatni információt. A TR esetében nagyrészt a citoplazmából a DNS-hez, míg az ST esetében főként a sejtmembrántól a sejtmag határáig. Mindhárom esettől (PPI, TR, ST) eltér a transzláció csendesítése, amely visszacsatolási és előrecsatolási szabályozó „körök" részeként a messenger RNS-ek koncentrációját módosítja.

(4) „A SignaLink adatbázis, amely elindításában a jelölt fontos szerepet játszott, az elmúlt években sokat bővült. Lehet-e ezt az adatbázist is elemezni a modulok szintjén? Jellemző-e a jelátviteli útvonalakra is az átfedő modulok megléte, illetve ezek összefüggésben hozhatóak-e a meglévő útvonalakkal, vagy pedig továbbra is döntően lineáris kaszkádnak kell tekintenünk a jelátviteli útvonalakat?

A SignaLink adatbázisban lévő jelátviteli hálózatot lehet a modulok szintjén elemezni. Az előző válaszban leírtak alapján a jelátviteli **hálózatok moduljai** jelentősen **eltérnek** a fehérje-fehérje kölcsönhatási hálózatok moduljaitól, ezért a k-klikk perkolációs klaszterektől eltérő módszerre van szükség. A jelátviteli hálózatok leggyakoribb hálózati modul felosztása az, amikor a jelátviteli fehérjéket a „klasszikus" jelátviteli útvonalak – mint modul középpontok – köré csoportosítjuk. A dolgozat ezt a módszert használja, például a Jak/STAT, a TGF-β és a Notch klasszikus útvonalakból kiindulva [11]. A dolgozat eredményei szerint ezek a modulok erősen átfednek, és a lineáris kaszkád kép helyett megfelelőbb az a kép, hogy a **sejtmembránt és a sejtmagot összeköti több háló** (általában egynél több kezdő- és végponttal), amik egymással erősen kölcsönhatnak és átfednek.

A **jelátviteli hálózatok** irodalmában a legtöbb modulkereső módszer a hálózattal együtt felhasználja a csúcspontok (fehérjék) számos biológiai tulajdonságát is. Az irodalomban vannak **tisztán hálózati alapú modul keresés**i módszerek is. Az egyik ilyen módszer kiinduló gondolata (a hálózat irányítottsága alapján) az, hogy a jelátvitel során két fehérje csoport akkor tekinthető különálló egységnek, ha a két csoport között legalább az egyik irányban kevés a kölcsönhatás [12]. Ezt a tulajdonságot a szerzők a „retroaktivitás" hiányának nevezik, és szemléletesen azt jelenti, hogy a két csoport vagy (közel) független, vagy csak az egyik irányban van erős hatás (vezérlés), és jelentős visszahatás nincsen. Ezzel a módszerrel egymással **nem vagy csak gyengén átfedő modulok** kaphatóak, amelyek a klasszikus útvonal alapú modulokon belüli alegységeket definiálnak.

Az itt leírtakon túl fontos megemlíteni, hogy a **modul** szónak a jelátvitelben van egy igen gyakori, és a hálózatoknál régebben használatos értelme. Az aminosav szekvencia, a térszerkezet és a funkció alapján együttesen megállapítható jellegzetes fehérje szakaszokat gyakran doménnek hívják, és a több élőlényben előforduló doméneket szokás modulnak nevezni.

(5) „Mivel a dolgozatból hiányzik egyfajta kitekintés, ezért ezzel kapcsolatban a következő kérdéseket fogalmazom meg: Hogyan látja a jelölt a hálózatkutatás biológiai rendszerekre való alkalmazásának jövőjét? Vannak-e még megválaszolatlan kérdések a területen? Vannak-e olyan molekuláris biológiai hálózatok, amelyek átfedő modulokkal nem vizsgálhatóak? Vannak-e olyan előrejelzései a modelleknek, amelyek a biológusok kísérletes igazolására várnak?"

Véleményem szerint a molekuláris biológiai kölcsönhatások **hálózatos vizsgálat**ának egyik legfontosabb eredménye egy **nézőpont váltás** (paradigma váltás). A kísérletes kutatásokban és a gyógyszerkutatásban a mai napig szokás egy vizsgált jelenséget néhány kiemelkedően fontos fehérje (vagy gén) hatásaként értelmezni. A transzkripció szabályozási (TR), a transzláció csendesítési és jelátviteli (ST) hálózatokban ezeknek a fontos géneknek és mikroRNS-eknek külön neve van: „driver gene" [13] vagy „master regulator" [14]. Hasonlóan a fehérje-fehérje kölcsönhatások esetén is – még napjainkban is – gyakori a fehérjék éles szétválasztása „fontos" és „nem fontos" fehérjékre. Ehhez a megközelítéshez képest jelentős változást hozott a molekuláris biológiai kölcsönhatások hálózatos vizsgálata. A hálózatok segítségével sok esetben (a korábbi, főként kvalitatív leírásoktól eltérően) kvantitatív módon sikerült megállapítani, hogy statisztikai értelemben hibás a fehérjéket (vagy géneket, vagy mikroRNS-eket) élesen szétválasztani „fontos" és „nem fontos" csoportba. Az ok a következő. A kölcsönhatási hálózatok alapján az látható, hogy ha minden egyes fehérje esetén megszámoljuk, hogy az adott fehérjének hány darab kölcsönhatása van (ez nagyjából jelzi a fontosságát), akkor a kapcsolatok száma sok esetben egy széles és folytonos eloszlás, amit nem lehet „szétvágni" fontos és nem fontos részekre [15]. Más szavakkal: a fokszámok valószínűség-sűrűség függvénye gyakran hatványfüggvényhez közeli, és így nincsen olyan jellegzetes (karakterisztikus) fokszám érték, ami fölött a fokszámok mind „magasak" (és alatta alacsonyak).

Természetesen vannak kivételek. Például (ld. [T5] 1c ábra) az élesztőgomba irányított TR hálózatában a befokszám (egy gént hány fehérje szabályoz) jóval gyorsabban csökkenő eloszlású, mint a kifokszám (egy fehérje hány gén transzkripcióját szabályozza). Itt a fokszám eloszlások eltérésének a biológiai oka a következő. A befokszámok eloszlását az határozza meg, hogy egy gént csak annyi fehérje tud szabályozni, amennyinek „elfér" a kötési helye a gén leolvasási szakasza „felett" (a DNS-en a gén 5' vége előtt lévő nagyjából 5-600bp hosszúságú szakaszon). Viszont arra jóval kevésbé van korlát, hogy egy fehérje kötőhelye hány másik gén leolvasási szakasza előtt található meg, tehát a transzkripciós faktor fehérjék kifokszáma kevésbé korlátozott. Ez alapján egy sejtben általában helyes azt a néhány gént elkülöníteni, amelyik a legerősebb transzkripciós szabályozás alatt áll, viszont **félrevezető lehet elkülönítve kezelni** azokat a transzkripciós faktor fehérjéket, amelyek a **legnagyobb szabályozó hatás**sal rendelkeznek a sejtben. Ebből a konkrét kivételből levonható egy fontos, általános érvényű következtetés. Több hálózatos mennyiség (például a fokszám) eloszlását – és ezen keresztül a résztvevő molekula típusok fontosságának eloszlását – megbecsülhetjük a **rendelkezésre álló erőforrások** (például hely vagy energia) eloszlása alapján. A konkrét példában a gén leolvasási szakaszok feletti hely a fő korlátozó tényező.

Véleményem szerint a molekuláris biológiai kölcsönhatási hálózatok esetén a kutatások egyik **központi kérdése a kezdetektől fogva az, és a jövőben is az lesz**, hogy mik **a hálózatok alkalmazhatóságának határai**. Ez a kérdés rögtön a mérési adatok kapcsán fel-

merül. A dolgozatban leírtak alapján látható, hogy a kölcsönhatásokat mérő módszerek (például a Tandem Affinity Purification) jelentős része nem molekula párokra vonatkozik, hanem csoportokra, és sokszor csupán a mérések utáni adatfeldolgozás készít ezekből a nyers adatokból a párokra vonatkozó kapcsolatokat (hálózati éleket). Az élek mérése helyett a csoportok mérése a legtöbb esetben biológiailag erősen megalapozott. Például a fehérje-fehérje kölcsönhatások esetén ismert, hogy a biológiai feladatokat nagyrészt kettőnél több fehérjéből álló komplex-ek (azonos időben összekapcsolódó molekulák) vagy funkcionális modulok végzik el [16]. A transzkripció szabályozásban ismert, hogy az egyes transzkripciós faktor (TF) fehérjék szabályozási hatása függ a már jelenlévő TF-ektől [17]. A transzláció szabályozásában és a jelátvitelben is gyakori ez a jelenség, és fizikai kapcsolat keletkezése esetén a fő oka a „cooperative binding". A jelenlegi ismeretek alapján a négyféle hálózat típusra közös, **általános elv**ként megfogalmazható, hogy a molekuláris biológiában a hálózatok jól alkalmazhatóak akkor, amikor a jelentős (például nagy fokszámú) csúcsok az egymáshoz való kapcsolódás helyett inkább a kevésbé jelentős (kis fokszámú) csúcsokhoz kapcsolódnak, azaz, a fokszámok (vagy a befolyást mérő egyéb mennyiségek) diszasszortatívak. A **PPI hálózatokban a fokszámok jellemzően disz-asszortatívak** [18], ezért a hálózatok a biológiai jelenségek jó közelítését adják. Ezzel szemben a hálózatok használatánál **pontosabb lehet a modulok használata** olyan esetekben, amikor a fokszámok asszortatívak. Erre egy konkrét példa a transzkripciós faktorok hálózata [19] (nem azonos a dolgozatban tárgyalt TR hálózattal), amikor a hálózatoknál pontosabb lehet a fehérje csoportok vizsgálata.

Kitekintésként érdemes megemlíteni a molekuláris biológiai hálózatok átfedő csoportjainak egy lehetséges alkalmazását. A gyorsan fejlődő élő szöveti mintavételezés (biopszia) [20, 21], DNS szekvenálás [22] és a daganatos sejtek klonális evolúciójának gráfokkal való rekonstrukciója [23] alapján a **rákkutatásban** a következő évtizedekben **folyamatos és jelentős fejlődésre számítok**. A daganatos sejtek növekedésének és osztódásának megváltozását a szabályozás változása teszi lehetővé. Az előző bekezdésben említett transzkripciós faktor hálózathoz [18] hasonlóan itt is hibás lehet csupán néhány kiemelt szabályozó fehérje (master regulator, driver gene) elemzése és a többi teljes elhanyagolása. Napjainkban még sokszor ez a kutatás fő elve. Véleményem szerint ha sikerül egyidejűleg több transzkripciós faktor fehérjét befolyásolni, azaz **transzkripciós faktorok egymással átfedő csoportjai**t, akkor lehet, hogy néhány évtized múlva a szervezetünkben folyamatosan jelen lévő, megváltozott szabályozású sejtekkel úgy fogunk együtt élni, mint jelenleg az influenza vírusokkal. Egy kis megfázással nem szoktunk foglalkozni, de ha a tüdőgyulladásnak vagy egy kezdődő asztmának a lehetősége felmerül, akkor vannak eszközeink a határozott cselekvésre. Ezen a területen a **konkrét előrejelzések** kidolgozásán és a **kísérletes kutatók számára javasolt kérdések**en túl kiemelkedően fontosnak tartom a heti rendszerességű kapcsolatot a kutató és betegekkel dolgozó orvosok között. Mi – számítógépes és elméleti kutatók – csak így követhetjük folyamatosan, hogy milyen matematikai és statisztikus fizikai eszközök használatára van lehetőség az **orvosi gyakorlat**ban.

Budapest, 2016. május 12.

Farkas Illés

Hivatkozott irodalom

A DOI számok a publikációkra mutató hiperlinkeket tartalmaznak.

[T2] B. Adamcsek, G. Palla, I. J. Farkas, I. Derényi, T. Vicsek: CFinder: locating cliques and overlapping modules in biological networks. Bioinformatics 22, 1021–1023 (2006) doi:10.1093/bioinformatics/btl039.

[T5] I. J. Farkas, C. Wu, C. Chennubhotla, I. Bahar, Z. N. Oltvai: Topological basis of signal integration in the transcriptional-regulatory network of the yeast, Saccharomyces cerevisiae. BMC Bioinformatics 7, Article Number: 478 (2006) doi:10.1186/1471-2105-7-478.

[1] R. Albert, H. Jeong, A.-L. Barabási. Error and attack tolerance of complex networks. Nature 406, 378–382 (2000) doi:10.1038/35019019.

[2] A.-L. Barabási. Network Science Book. A könyv 8. fejezetében található 8.10. számú ábra. http://barabasilab.neu.edu/networksciencebook/downlPDF.html.

[3] L. Lü, T. Zhou. Link prediction in complex networks: A survey. Physica A 390, 1150 (2011) doi:10.1016/j.physa.2010.11.027.

[4] B. Barzel, A.-L. Barabási. Network link prediction by global silencing of indirect correlations. Nature Biotechnology 31, 720–725 (2013) doi:10.1038/nbt.2601.

[5] C. V. Cannistraci, G. Alanis-Lobato, T. Ravasi. From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. Scientific Reports 3, Article Number: 1613 (2013) doi:10.1038/srep01613.

[6] U. Guldener, et. al. CYGD: the Comprehensive Yeast Genome Database. Nucleic Acids Research 33, D364–D368 (2005) doi:10.1093/nar/gki053.

[7] U. Alon. Network motifs: theory and experimental approaches. Nature Reviews Genetics 8, 450–460 (2008) doi:10.1038/nrg2102.

[8] A. Ma'ayan. Formation of Regulatory Patterns During Signal Propagation in a Mammalian Cellular Network. Science 309, 1078–1083 (2005) doi:10.1126/science.1108876.

[9] F. Fazi, et. al. A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPα regulates human granulopoiesis. Cell 123, 819–831 (2005) doi:10.1016/j.cell.2005.09.023.

[10] K. Chen, N. Rajewsky. The evolution of gene regulation by transcription factors and microRNAs. Nature Reviews Genetics 8, 93–103 (2007) doi:10.1038/nrg1990.

[11] A. Pires-daSilva, R. J. Sommer: The evolution of signalling pathways in animal development. Nature Reviews Genetics 4, 39–49 (2003). DOI:10.1038/nrg977.

[12] J. Saez-Rodriguez, et. al. Automatic decomposition of kinetic models of signaling networks minimizing the retroactivity among modules. Bioinformatics 24, i213–i219 (2008) doi:10.1093/bioinformatics/btn289.

[13] D. Tamborero, et.al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. Scientific Reports 3, Article number: 2650 (2013) doi:10.1038/srep02650.

[14] S. S. Chan, M. Kyba. What is a Master Regulator? J. Stem Cell. Res. Ther. 3, 2 (2013) doi:10.4172/2157-7633.1000e114.

[15] R. Albert: Scale-free networks in cell biology. Journal of Cell Science. 118, 4947–4957 (2005) doi:10.1242/jcs.02714.

[16] L. H. Hartwell, J. J. Hopfield, S. Leibler, A. W. Murray: From molecular to modular cell biology. Nature 402, C47–52 (1999) Link a cikk ingyenes PDF file-jára.

[17] G. K. Ackers, A. D. Johnson, M. A. Shea. Quantitative model for gene regulation by lambda phage repressor. PNAS 79, 1129–1133 (1982) doi:10.1073/pnas.79.4.1129.

[18] S. Maslov, K. Sneppen: Specificity and stability in topology of protein networks. Science 296, 910–913 (2002) doi:10.1126/science.1065103.

[19] D. A. Pechenick, J. L. Payne, J. H. Moore. Phenotypic Robustness and the Assortativity Signature of Human Transcription Factor Networks. PLoS Computational Biology. 10, e1003780 (2014) doi:10.1371/journal.pcbi.1003780.

[20] M. Basik, et. al. Biopsies: next-generation biospecimens for tailoring therapies. Nature Reviews Clinical Oncology 10, 437–450 (2013) doi:10.1038/nrclinonc.2013.101.

[21] C. S. Ross-Innes, et.al. Whole-genome sequencing provides new insights into the clonal architecture of… . Nature Genetics 47, 1038–1046 (2015) doi:10.1038/ng.3357.

[22] K. A. Wetterstrand. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP): http://www.genome.gov/sequencingcosts.

[23] N. Beerenwinkel, R. F. Schwarz, M. Gerstung, F. Markowetz. Cancer evolution: mathematical models and computational inference. Syst. Biol. 64, e1 (2015) doi:10.1093/sysbio/syu081.