

A jelátviteli hálózatok regulációja és tumorokban való vizsgálata

Doktori tézisek

Dr. Módos Dezső

Semmelweis Egyetem

Patológiai Tudományok Doktori Iskola



Témavezető: Dr. Földvári-Nagy Lászlóné Dr. Lenti Katalin, PhD,
főiskolai tanár

Konzulens: Dr. Korcsmáros Tamás, PhD, tudományos főmunkatárs

Hivatalos bírálók: Dr. Kapuy Orsolya, PhD, egyetemi adjunktus
Dr. Egyed Balázs, PhD, egyetemi adjunktus

Szigorlati bizottság elnöke: Dr. Tóth Sára, PhD, habilitált egyetemi
docens

Szigorlati bizottság tagjai: Dr. Sipeki Szabolcs, PhD, egyetemi adjunktus
Dr. Hegedűs Tamás, PhD, tudományos
főmunkatárs

Budapest

2017

Bevezetés

Az emberi szervezet bonyolult működéséhez rengeteg folyamatnak kell megfelelő szabályozottsággal végbemennie. Egymással kommunikáló rendszerként fonódnak össze a szervek, szövetek és az élővilág legkisebb önállóan működőképes egysége, a sejt is.

Egy sejt sorsa számtalan bejövő jeltől, ingertől függ, és ezekre millióféleképpen válaszolhat pillanatnyi állapotától függően. Ezeknek a jeleknek a feldolgozását végzi a jelátviteli rendszer. Egy ilyen bonyolult rendszer megértéséhez szükségünk van arra, hogy ne csak az egyes biokémiai reakciókat figyeljük meg, hanem a bejövő jeleket feldolgozó rendszert összességét. Ehhez először is olyan gyűjteményekre, adatbázisokra van szükség, amelyek leírják a jelátvitelt, és azok transzkripcionális regulációs visszacsatolásait is tartalmazzák.

Az így felállított adatbázisokból elkészíthető a jelátvitel interakciós modellje. Ennek egyik legegyszerűbb reprezentációja az egyes fehérjéket és a köztük található interakciókat tartalmazó fehérje interakciós hálózat.

A jelátvitel, és ezzel együtt a sejt, mint rendszer megértéséhez és beavatkozások tervezéséhez meg kell találnunk a rendszer kulcsfontosságú, kritikus szabályozási pontjait. Doktori munkámban megvizsgáltam, hogy van-e kitüntetett szerepe az evolúció során duplikálódott és azonos fajban megtalálható géneknek – a paralógoknak – a jelátvitelben. A paralógok lehetnek azon gének, melyek leginkább érintettek a betegségek, különösen a daganatos betegségek kialakulásakor. Ha meghatároztuk a jelátvitelben történő változásokat a különböző kórfolyamatokban, akkor elemezhetjük a

változások hatásait, megtudva, hogy ezek a jelátviteli hálózat mely területén mennek végbe. A jelátvitel daganatban megváltozó elmei közül legtöbbször magukat a jelátvitelben megváltozó fehérjéket elemzik, de figyelmen kívül hagyják a környezetüket. Ha sikerül megismerni a daganatos betegségeket okozó gének jelátviteli környezetét, lehetővé válhat a kórfolyamatokhoz hasonlóan, de ellenkező előjellel felülírni a jelátvitelt, ezzel gyógyítva a rendszerszintű betegségeket.

Célkitűzések

Doktori munkám során célul tűztem ki, hogy:

1. elkészítem a Signalink 2 jelátviteli adatbázis transzkripcionális faktor - célgén kapcsolatokat és miRNS- célgén kapcsolatokat tartalmazó rétegeit,
2. elkészítsek egy algoritmust a transzkripció faktor célgén kapcsolatok jóslásához, és ezt alkalmazom két másik fontos sejtbiológiai folyamat – az autofágia és az NRF2 stressz válasz- regulációs vizsgálatára,
3. meghatározom a jelátvitelben a kritikus paralógokat, és elemzem azok hálózati és biológiai jelentőségét,
4. azonosítom négy nagy prevalenciájú daganatban (colorectalis carcinoma, mell daganat, nem-kissejtes tüdőrák és hepatocellularis carcinoma) érintett fehérjéket a jelátvitelben, és bemutatom környezetüket első szomszédjaikon keresztül, ezzel illusztrálva a daganatok jelátvitelt áthuzalozó stratégiáját,
5. megkeresem és bemutatom mind a kritikus paralógok, mind pedig az első szomszédok gyógyszertervezési és alkalmazási lehetőségeit.

Módszerek

A SignaLink 2.0, az ARN és az NRF2Ome adatbázis regulációs rétegei

Az adatbázisok transzkripcionális-faktor célgén interakcióinak a meghatározásához a JASPAR adatbázisból töltöttem le az egyes transzkripció faktorok (TF) pozíció súlyozott mátrixait. Ezekből meghatároztam, hogy az egyes génekben, illetve tőlük *upstream* illetve *downstream* 2000 nukleotidra hol találhatóak a transzkripció faktor kötőhelyek.

A miRNS-célgén adatbázisok integrálásakor a mikroRNS-eket (miRNS) MIRBASE ID-ra fordítottam a miR2Disease, a doRiNA, a DianaMicroT, a miRDeathDB, a miRanda a PicTar a TargetScan és a TarBase adatbázisokból. A PuTmiR a Transmir és az Encode adatbázis TF-miRNS adatai esetén szintén hasonlóan jártam el. A célgéneket Uniprot ID-ra fordítottam.

A jelátvitelen belüli kritikus paralóg csoportok meghatározása

Egy kritikus paralóg csoport olyan fehérjecsoport, mely

1. rendelkezik paralógokkal a jelátvitelben (evolúciós kritérium)
2. legalább egy tagján átmegy egy szövetspecifikus jelátviteli út (szövetspecifikus gráf elméleti kritérium)
3. legalább egy tagja cross-talkban vesz részt (jelátviteli-biológiai kritérium)

Ennek megfelelően megkerestem a paralógokat az egyes SignaLink 2 jelátviteli fehérjéknek az Inparanoid és az OrthoDB adatbázis segítségével. Egy jelátviteli útnak a ligandtól a transzkripció faktorig tartó kaszkádot

tekinttem. A szövet specifikusságot az ESEMBL adatbázisból töltöttem le, majd nagyobb szervrendszereknek megfelelően összevontam. A cross-talk információt a Signalink adatbázis tartalmazta. A hálózati tulajdonságokat iGraph programmal vizsgáltam meg. A specificitási vizsgálatokhoz a GeneOntology, OMIM, ChEMBL, Cancer Gene Census (CGC) adatbázisokat használtam. A folytonos változók statisztikai tesztjéhez Komogorov-Szmirnov és Wilcoxon rang próbákat, a diszkrét változók függetlenségének vizsgálatára khi négyzet tesztet alkalmaztam.

A daganatos fehérjék első szomszédjainak meghatározása és vizsgálata

Munkám során megkerestem négy szolid daganatban (colrectalis carcinoma, hepatocellularis carcinoma, nem-kissejtes tüdőrák, mell daganat) a CGC szerint mutálódó géneket, illetve a differenciáltan expresszáldó géneket, összesen 1557 Affymetrix HGU133 plus2 microarray chip adatainak felhasználásával. A micorarray chipek normálása RMA módszer segítségével történt, ezáltal összehasonlíthatókká váltak egymással. Ezeket a géneket projektáltam daganatonként különböző jelátviteli (Signalink 2.0, Cui és munkatársai, Reactome) és fehérje interakciós (HPRD, BIOGRID-DIP-INTACT) hálózatokra, hogy megkeressem az első szomszédjaikat. Ezt követően lemértem különböző hálózati paramétereket (fokszám, köztiség, klaszterezettségi együttható) az iGraph program segítségével. A statisztikai tesztekhez Komogorov-Szmirnov és Wilcoxon rang próbákat alkalmaztam. Munkám során gyógyszer-célpontokat is kerestem a daganatban érintett fehérjék első szomszédjai és interakciós partnereik között. Ehhez a ChEMBL adatbázist

használtam. Az adatbázisban akkor tekintetem egy molekulát fehérjéhez kapcsolódónak, ha a kötési affinitás $10 \mu\text{mol/l}$ -nél kisebb volt.

Eredmények

A Signalink 2.0, az ARN és az NRF2Ome adatbázis regulációs rétegei

Az egyes adatbázisokba integrált kapcsoltokat az 1. táblázat tartalmazza. A transzkripciós faktor célgén kapcsolatokban csupán jóslt adatokat tüntettem fel, mivel ezt határoztuk meg az általam kidolgozott módszer segítségével.

Táblázat 1 A regulációs integrált kapcsolatok az egyes adatbázisokban

	Signalink 2.0	NRF2ome	ARN
Jóslt TF-célgén kapcsolatok	23098	8217	2911
miRNS-célgén kapcsolatok	7381553	8157	325680
TF-miRNS kapcsolatok	5227	4982	2192

A jelátvitelen belüli kritikus paralóg csoportok meghatározása

A módszerekben ismertetett definíciónak megfelelően összesen 75 darab kritikus paralóg csoportot és 265 darab kritikus paralógot határoztam meg. Elemzésem kimutatta, hogy ezek a fehérjék a többi jelátviteli fehérjéhez hasonlóan fontos és központi szereppel bírnak, magasabb köztiséggel, fokszámmal rendelkeznek, mint a többi paralóg fehérje ($p \leq 0,05$ Kolmogorov-Szmirnov teszt és Wilcoxon rang próba). A kritikus paralógok gyakrabban vesznek részt daganatos és más betegségek kialakításában a CGC és az OMIM adatbázis adatai alapján, mint más paralóg fehérjék (Khi négyzet teszt $p < 0,05$). A kritikus paralógok különbözőbb Gene Ontology biológiai funkcióval rendelkeznek az egyéb paralógokhoz képest ($p \leq 0,05$ Kolmogorov-Szmirnov teszt és Wilcoxon rang próba). A kritikus paralógok között gyakrabban találhatóak gyógyszercélpontok is, mint más gyógyszercélpontok is, mint más paralóg fehérék között (Khi négyzet teszt

$p < 0,05$). Vizsgálatom alapján tehát sikerült elkülöníteni kétféle paralógot az emberi jelátvitelben, a kritikus paralógot, melyek egymással nem felcserélhetők, míg a paralóg fehérjék egy csoporton belül hasonlóak, és felcserélhetőek egymással.

A daganatos fehérjék első szomszédjainak meghatározása és vizsgálata

Vizsgálatom során egymástól függetlenül négy különböző daganatban (colorectalis carcinoma, hepatocellularis carcinoma, nem-kissejtes tüdőrák, mell daganat) három jelátviteli (SingaLink 2, Reactome, Cui és munkatársai) és két fehérje-fehérje interakciós hálózatban (HPRD, DIP-Intact-BioGrid összevont hálózat) megállapítottam, hogy a daganatban érintett fehérjék első szomszédjai legalább annyira fontosak, mint a daganatban érintett fehérjék: magasabb vagy ugyanakkora a köztiségük, fokszámuk, klaszterezettségi együtthatójuk ($p \leq 0,05$ Kolmogorov-Szmirnov teszt és Wilcoxon rang próba).

Az első szomszédok egy jobban összetartozó hálózatot alkotnak, mint a daganatban részt vevő fehérjék mind a négy általam vizsgált daganat mind az öt hálózatában (Z analízis óriás komponens százaléka $p < 0,001$). Sőt, minden esetben azt találtam, hogy a differenciáltan expresszálódó fehérjék első szomszédjai magasabb hálózati centralitásokkal rendelkeznek ($p < 0,05$ Kolmogorov-Szmirnov teszt Wilcoxon rang próba), míg a mutálódóké ugyanakkorával ($p > 0,05$ Kolmogorov-Szmirnov teszt). Ez alapján elmondható, hogy a daganatok kétféleképpen huzalozzák át a jelátvitelt. Egyrészt direkt mutációval, amikor centrális fehérjékre hatnak,

másrészt indirekt módon, az első szomszédjaikkal a differenciáltan expresszáladást követően. Megvizsgálva a gyógyszer-célzási teret az szignifikánsan megnő, ha figyelemebe vesszük az első szomszédokat. 122 gyógyszer és 3585 egyéb molekula célozza a 105 daganatban érintett fehérjét. Arányaiban szignifikánsan kevesebb gyógyszer (233) és molekula (6699) célozza a 295 első szomszédot (Bernoulli teszt $p < 0.005$). Ez mutatja a lehetőséget a gyógyszertervezési tér kiterjesztésére a centrális első szomszédok irányában.

Következtetések

Doktori munkám során sikeresen felépítettem a Signalink, az Autofágia Regulációs Hálózat (ARN) és az NRF2ome adatbázis regulációs rétegeit. A fenti adatbázisok közreműködéssel alkalmassá váltak a regulációs körök vizsgálatára, amik egyes lehetőségeit a dolgozat diszkussziójában ismertettem.

A három adatbázis közül a Signalink2-t használtam a jelátviteli hálózatok analízisének bemutatására. Sikeresen azonosítottam, és analizáltam a kritikus paralóg csoportokat és kritikus paralógokat az ember legfontosabb 7 jelátviteli útvonalában. Ehhez két, egymást kiegészítő paralóg kereső adatforrást használtam. Az általam bemutatott metodika képes lehet a jelenleg nem a jelátvitelbe tartozó paralógok jelátviteli útvonalakhoz való rendelésére.

A paralógok megtalálása után összeállítottam egy munkafolyamatot a kritikus paralógok és paralóg csoportok azonosítására. Ehhez jelátviteli tulajdonságként a cross-talkokban betöltött szerepet használtam, és esszenciálisnak vettem azokat a fehérjéket, melyek nélkül eltűnik a jelátvitelben egy adott szövetben a ligandtól transzkripciós faktorig tartó jelátviteli út. Mindössze ez a két kritérium elegendőnek bizonyult, hogy megkülönböztethessük a kritikus és a nem kritikus paralóg csoportokat egymástól. A kritikus paralógok azonosításával lehetőség nyílik a gyógyszercélpontok keresésére, hiszen a kritikus paralóg fehérjék között jóval több a betegségben részt vevő fehérje, és különbözőbb funkcióval rendelkeznek, mint az egyéb paralóg fehérjék. Mindez azt mutatja, hogy a

kritikus paralógok egymással nem felcserélhetők, míg a paralóg fehérjék képesek lehetnek átvenni egymás funkcióját.

A daganatban érintett első szomszédok vizsgálatával rámutattam, hogy az alkalmazott gyógyszercélpontok körét érdemes kiterjeszteni a daganatban érintett fehérjék első szomszédjaival. Az első szomszédok ugyanolyan vagy magasabb centralitás értékekkel rendelkeznek, mint a tumor fehérjék, és mint ragasztó tartják össze a jelátviteli hálózatot, és küldik tovább a daganatban érintett fehérjék változásait.

Az egyes daganatokban a daganatban érintett fehérjék kétféle stratégiával érik el a teljes hálózatot: a drasztikusabb mutáció közvetlenül, jelátvitel centrálisabb elemeire hat. Ezzel szemben a differenciáltan expresszálandó daganatban érintett fehérjék közvetetten befolyásolják a jelátvitelt az első szomszédjaikon keresztül. Az első szomszédok figyelembe vételével egyes, már használt gyógyszereknek adhatunk új indikációt, és megnagyobbíthatjuk a gyógyszercélzási teret.

Doktori munkám során tehát részt vettem több adatbázis létrehozásában, melyekben specifikus intracelluláris folyamatok vizsgálhatók, úgymint az autofágia, a jelátvitel és az oxidációs stressz. Az adatbázisok közül a SignaLink 2 adatbázis elemzése során megállapítottam, melyek lehetnek azok a paralóg csoportok, melyek jó gyógyszercélpontok lehetnek. Végezetül a jelátviteli és fehérje-fehérje interakciós hálózatok figyelembe vételével bemutattam a daganatos fehérjék első szomszédjainak fontosságát, és azt, hogy ezek a fehérjék hogyan képesek a gyógyszercélpontok körét kitágítani.

Saját publikációk

1. A dolgozathoz kapcsolódó publikációk

A SignaLink 2.0, az ARN és az NRF2Ome adatbázis regulációs rétegei:

Papp D, Lenti K, **Módos D**, Fazekas D, Dúl Z, Türei D, Földvári-Nagy L, Nussinov R, Csermely P, Korcsmáros T (2012) The NRF2-related interactome and regulome contain multifunctional proteins and fine-tuned autoregulatory loops *FEBS Letters* 586, 13, 1795–1802.

IF:3,58

Fazekas D.*, Koltai M.*, Türei D.*, **Módos D**, Pálffy M, Dúl Z, Zsákai L, Szalay-Bekő M, Lenti K, Farkas I, Vellai T, Csermely P és Korcsmáros T (2013) SignaLink 2 – a signaling pathway resource with multi-layered regulatory networks *BMC Systems Biology* 7:7.

IF: 2,85

Türei D., Papp D., Fazekas D., Földvári-Nagy L., **Módos D.**, Lenti K., Csermely P. és Korcsmáros T. (2013) NRF2-ome: an integrated web resource to discover protein interaction and regulatory networks of NRF2 *Oxidative Medicine and Cellular Longevity*, 2013: 737591.

IF: 3,36

Türei D, Földvári-Nagy L, Fazekas D, **Módos D**, Kubisch J, Kadlecsek T, Demeter A, Lenti K, Csermely P, Vellai T és Korcsmáros T. (2015)

Autophagy Regulatory Network – a systems-level bioinformatics resource for studying autophagy components and their regulation *Autophagy*, 11(1):155-165.

IF: 9,11

A jelátvitelen belüli kritikus paralóg csoportok meghatározása:

Módos D., Brooks J., Fazekas D., Ari E., Vellai T., Csermely P., Korcsmaros T és Lenti K. (2016) Identification of critical paralog groups with indispensable role in the regulation of signaling flow *Scientific Reports* 6: 38588 13p

IF: 4,26

A daganatos fehérjék első szomszédjainak meghatározása és vizsgálata:

Módos D., Bulusu K. C., Fazekas D, Kubisch J., Brooks J., Marczell I, Szabó P. M., Vellai T., Csermely P., Lenti K., Bender A. és Korcsmáros T. (2017) Neighbours of cancer-related proteins have key influence on pathogenesis and could increase the drug target space for anti-cancer therapies *npj Systems Biology and Applications*, 3: (1) 2 13 p.

2. *Nem kapcsolódó publikációk*

Földvári-Nagy Cs, Földvári-Nagy Cs. K., **Módos D**, Lenti K. (2016) Comparative study of hygiene habits in three different groups in Hungary *New Medicine* 20: (4) 141-147

Csályi K, Fazekas D, Kadlecsek T, Türei D, Gul L, Horváth B, **Módos D**, Demeter A, Pápai N, Lenti K, Csermely P, Vellai T, Korcsmáros T és Varga

M (2016) SignaFish: A Zebrafish-Specific Signaling Pathway Resource. *Zebrafish* 13: (6) 541-544

IF:2,24

Földvári-Nagy Cs, **Módos D.**, Feith H. J., Lenti K. (2015) Quantitative study of the generational changes among relationship habits in highly educated Hungarian population *New Medicine* 19: (3) 100-109.

Perez-Lopez Á. R., Szalay K. Z., Türei D., **Módos D.**, Lenti K., Korcsmáros T, és Csermely P. (2015) Targets of drugs are generally, and targets of drugs having side effects are specifically good spreaders of human interactome perturbations. *Scientific Reports* 5 10182. 9p

IF: 5,23

Csermely P, Hódsági J, Korcsmáros T, **Módos D.**, Perez-Lopez Á. R, Szalay K, Veres D. V, Lenti K, Wu L.Y., Zhang X. S. (2014) Cancer stem cells display extremely large evolvability: alternating plastic and rigid networks as a potential Mechanism: Network models, novel therapeutic target strategies, and the contributions of hypoxia, inflammation and cellular senescence. *Seminars in Cancer Biology* pii: S1044-579X(13)00130-2.

IF: 9,33

Gyurkó D M, Veres D V, **Módos D.**, Lenti K, Korcsmáros T és Csermely P.: Adaptation and learning of molecular networks as a description of

cancer development at the systems-level: Potential use in anti-cancer therapies. (2013) *Seminars in Cancer Biology* 23: 262–9

IF: 9,14

Köszönetnyilvánítás

Szeretném megköszönni a rengeteg támogatást Lenti Katalinnak témavezetőmnek, akihez nem csak szakmai, hanem lelki segítségért is fordulhattam. Korcsmáros Tamásnak, konzulensemnek, akivel, ha kellett az éjszaka közepén is alkottunk a dolgozat alapját szolgáló TDK előadást és a cikkeket. Csermely Péternek, aki átadta hálózatos szemléletét és megismertetett Tamással. Dörnyei Gabriellának, aki lehetőséget adott gyakorlatokat és előadásokat tartani PhD hallgatóként az Egészségtudományi Kar Morfológiai és Fiziológiai Tanszékén, ezzel elősegítve előadói képességem fejlődését. Vellai Tibornak, aki lehetőséget adott, hogy az ELTE Genetikai Tanszéken is dolgozhassak. Fazekas Dávidnak, aki segített első lépéseimnél a programozás területén, sok hasznos ötletet adott és közben pár sört is elfogyasztottunk. Türei Dénesnek, aki elkészítette az adatbázisok magját alkotó SQL honlapokat és a dizájnt. Szuromi Gábornak, aki segített nekem a TF-célgének szekvenciáinak a legenerálásában. Marczell Istvánnak és Szabó Péternek, akik segítettek megoldani a microarrayek egymáshoz normálását. Kubish Jánosnak, akivel együtt kerestük ki azt az ezerötszáz microarrayt, amivel dolgoztunk. Ari Eszternek hasznos megjegyzéseiért és a közös, gyümölcsöző ebédékért. Földvári-Nagy Lászlónak, aki átnézte az autofágia fejezetet az ő irodalmi ismereteivel tudtuk elkészíteni az autofágia regulációs adatbázist. Johanne Brooksnak a cikkeim angol lektorálásáért és megjegyzéseiért. Krishna Bulusunak, aki megtanított, hogy működik a ChEMBL adatbázis, és Andreas Bendernek, aki alkalmazott, még PhD nélküli post-docként.

Hálával tartozom a Morfológiai és Fiziológiai Tanszék valamennyi munkatársának, akik meghallgattak és segítettek. A NetBiol csoport és az ELTE Genetikai Tanszék valamennyi tagjának a rengeteg hasznos tanácsért, ötletért, közös kávéért.

Végül családomnak és barátaimnak, akik elviselték, hogy távol vagyok tőlük, mert a kutatásaimon dolgozom, és a végén ők is hajtottak, hogy befejezzem disszertációm.