

Biológiai hálózatok dinamikájának vizsgálata

Doktori értekezés

Szalay Kristóf Zsolt

Semmelweis Egyetem

Molekuláris Orvostudományok Doktori Iskola



Témavezető: Dr. Csermely Péter egyetemi tanár, az MTA tagja

Hivatalos bírálók: Dr. Hamar Péter egyetemi docens, Ph.D.

Dr. Oborny Beáta egyetemi docens, Ph.D.

Szigorlati bizottság elnöke: Dr. Falus András egyetemi tanár, az MTA tagja

Szigorlati bizottság tagjai: Dr. Enyedi Péter egyetemi tanár, az MTA doktora

Dr. Palla Gergely tudományos főmunkatárs, Ph.D.

Budapest

2014

Tartalomjegyzék

1. Rövidítések jegyzéke	4
2. Bevezetés	5
2.1. Hálózatok	5
2.1.1. Gráfok	5
2.1.2. Kisvilágság	9
2.1.3. Skálafüggetlenség	10
2.1.4. Csoportok	13
2.1.5. Központiság	13
2.1.6. Rendszerjellemezők	15
2.2. Biológiai hálózatok	16
2.2.1. Fehérjeszerkezet-hálózatok	16
2.2.2. Metabolikus hálózat	17
2.2.3. Jelátviteli hálózatok	17
2.2.4. Gén-kölcsönhatási hálózatok	18
2.2.5. Interaktómok	18
2.2.5.1. Az élesztő két-hibrid rendszer	19
2.2.6. Ideghálózatok	20
2.3. Hálózatok mozgásban	21
2.3.1. Hálózatos rendszerelmélet	21
2.3.2. Attraktorok	22
2.3.3. Biológiai rendszerek attraktorai	25
2.4. Mesterséges intelligencia	27
2.4.1. Keresés gráfokban	27
2.4.2. Lokális keresés	29
2.4.3. Nyaláb keresés	30
2.4.4. Genetikus algoritmus	31

3. Célkitűzések	33
4. Módszerek	34
4.1. A Turbine program	34
4.1.1. Szimulátor	37
4.1.2. Iterált szimulációk	38
4.1.2.1. Attraktorkeresés	39
4.1.2.2. Beavatkozás-tervezés	39
4.1.2.3. Attraktortervezés	40
4.1.2.4. Sebesség és pontosság	40
4.2. Alkalmazott hálózatok	42
4.2.1. Metionil-tRNS szintetáz fehérje szerkezet	42
4.2.2. Citokróm P450 2B4 fehérje szerkezet	42
4.2.3. Élesztő fehérje-fehérje kölcsönhatás	43
4.2.4. T-LGL jelátvitel	43
4.2.5. Humán daganatos jelátvitel	43
4.3. Dinamikus modellek	44
4.3.1. Közlekedőedények-modell és terjesztési központosság	44
4.3.2. Folytonos jelátviteli modell	46
4.3.3. Boole-modell	46
4.3.4. Statisztikai módszerek	47
5. Eredmények	48
5.1. Általános modellezés	48
5.1.1. tRNS kötőhelyek meghatározása	48
5.1.2. Másodlagos struktúrák megkülönböztetése	50
5.1.3. Élesztő stresszben tapasztalható hálózatos funkcióváltozásai	52
5.2. A jelátvitel specifikus modellezése	54
5.2.1. A jelátviteli modell	54
5.2.2. Verifikáció a T-LGL modellen	56
5.2.3. Humán daganatos jelátviteli modell	58
5.2.3.1. Daganatos mutációs profilok	61

5.2.3.2.	Specifikus célponttervezés	63
5.2.3.3.	Tervezés mellékhatások minimalizálásával	65
5.2.3.4.	Zavaró mutációk	67
6.	Megbeszélés	69
6.1.	A Turbine program	69
6.2.	Terjesztési központiság	72
6.3.	Attraktorok és beavatkozás-tervezés	74
7.	Következtetések	82
8.	Összefoglalás	85
9.	Summary	86
10.	Irodalomjegyzék	87
	Ábrák jegyzéke	103
	Táblázatok jegyzéke	105
	Függelékek	106
A.	Königsberg hídjai	106
B.	A Milgram-kísérlet	107
C.	A MetRS és a CyP450 aminosavainak closeness központiságai	108
D.	A MetRS és a CyP450 aminosavainak betweenness központiságai	109
E.	Élesztő fehérje-fehérje kölcsönhatás-hálózatában a legmagasabb terjesztési központiságú pontokban dúsuló GO funkciók nyugalmi és stresszelt állapotokban	110
F.	Élesztő fehérje-fehérje kölcsönhatás-hálózatában a stressz hatására legjobban növekedő terjesztési központiságú pontokban dúsuló GO funkciók különféle stresszelt állapotokban	113
G.	Élesztő fehérje-fehérje kölcsönhatás-hálózatában a stressz hatására legjobban csökkenő terjesztési központiságú pontokban dúsuló GO funkciók különféle stresszelt állapotokban	116

1. Rövidítések jegyzéke

BP	belief propagation ("gyanúterjedés" algoritmus)
CDAT	Turbine complex network data (adatfile)
CNET	Turbine complex network (hálózatfile)
CyP450	citokróm P450
DAG	directed acyclic graph (irányított körmentes gráf)
GA	genetikus algoritmus
GCN	gene co-expression network (gén-kölcsönhatási hálózat)
GEXF	Graph Exchange XML Format (hálózatfile)
GML	Graph Markup Language (hálózatfile)
GO	Gene Ontology (adatbázis)
GWAS	Genome-wide association study (genomi asszociációs vizsgálat)
MetRS	metionil-tRNS szintetáz
MI	mesterséges intelligencia
MPI	Message Passing Interface (szuperszámítógépek belső üzenetküldő rendszere)
PC	perturbation centrality (terjesztési központiség)
PDB	Protein Data Bank (adatbázis)
PPI	protein-protein interaction network (fehérje-fehérje kölcsönhatási hálózat)
PSN	protein structure network (fehérjeszerkezet-hálózat)
RIN	residue interaction network (fehérjeszerkezet-hálózat)
T-LGL	T - large granular lymphocyte (nagy granulált T-limfocita)
TCP	transmission control protocol (Internet-protokoll)
vdW	van der Waals

2. Bevezetés

„A tudás a világ alkotóelemeiről szerzett ismeret. A bölcsesség az elemek kapcsolódásának az ismerete.” —Jiddu Krishnamurti

2.1. Hálózatok

A hálózatok körülvesznek minket. Hálózatot alkotnak a baráti társaságaink, az emberi társadalmak, az internet, a gazdasági rendszerek, az élővilág, és – nem utolsósorban – hálózatok találhatóak bennünk is. A hálózat tulajdonképpen egy komplex rendszerek leírására széleskörűen használt modell, melynek alapfeltevése, hogy bonyolult rendszerek működése megérthetővé válik, ha a rendszert nem egyben, fekete dobozként, hanem elemeire és az elemek közötti kölcsönhatásokra szétbontva vizsgáljuk. A hálózatos megközelítés nem új, de mások mellett különösen Barabási–Albert László munkássága által az elmúlt másfél évtizedben vált különösen népszerű módszerré. Nagy sikerrel alkalmaztak hálózatokat társadalomtudósok (társadalmi, szociális hálók), ökológusok (táplálékhálózatok), közgazdászok (tőzsdei, banki hálózatok) és persze az orvostudomány művelői is.

2.1.1. Gráfok

Hogyan is kell elképzelni egy hálózatot? A hálózat magát a modellezett rendszert jelenti, ezt általában egy gráffal reprezentáljuk. Definíció szerint egy gráf pontok vagy csúcsok (node) és élek (edge, link) halmaza. A pontok a hálózat elemeit, az élek pedig az elemek kölcsönhatásait ábrázolják. Ha egy hálózat két eleme valamilyen – az adott hálózattól függő – módon *közvetlen* kölcsönhatásban van egymással, akkor az adott két pont között a hálózatot ábrázoló gráfban is futni fog egy él. A közvetlenség itt fontos; a modellezett rendszer minden közvetett kölcsönhatásának adódnia kell a közvetlen kölcsönhatásokból összeköttetések valamilyen láncolatán keresztül, különben valahol hiányzik egy vagy több csúcs a gráfból. Ilyen hiányzó elemek keresése önmagában is fontos kutatási kérdés lehet. Ez az egyik kulcsa a hálózattudomány megvilágító erejének is: hálózatok használatával

megtudhatjuk, hogy az általában keresett közvetett hatás közvetlen kapcsolatok milyen láncolatán keresztül valósul meg, így a vizsgált folyamat háttere, és a lehetséges egyéb zavaró tényezők megérthetővé és szabályozhatóvá válnak.

A gráfokat általában síkban ábrázoljuk, a gráf pontjait körökkel, az éleket pedig a körök között futó vonalakkal vagy nyilakkal jelöljük. A gráfelmélet alapjait Leonhard Euler fektette le 1735-ben a Königsbergi hidak híres problémájából kiindulva (A. függelék), és azóta a matematika egy teljesen önálló ágává nőtte ki magát, így itt csak egy nagyon rövid bevezetőt tudok adni róla azon fogalmak ismertetésével, amelyeket a dolgozatban használok.

A gráf **pontszáma** vagy mérete a teljes gráf pontjainak száma. Egy pont **fokszáma** a pontból kiinduló vagy abba belépő élek száma. Ha csak a kiinduló vagy csak a belépő éleket vizsgáljuk, definiálhatunk **bemeneti fokszámot** (indegree) és **kimeneti fokszámot** (outdegree) a két fogalom elkülönítésére.

A gráfoknak viszonylag sok fajtáját különböztetjük meg. **Egyszerű gráfoknak** nevezünk azokat a gráfokat, melyekben két pont között legfeljebb egy él fut, és nem tartalmaznak **hurokét**, azaz olyan élt, melynek kezdő-, és végpontja ugyanaz a pont (1. ábra A panel). Az ezen tulajdonsággal nem rendelkező gráfokat **összetett gráfoknak** (1. ábra B panel) nevezünk. Ez a megkülönböztetés inkább csak a matematikában jelentős.

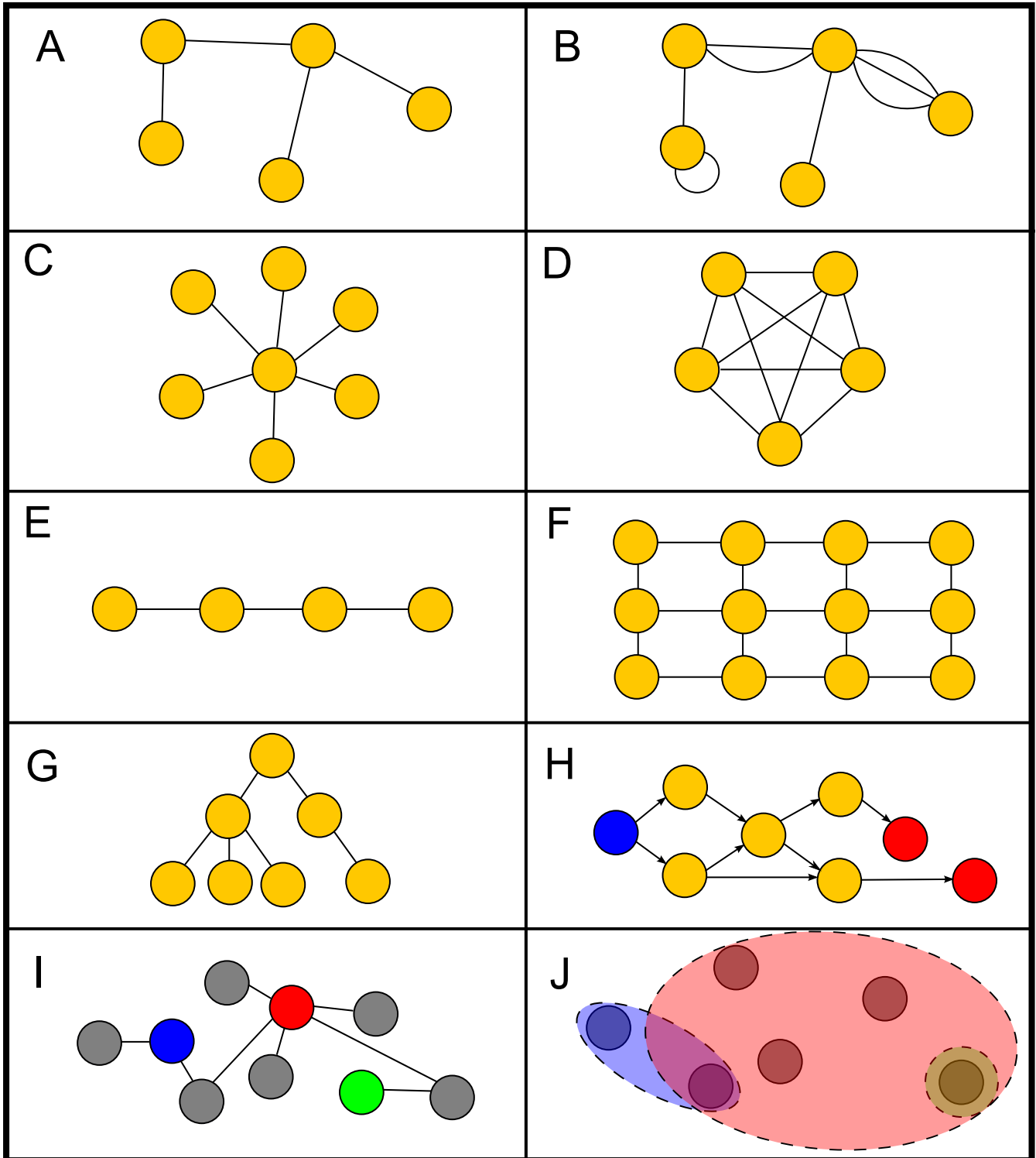
Fontosabb kérdés a gráf összefüggősége. Egy **összefüggő gráfban** az éleken történő lépések sorozatával egy pontból bármely másik pont elérhető, azaz nem tudunk különálló **részgráfokat** leválasztani róla. Ha a gráf nem is egészen összefüggő, kérdés, hogy található-e benne **óriáskomponens**, azaz néhány különálló kis csoportot leszámítva a gráf pontjainak túlnyomó része egy összefüggő csoportban van-e. A legtöbb valós hálózatban találunk ilyen óriáskomponenst. Chung és Lu 2002-ben megmutatta [1], hogy amennyiben egy Erdős-Rényi véletlen gráf [2] pontjainak átlagos fokszáma nagyobb, mint 1, már jó eséllyel találunk benne óriáskomponenst. A valós hálózatok átlagos fokszámára gyakorlatilag mindig teljesül ez a feltétel (amellett, hogy később kiderült, hogy a "véletlen gráf" modellnél jobb leírásukat is adhatjuk), ezért általában csak pár darab "leszakadt" elemet találunk bennük, és azok is inkább az elem elégtelen kutatottságára utalnak.

Egy másfajta csoportosítási lehetőség a gráf irányítottsága. Egy él **irányított**, ha a két elem egymásra hatása nem kölcsönös, csak az egyik elem hat a másikra, visszafe-

le irányuló effektus nincs. Jelátviteli hálózatokra többnyire ilyen típusú kölcsönhatások jellemzőek. Az él ezzel szemben **irányítatlan**, ha a hatás kölcsönös, vagy nem ismert az iránya. Az irányított éleket többnyire nyíllal, az irányítatlanokat pedig egyszerű vonallal jelöljük. Technikai okokból általában az egész gráfot vagy irányítatlanként vagy irányítottként ábrázoljuk. Irányított gráfban irányítatlan élnek megfelelő struktúrát létrehozhatunk úgy, hogy az adott két pont között két irányított élt alakítunk ki, ellentétes irányítással.

Vannak bizonyos speciális struktúrájú gráfok, melyek az általános esetnél könnyebben kezelhetőek. Ilyen például a **csillagháló** (1. ábra C panel), ahol az összes elem kizárólag egy kiemelt középponthez csatlakozik, a **teljes gráf** vagy klikk (clique) (1. ábra D panel), ahol minden pont minden másikkal össze van kötve, a **lánc** (1. ábra E panel), ahol a pontok csak a két szomszédjukhoz vannak kötve, a **rácsháló** (1. ábra F panel), amelyben hasonlóan csak a közvetlen "földrajzi" szomszédokkal vannak összekötve a pontok, de a pontok egy kétdimenziós rácson helyezkednek el, a **fa** (1. ábra G panel), melynek tulajdonsága, hogy nem tartalmaz kört, illetve az irányított körmentes gráf (**DAG**, directed acyclic graph, 1. ábra H panel), melyben csak az a megkötés, hogy akkor nem bukkanhatunk körre, ha az él mentén az irányítást figyelembe véve haladunk végig. Sok helyen használnak fákat az informatikában különféle rendezett adatstruktúrák (halmaz, szótár) létrehozására, a DAG-k pedig verziókövető rendszerekben, vagy egyéb, valamilyen történetet vagy tervet feldolgozó esetekben bukkanak fel természetes módon.

Érdekes struktúra a **páros gráf** (1. ábra I panel), melynek pontjai egyértelműen két halmazra bonthatóak úgy, hogy egyik halmazon *belül* sem fut él. Ilyen gráfok adódnak, ha pontok két különböző típusát kapcsoljuk össze egymással, például filmeket a bennük játszó színészekkel vagy metabolitokat kémiai reakciókkal.



1. ábra. Különböző gráf típusok. **A:** egyszerű gráf. **B:** összetett gráf. **C:** csillagháló. **D:** teljes gráf. **E:** lánc. **F:** rácsháló. **G:** fa. **H:** irányított körmentes gráf (DAG). Az ilyen gráfoknak mindig van egy vagy több forrás (kék) és nyelő (piros) elemük. **I:** páros gráf. A szürke pontok csak színesekkel, a színesek csak szürkékkel vannak összekötve. **J:** Az I-vel ekvivalens hipergráf, ellipszisekkel jelölve a **hiperéleket**.

A **hipergráfok** (1. ábra J panel) olyan struktúrák, melynek élei nem két, hanem tetszőleges számú pontot kötnek össze. Az így létrejövő **hiperéleket** Venn–diagram-szerűen, ellipszisekkel szokás ábrázolni. Észrevehető, hogy a hipergráfok és a páros gráfok megfeleltethetőek egymásnak, azaz egy hipergráf ábrázolható páros gráfként, és fordítva.

2.1.2. Kisvilágság

A hálózattudomány számára a kiemelkedést az a megfigyelés hozta meg, hogy az organikus, önállóan szerveződött hálózatoknak nagyon sok közös, nem triviális tulajdonságuk van, melyek többnyire a hálózat típusától függetlenül teljesülnek. Az első ilyen tulajdonság a kisvilágság. Kísérletesen először Stanley Milgram mutatta meg, hogy az emberi kapcsolatok hálójában két ember sokkal közelebb van egymáshoz, mint gondolnánk ([3], B. függelék). Az ennek eredményeképp becsült átlagos hat lépés távolság azóta gyakorlatilag szállóigévé vált a köznyelvben is. Karinthy Frigyes ezt az eredményt már 1929-ben megsejtette Láncszemek című novellájában:

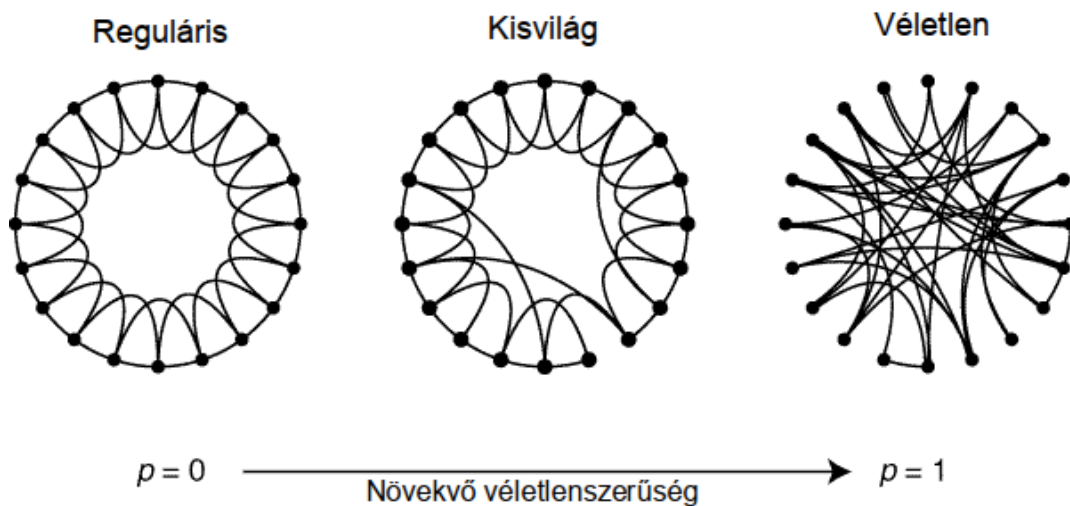
„Annak bizonyításául, hogy a Földgolyó lakossága sokkal közelebb van egymáshoz, mindenféle tekintetben, mint ahogy valaha is volt, próbát ajánlott fel a társaság egyik tagja. Tessék egy akármilyen meghatározható egyént kijelölni a Föld másfél milliárd lakója közül, bármelyik pontján a Földnek - ő fogadást ajánl, hogy legföljebb öt más egyéneken keresztül, kik közül az egyik neki személyes ismerőse, kapcsolatot tud létesíteni az illetővel, csupa közvetlen – ismeretség – alapon. . . ”

Ez valóban így van, nem csak a társadalmi hálózatokban, hanem az interneten, az ökológiában és a sejtek működését leíró hálózatokban is. A hálózat **átmérője** – a leghosszabb út, ami a hálózatban bármely két pont összekötéséhez kell – ugyanis a hálózat méretének növekedésével *logaritmikusan* nő, tehát jóval lassabban, mint maga a hálózat. Ennek a jelenségnek az egyik fő oka, hogy a természetesen kialakult hálózatokban az egymáshoz közeli elemeket összekötő kapcsolatokon kívül van néhány ”hosszútávú” kapcsolat is, mely kapcsolatok hatására az általuk összekötött teljes csoportok relatív közel kerülnek egymáshoz.

Watts és Strogatz 1998-ban készített egy modellt, melynek segítségével kisvilág-hálózatokat lehet alkotni, ezek a Watts–Strogatz hálók [4]. A hálózat a következő algoritmus szerint készül N pontból álló hálózat és K fokszám mellett:

1. Készítsünk egy tetszőlegesen választott N pontból álló szabályos gyűrű-gráfot, melyben minden pont mindkét oldalán $\frac{K}{2}$ szomszédjával van összekötve.
2. Ezután minden élt valamilyen konstans p valószínűséggel kivesszünk, és véletlen helyre helyezünk át.

Ha p érték 0, akkor a hálózatunk egy egyszerű K -szor összekötött gyűrű marad. Ha a p érték közel 1, akkor egy véletlen gráfot kapunk. p alacsony, de nem 0 értékeinél kapunk kisvilág-hálózatokat, a 2. ábrának megfelelően ($K = 4$).



2. ábra. Watts–Strogatz hálózatok különböző β értékek mellett. $p=0$ -nál reguláris gráfot, $p=1$ -nél véletlen gráfot kapunk. $p=0,1$ körüli értékek mellett jön létre a közepén látható kisvilág-hálózat. Az ábra az eredeti cikkből származik [4].

2.1.3. Skálafüggetlenség

A skálafüggetlenség fogalma, illetve annak általánossága Barabási–Albert László és Albert Réka munkája révén lett közismert [5]. Maga a definíció első ránézésre nem túl sokatmondó:

Skálafüggetlen egy hálózat, ha a pontok fokszámának eloszlása hatványtörvényt követ, azaz a következő képlet szerint alakul:

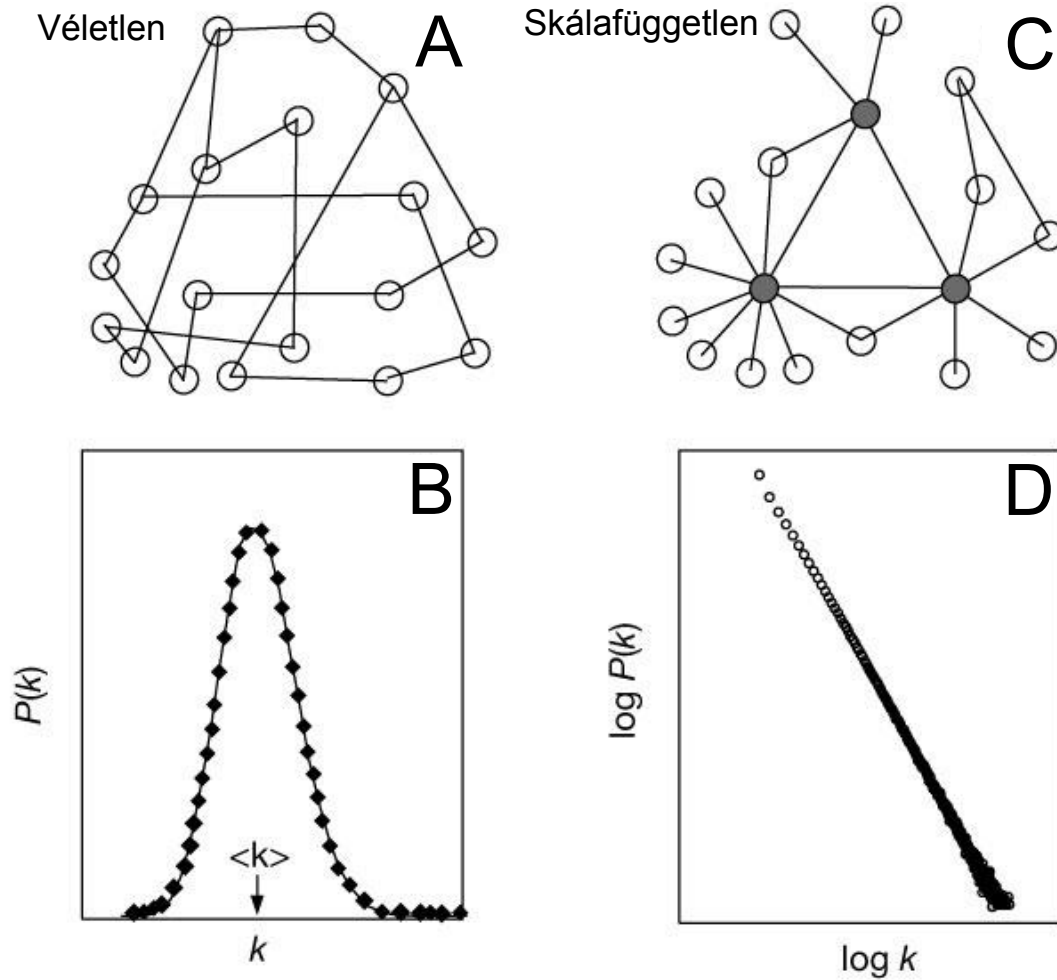
$$P(k) \sim k^{-\gamma}$$

ahol a γ hatványtényező többnyire 2 és 3 között mozog. $P(k)$ jelöli a fokszámeloszlásfüggvényt, azaz az adott k fokszámú pontok számát, k pedig a vizsgált fokszámot.

Ennek következményei viszont igen jelentősek. Azt mutatja a képlet, hogy az ilyen hálózatok fokszámeloszlása az Erdős és Rényi által 1959-ben [2] bevezetett és a matematikai bizonyításokban addig gyakorlatilag kizárólagosan használt **véletlen gráfok** fokszámeloszlásához képest **vastag végű**, azaz jelentős mennyiségű kiemelkedően sok éllel rendelkező pont található bennük. További következmény, hogy a pontok legnagyobb része viszont igen kevés kapcsolattal rendelkezik. 1965-ben írta le először Derek de Solla Price, hogy a kutatók egymásra hivatkozásának hálózatának (citációs hálók) fokszámeloszlása vastag végű [6], és 1999-ben mutatták meg Barabásiék, hogy ez az Internetre, társadalmi és biológiai hálózatokra is igaz [5]. A 3. ábra B és D tábláján látható a véletlen és a skálafüggetlen hálózatok fokszámeloszlásának különbsége. Egy véletlen hálózat Gauss-görbe alakú fokszámeloszlással rendelkezik, ezért az eloszlás végei exponenciális lecsengést mutatnak; a skálafüggetlen hálózat hatványtörvény szerinti lecsengése ennél jóval lassabban történik. A "skálafüggetlen" szó onnan ered, hogy a véletlen hálózatokban a Gauss-görbe csúcsánál, illetve körülötte található egy tipikus fokszámérték, a *skála*, melytől a pontok fokszáma csak kis mértékben tér el. Skálafüggetlen hálózatban nincs ilyen mérték, a hatványtörvény szerint bármelyik pontot is nézzük, egy tízszer akkora fokszámút mindig századannyi eséllyel találunk ($\gamma = 2$ esetén). A hálózatainkban tehát találunk csomópontokat, eredeti szóval **hub**-okat, melyeken a hálózati forgalom nagy része áthalad – legyen az pletyka vagy molekula – és fontos koordinációs szereppel bírnak (a 3. ábra C táblájának szürke pontjai).

Az [5] cikkben Barabásiék egy modellt is bemutattak, melynek segítségével skálafüggetlen hálózatokat lehet létrehozni. Az így készített gráfokat nevezzük Barabási–Albert hálózatoknak. A modell lényege, hogy a „gazdag még gazdagabbá válik”, azaz az újonnan létrejött pontok szívesebben kapcsolódnak a már eddig több kapcsolatot gyűjtött pontokhoz. Az algoritmus a következő (k fokszám-paraméter mellett):

1. Vegyünk egy tetszőleges méretű kiindulási hálózatot.



3. ábra. A véletlen és a skálafüggetlen hálózatok közötti különbség szemléltetése. Míg a véletlen hálózatnak van egy tipikus skálája (a B táblán nyíllal jelölve), a skálafüggetlen hálózatnál nincs ilyen jellemző skála. A skálafüggetlen hálózatok ebből következően tartalmaznak nagy foksámú csomópontokat (a C táblán szürkével jelölve). Az ábra a [7] cikkből származik.

2. Adjunk a hálózathoz egy új pontot.
3. Az új pontot kössük össze k meglévő ponttal úgy, hogy egy adott i . meglévő ponthoz való kapcsolódás esélye $p_i = \frac{k_i}{\sum_j k_j}$ legyen, ahol k_i az adott pont foksáma, j pedig az összes ponton végigmenő változó, így a nagyobb foksámú pontoknak a foksámukkal arányosan nagyobb esélyük van az új él befogadására.
4. Ismételjük a 2–4. lépést, amíg el nem érjük a kívánt méretű hálózatot.

2.1.4. Csoportok

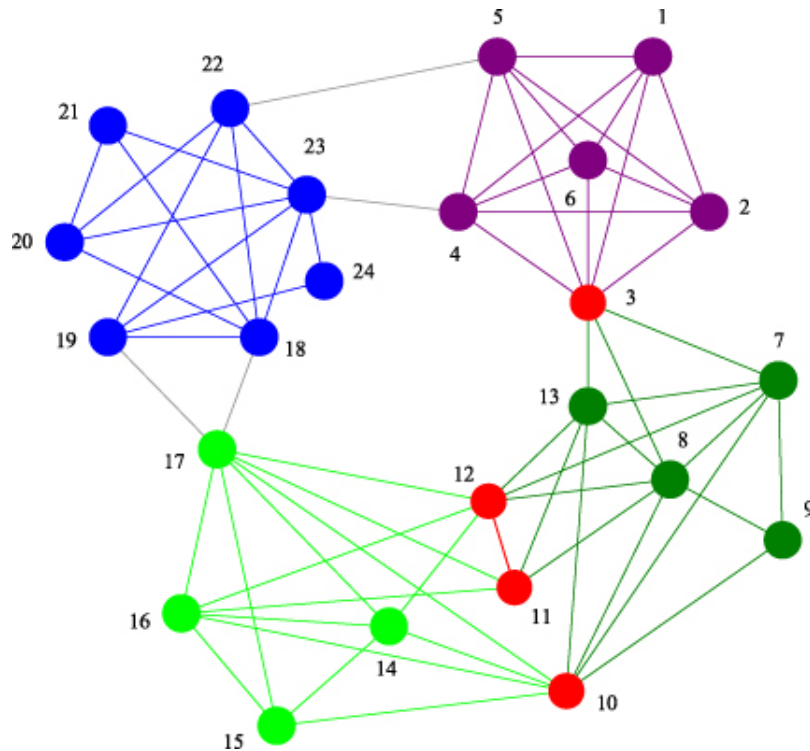
A hálózatos **csoportok** vagy **modulok** (community) pontosabb elemzésének lehetősége már a XXI. században született meg. 2002-ben mutatta meg Girvan és Newman [8], hogy a hálózatok elemei nem egyforma sűrűn vannak összekötve, hanem meghatározhatóak egymással "sűrűbben" összekötött csoportok úgy, hogy a csoportok közötti összeköttetések "lazábbak", belőlük számszerűleg is kevesebb található. Több különféle, egymástól eltérő matematikai definíció is született a modularitás meghatározására, de általánosan továbbra is a fenti egyszerű leírást használjuk. Ezen modulok *megtalálása* azonban – a pontosabb definíciók sokrétűségéből adódóan – messze nem egyértelmű feladat. Az elmúlt években több, mint 100 különféle módszer született modulok azonosítására, közülük egy a munkacsoportunkban [9].

Mire jó a modulok megkeresése? Egy modul tagjai általában egyfajta feladat elvégzésére alkalmasak, esetleg egyfajta környezeti hatásnak vannak kitéve, vagy valamilyen más módon összefüggés, hasonlóság van közöttük. Ha találunk egy ismeretlen pontot, és tudjuk, hogy az melyik modulba tartozik, akkor a modul többi tagja alapján tudunk következtetni az ismeretlen pont funkciójára vagy tulajdonságaira.

Szintén érdekesek és fontosak a modulokat összekötő kapcsolatok, vagy *hidak*. Híd-elem lehet két modul közötti él, vagy egy olyan pont is, mely mindkét modulba többé-kevésbé egyformán beletartozik. Ezek a hidak adják a hálózat kisvilág voltát, és szerepük többnyire igen lényeges a hálózatban. Girvan és Newman eredeti módszere [8] egyébként a következő szakaszban bemutatott köztiségi (betweenness) központiság mérőszám használatára épül, ami a modulokat összekötő hídelemeket képes kiemelni. Ezen hídelemek iteratív eltávolításával a hálózatból ugyanis a hálózat szétesik különálló részgráfokra, mely részgráfok az eredetileg keresett moduljaink. Egy ilyen moduláris hálózatot mutat a 4. ábra, a különböző modulokat különböző színekkel, a hídpontokat pirossal jelölve.

2.1.5. Központiság

A hálózatok vizuális ábrázolása mellett sokak fejében megfogalmazódott, hogy szükséges volna egy mérőszám kidolgozása a pontok "fontosságának" mérésére a hálózatban elfoglalt helyük alapján. Mivel a fontosság igen sok mindent takarhat, ennek megfelelően



4. ábra. Moduláris hálózat szemléltetése. A különböző modulok különböző színekkel vannak jelölve az ábrán, a modulok közötti hídelemek pedig pirossal. Az ábra a [10] cikkből származik.

igen sok különféle, ezt mérő **központisági** mérőszám került kidolgozásra, melyek közül most hármat mutatok be részletesebben. Ebből a szempontból természetesen maga a pont fokszáma is egy központisági mérőszám, mely a csomópontokat (hubokat) jelöli meg.

Az első ilyen mérőszám a közelségi központiság [11] (**closeness centrality**). Egy pont hálózatos **távolságát** (farness) az adott pontból a gráf összes többi pontjába vezető legrövidebb utak hosszának átlagaként számoljuk. A közelségi központiság ennek az értéknek a reciproka. Az ezen módszer által megjelölt pontok a gráf ábrázolása során is többnyire az ábra közepén helyezkednek el, ezért nevezik földrajzi központiságnak is.

Másik népszerű mérőszám a köztiségi központiság [12] (**betweenness centrality**). A két szó hasonlósága miatt a továbbiakban az angol megnevezéseket fogom használni. Egy pont **betweenness** értékét úgy számítjuk, hogy meghatározzuk a gráfban minden pontpár között a legrövidebb utat, és kiszámítjuk, hogy ezen legrövidebb utak közül mennyi halad át a kérdéses ponton. Ezen értékek általában a csoportokat összekötő hídelemeken a legmagasabbak.

Irányított hálózatokra alkalmazható a Google által eredetileg weboldalak fontosságának meghatározására kidolgozott **PageRank** módszer [13]. A módszer lényege, hogy annál fontosabb egy weboldal, minél több, minél fontosabbnak ítélt weblap mutat rá. A leírásból észrevehető, hogy ez egy iteratív eljárás, de sajátérték-módszerekkel egy lépésben, algebrailag is kiszámítható. Az algoritmus egy egyszerűsített leírása a következő:

Elsőként $t = 0$ időpontban minden ponthoz egyforma, $\frac{1}{N}$ értékű PageRank értéket rendelünk, ahol N a hálózat pontjainak száma. Ezután a pontok a kimenő éleikkel "szavaznak" egymásra, azaz az aktuális PageRank értéküket szétosztják a kimenő éleik között. A következő körben a pontok az új PageRank értékeikkel szavaznak egymásra, egészen addig, amíg két kör között a PageRank változások értéke elhanyagolhatóvá válik. A teljes algoritmus egy disszipációs (damping) faktort is alkalmaz, melynek értéke általában 0,85. Az algoritmus konvergenciáját matematikailag is bizonyították [14].

2.1.6. Rendszerjellemzők

A rendszerjellemző tulajdonságok (emergent properties) megjelenése csaknem varázslatos jelenség. A rendszerjellemzők olyan tulajdonságai a rendszernek, melyekkel az egyes elemek nem rendelkeznek, csak azok bizonyos konfigurációi esetén állnak elő, és olyan viselkedést mutatnak, melyre az egyedi elemekből nagyon nehéz következtetni. A legszébb példa erre maga az élet, ami kémiai molekulák bizonyos elrendezésének rendszerjellemző tulajdonsága. Ilyen a hangyaboly felderítőinek viselkedése¹, a vadludak V alakba rendeződése (a ludak szárnya mögött felhajtóerő képződik, ahol könnyebb repülni), a hókristályok formája, a sűrűlódás jelensége (az egyes részecskék közötti erők konzervatívak, a sűrűlódás viszont nem konzervatív erő), a konvekció és a turbulencia jelensége folyadékokban és gázokban, és még sok hasonló példát lehetne hozni. Saját véleményem szerint a hálózat kutatás legszébb része, amikor a rendszer összeállított modelljét elindítva ilyen tulajdonságok bukkannak elő. A szépségen túl persze fontos is, mert ezen jelenségek megismerése nélkül csak igen korlátozott vagy pontatlan következtetéseket tudunk levonni a rendszer viselkedésével kapcsolatban.

¹A hangyaboly felderítői gyakorlatilag véletlenszerűen bolyonganak, miközben feromoncsíkot hagynak maguk után. Ha egy másik hangya által hagyott feromoncsíkra bukkannak, akkor annak erősségétől függően bizonyos eséllyel követik azt. Ahol nagyon sok feromon van, ott valószínűleg táplálékot találtak.

2.2. Biológiai hálózatok

A hálózatokról és a dinamikáról összegyűjtött tudás a biológiában egyre nagyobb sebességgel szaporodó adatmennyiség ismeretében igen jól kamatoztatható, és érdekes, újfajta vizsgálatokra nyújt lehetőséget. A biológiában gyakorlatilag csak bonyolult, komplex rendszerekkel foglalkozunk, melyek működésének minél mélyebb megértése kulcsfontosságú feladat. Ezzel párhuzamosan igen sok módon tudunk egy élőlény működésével kapcsolatban hálózatot felírni.

2.2.1. Fehérjeszerkezet-hálózatok

Az atomi szintről felfele építkezve az irodalomban előforduló legalsó szint, ahol már használnak hálózatokat, a fehérjeszerkezet-hálózatok (a megnevezések elég vegyesek a szakirodalomban: protein structure network (PSN), amino acid network és residue interaction network (RIN) is előfordul). Ezek egy fehérje struktúráját írják le, pontjai az adott fehérje egyes aminosavai, élei pedig az aminosavak atomjai között létrejövő fizikai kölcsönhatások valamilyen egyszerűsített ábrázolása. A fizika törvényei szerint elvileg a fehérje minden atomja kölcsönhatásban van minden másik atomjával, de az így létrehozott teljes gráf nem sok használható információt ad. Ezért valamilyen módon ki kell választani a fontos kapcsolatokat, azaz azokat az aminosavakat, amelyek között lényegesen erősebb kölcsönhatás jön létre.

Az első módszer erre, hogy meghatározunk egy minimális fizikai távolságot az aminosavak között (pontosabban az α vagy β szénatomjaik között) amelynél távolabb eső aminosavakat nem tekintjük egymásra hatónak [15]. Ennél modernebb eljárás a RINerator program [16] által használt módszer, melyben először egy Reduce nevű program segítségével visszahelyezzük a röntgen-kristallográfia által nem látható hidrogénatomokat a megfelelő pozíciókba, majd egy Probe nevű program segítségével egy 0.25 \AA sugarú gömböt végiggörgetünk a fehérje atomjainak van der Waals (vdW) sugara által meghatározott felületén. Amennyiben az így görgetett gömb egynél több aminosav atomjaival is érintkezik, úgy kölcsönhatás van a két aminosav között: vonzó, ha az aminosavak van der Waals felületei nem érnek össze, ellenkező esetben taszító. Ezzel a módszerrel a kölcsönhatások erőssége pontosabban feltérképezhető.

Ezeket a hálózatokat többnyire adott fehérjék aktív centrumainak [17] vagy fehérjén belüli kommunikációs útvonalainak [18] felderítésére alkalmazzák.

A fehérjeszerkezet-hálózatok általában nem skálafüggetlenek, mivel egy bizonyos mennyiségnél fizikailag nem fér el több aminosav egy másik környezetében.

2.2.2. Metabolikus hálózat

A metabolikus hálózatok – nevükből adódóan – a sejtek anyagcserejének folyamatait ábrázolják. A hálózat pontjai az egyes molekulák, illetve a molekulák között létrejövő kémiai reakciók. A kétféle elemtípusból adódik, hogy az ilyen hálózatok páros vagy hipergráfként ábrázolhatóak: az egyes molekulákból kiinduló élek az általuk létrehozott reakciókra mutatnak, míg a reakciók a termékként létrejövő molekulákra. A metabolikus reakciók kvantitatívan is jól ismertek és modellezhetőek, ennek ellenére – talán pont az anyagcserefolyamatok jól ismertsége, és kutatási szempontból ”túlhaladottnak” tekintettsége miatt – viszonylag kevés aktív kutatás folyik ebben az irányban. Különböző élőlények metabolikus hálózatai elérhetőek a KEGG [19] és a BioCyc [20] rendszerében, legjobban az *E. coli* metabolikus hálózata tanulmányozott. Átfogó emberi metabolikus hálózat a Recon2, amely 2013-ban került publikálásra [21]. Az *Ectocarpus siliculosus* alga metabolikus hálózata a napokban, 2014. július 25-én került publikálásra [22]. Irodalmi adatok felhasználásával vagy meglévő adatbázisok alapján építenek ilyen hálózatokat.

2.2.3. Jelátviteli hálózatok

A jelátviteli útvonalakat mindig is természetes módon irányított gráfként ábrázolták, azaz az útvonalakban résztvevő fehérjék egyszerűsített ábrázolása mellett a különböző fehérjék közötti kölcsönhatásokat nyilakkal jelölték. A kifejezett hálózatos megközelítés jelentősége tovább nőtt, ahogy egyre több és több útvonalak közötti összeköttetést (átvezetés, cross-talk) fedeztek fel. Egy mai jelátviteli hálózatot – bár a viselkedésében a fő útvonalak komoly szerepet töltenek be – már nem is lehet látványosan útvonalakra szétbontani úgy, hogy közöttük csak néhány átvezetés legyen található, olyan sok útvonalközi kölcsönhatás létezésére derült fény.

Jelátviteli hálózatból az egészen kicsi, egy alrendszer működését modellező hálózat-tól kezdve [23, 24] az egészen nagy, lehetőség szerint mindent leírni igyekvő rendszerig [25, 26] igen sokfélélt találhatunk. A méret növekedésével párhuzamosan az adatok részletessége sajnos csökken; jelenleg csak a legkisebb hálózatok tartalmazznak dinamikai modellezésre alkalmas élsúlyokat.

A jelátviteli hálózatok dinamikus modellezése ettől függetlenül hálás feladat, mivel a jelátviteli fehérjék túlnyomó többsége "kapcsolószerűen" működik, azaz egy csak BE-KI állapotokat megkülönböztetni képes **Boole-modell** segítségével, és logikai (ÉS, VAGY) függvények használatával működésük kielégítően modellezhetőek. Ilyen Boole-modellek használatával több ízben születtek igen prediktív, kísérletesen is alátámasztott eredmények [23, 27].

Jelentőségük igen nagy; mivel a sejt sorsa főként jelátviteli események során dől el, ezért ezen hálózatok minél pontosabb feltérképezésével és elemzésével válaszokat kaphatunk a különböző fejlődéstani, sejtpusztulási vagy túlszaporodási patológiás folyamatok biomechanizmusára, és lehetőség nyílhat célzott módon történő kezelésükre is.

2.2.4. Gén-kölcsönhatási hálózatok

A gén-kölcsönhatási hálózatok (gene co-expression network, **GCN**) a gének egymás közötti (a fehérjehálózat által közvetített) hatásait próbálják feltérképezni. Többnyire microarray vizsgálatokból határozzák meg őket, a vizsgálat során jelentősen korrelált vagy antikorrelált expressziójú gének kerülnek összekötésre [28]. Ilyen hálózatház általában nehezen rendelhető egyértelmű dinamika, inkább csak topológiai, statisztikai vizsgálatokra alkalmazzák ezeket. Nagy előnyük, hogy az aktuálisan vizsgált sejtvonalhoz könnyen elkészíthetőek, akár egy kísérlet során több alkalommal is [29], és közvetlen a genetikai változások által okozott hatásokba nyújtanak betekintést.

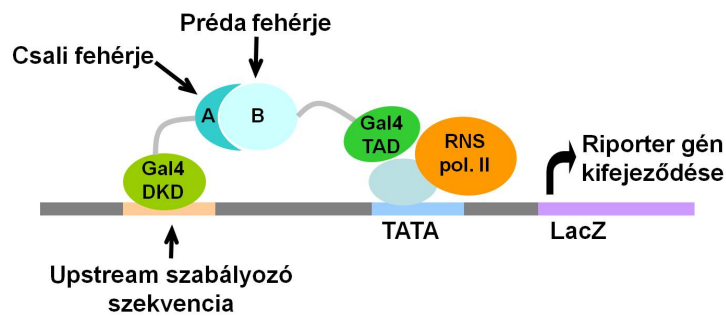
2.2.5. Interaktómok

A fehérje-fehérje kölcsönhatási hálózatok (**PPI**, protein-protein interaction network), vagy más nevükön interaktómok olyan hálók, melyek elemei a jelátviteli hálózatokhoz hasonlóan a sejtben található fehérjék típusai, azzal a különbséggel, hogy a fehérje-fehérje köl-

csönhatási hálózatok törekednek arra, hogy lehetőség szerint az adott sejt összes fehérje-típusát tartalmazzák. Két fehérje között egy ilyen hálózatban általában akkor fut él, ha a két fehérje közvetlenül, fizikailag kapcsolódni képes egymással. Gyengébb értelemben vett PPI-kben más esetekben is futhat él két fehérje között, koncentrációjuk közös változása, vagy egyéb közvetett kapcsolatok alapján, ekkor a gén-kölcsönhatási hálózatok és a fehérje-fehérje kölcsönhatási hálózatok közötti különbség elmosódik. A PPI-k általában irányítatlanok, és jelenleg a gén-kölcsönhatás hálózatokhoz hasonlóan nem alkalmazható egyértelmű dinamikai modell rájuk, ezért inkább topológiai elemzésre alkalmasak, azaz egy érintett fehérje hatását a szervezet működésére a központi értékei, a közvetlen szomszédai, vagy az őt tartalmazó modul funkciója alapján próbálják meghatározni. Ilyen hálózatok a BioGRID [30] és a STRING [31].

2.2.5.1. Az élesztő két-hibrid rendszer

Fehérje-fehérje kölcsönhatási hálózatok összeállítására többnyire automatizált élesztő két-hibrid rendszert használnak. Ezt a rendszert Fields és Song mutatta be 1989-ben [32]; azóta is nagy sikerrel használják, és számtalan változata született. A módszer lényege, hogy az élesztő Gal4 transzkripciós aktivátorának két doménje külön-külön ugyan működésképtelen, de amennyiben az egyik (AD, DKD) domént egy választott "csali" (bait) fehérjével fúziós fehérjeként expresszáltatjuk, a másik (BD, TAD) domént pedig egy választott "préda" (prey) fehérjével fuzionáljuk, és a csali és a préda fehérje egymással kapcsolódni tud, akkor a Gal4 két doménje is elég közel kerül egymáshoz ahhoz, hogy az eredeti transzkripciós aktivátor funkcióját be tudja tölteni (5. ábra). Így a választott reporter gén expressziója csak akkor jön létre, ha a csali és a préda fehérje egymással fizikailag kapcsolódik. A módszer nagysebességű elemzésekre kiterjeszhető, ha az egyik, vagy mindkét partner helyett egy egész cDNS-AD vagy cDNS-BD fúziós könyvtárat alkalmazunk [33]. Az így megtalált interakciókat érdemes további vizsgálatnak is alávetni, mert nem biztos, hogy az interakció a vizsgált élőlényben is létrejön *in vivo* (specifikus poszt-transzlációs módosítások vagy az *in vivo* alacsonyabb expressziós szintek miatt), illetve a fúziós partner letakarhat fontos kötőfelszíneket.



5. ábra. Az élesztő két-hibrid rendszer. A Gal4 transzkripciós aktivátor két doménjét külön-külön, a két vizsgált fehérjével fúziós fehérjeként expresszáltatjuk. Így a Gal4 két doménje csak akkor tud egymáshoz olyan közel kerülni, hogy a riporter gén transzkripcióját aktiválja, ha a két vizsgált fehérje képes egymás kötésére.

Ábra: <http://elte.prompt.hu/sites/default/files/tananyagok/gentechnologia/ch17.html>

2.2.6. Ideghálózatok

Természetes, élő hálózat szervezetünkben az idegsejtjeink hálózata, melyben az idegsejtjeink egymással elektromos vagy kémiai úton kommunikálnak. Az idegsejtek közötti kommunikáció, információterjedés módja igen jól kutatott és ismert, így az idegsejt-hálózatok dinamikája jól modellezhető. Az egyszerű integráló-tüzelő (integrate and fire) modelltől kezdve – amit már 1907-ben (!) leírtak [34] – melyben a modellezett sejt adott disszipáció mellett folyamatosan gyűjti és összeadja a bemenő jeleket, majd ha ez egy küszöbérték felé kerül, akkor "tüzel", a membrán-depolarizáció folyamatát részletesen, ioncsatornánként leíró modellekig [35] sok különböző finomságú elemzés került kidolgozásra. A legelső teljes ideghálózat-modell (**connectome**) a *C. elegans* idegsejtjeinek hálózata, melyet először White és mtsai. írtak le 1986-ban [36]. A *C. elegans* hermafrodita mindössze 302 idegsejttel rendelkezik, de alapvető tanulási és memóriafolyamatok ezzel együtt is tapasztalhatóak *C. elegans*ban [37]. A következő modellorganizmus ideghálózatának rekonstruálása már jóval nagyobb kihívást jelent ma is, hiszen a *D. melanogaster* agya körülbelül 135.000 neuront tartalmaz. Ennek a rekonstrukciója folyamatban van [38], befejezése a következő években várható. Egerek agyi idegsejt hálózatának egyes apróbb részleteit is sikeresen rekonstruálták már [39]. Az emberi agy becslések szerint 86 milliárd neuronjának és 1 trillió kapcsolatának feltérképezése még évtizedeket vehet igénybe.

Idegsejtek hálózatának mesterséges, egyszerűsített változatát (**neurális háló**) sikeresen használják a mesterséges intelligencia területén osztályozásra és mintázatfelismerésre, főleg kézírás- és arcfelismerő rendszerekben gyakori, de robotok vezérlésére is alkalmazzák.

2.3. Hálózatok mozgásban

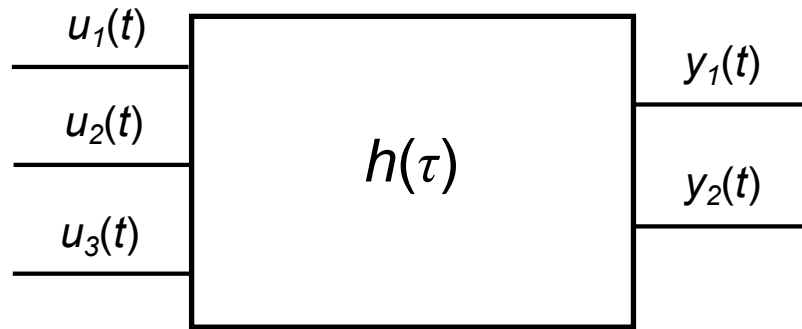
Ahhoz tehát, hogy a hálózatos modell segítségével a rendszert annak rendszerjellemzőivel együtt megismerhessük, és erős predikciókat tudjunk tenni a rendszer várható viselkedésével kapcsolatban, elkerülhetetlenül "be kell kapcsolni" a hálózatot, azaz egy topológiai nézőponton kívül a dinamikát, a rendszer állapotainak változását is figyelembe szükséges vennünk.

2.3.1. Hálózatos rendszerelmélet

Az elektronikai rendszerelmélet (avagy jelek és rendszerek elmélete) a rendszert úgy határozza meg, mint alkotóelemek egy fizikailag összekötött csoportját, mely valamilyen bemeneti jel hatására valamilyen kimeneti jelet produkál. A jel itt bármilyen időben változó fizikai mennyiséget jelent, amely segítségével információ közvetíthető.

Egy ilyen rendszer vázlatosan a 6. ábrán látható módon épül fel (feketedoboz-modell), és lehet egy vagy több **bemenete**, egy vagy több **kimenete**, illetve belső **állapotváltozói**, melyek a rendszer "memóriái", annak belső állapotát határozzák meg, így az állapotváltozók pillanatnyi állapotától függően egy adott bemenetre a rendszer más-más kimenetekkel válaszolhat. A bemenetek hatást gyakorolhatnak közvetlenül a kimenetre, vagy az állapotváltozókra, illetve az állapotváltozók is hatást gyakorolhatnak egymásra, és a kimenetre.

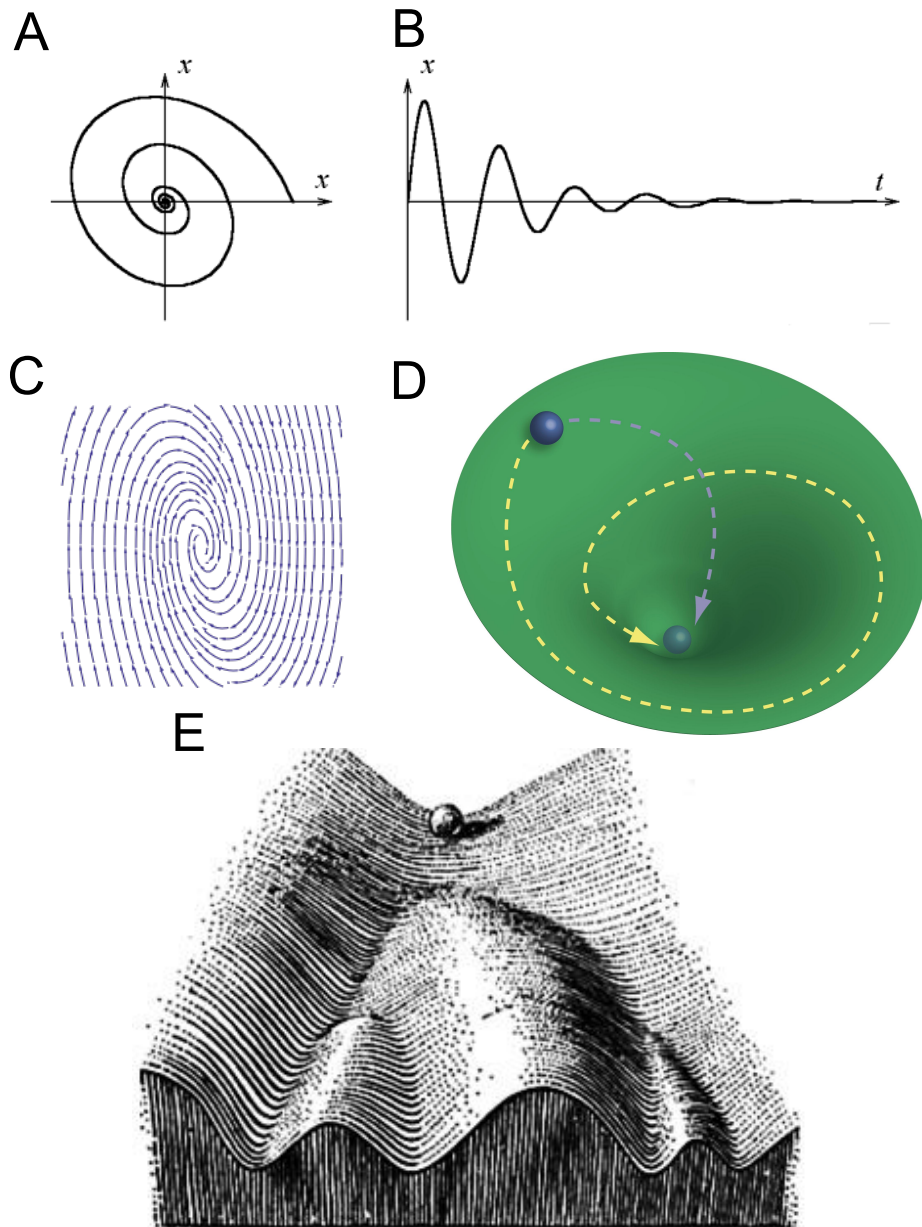
Az elektronikai rendszerelmélet nagyon természetes módon átültethető a hálózatos világba. A hálózat pontjainak állapotai lesznek az állapotváltozóink, az élek pedig az állapotváltozók közötti kölcsönhatások. A hálózatot kívülről tetszőleges ponton érő zavarok lesznek a bemenetek, illetve szintén tetszőlegesen vizsgált pontok értékei, vagy azok valamilyen kombinációban számolt összegei/átlagai a kimenetek.



6. ábra. Dinamikus rendszer feketedoboz-ábrája. Az ábrázolt rendszernek 3 bemenete ($u_1(t), u_2(t), u_3(t)$) és két kimenete ($y_1(t), y_2(t)$) van, melyek időben változnak (a villamosmérnöki konvenciók szerint u -val jelöljük a bemeneteket, y -nal a kimeneteket). A doboz belseje itt ismeretlen, a válasz kiszámításához (lineáris, időinvariáns esetben) a $h(\tau)$ -val jelzett impulzusválasz-függvény ismerete elégséges.

2.3.2. Attraktorok

A vizsgált rendszer működését, adott bemenet hatására történt változásait a leglátványosabb módon a **fázisgörbe** segítségével követhetjük. A fázisgörbe azon pontok összessége, melyeket a rendszer a **fázistérben** bejár a vizsgálat során. A fázistér egy olyan tér, melyben minden állapotváltozó egy dimenzió, azaz egy tengely, tehát az állapotváltozók egy adott időpontban felvett értékeinek összessége pontosan egy pontnak felel meg a fázistérben. Két vagy három állapotváltozó esetén így a teljes fázistér ábrázolható, ennél több esetén annak csak valamilyen vetületét tudjuk megjeleníteni. A fázisgörbéket és a fázisteret több módon is lehet ábrázolni, ezt mutatja a 7. ábra. Egy kétdimenziós fázistér (vagy egy nagyobb fázistér kétdimenziós vetületének) három dimenziós ábrázolása rendkívül intuitív módon, "terepszerűen" (landscape) írja le a rendszer viselkedését (D tábla). Ilyen ábrázolás mellett a rendszerünk többé-kevésbé azt az utat fogja követni a fázistérben, amelyet egy hasonló alakú tárgyra helyezett labda is követne. Waddington már 1957-ben ilyen formában képzelte el a sejtek differenciációjának módját [40] (E tábla). Az ábra A és C tábláin látható, hogy a pontok mozgása egy kiemelt pont felé tart, mintha az "vonzaná" őket. A fázistér ilyen pontjait nevezzük **attraktoroknak**. Ha egy lineáris, elsőrendű rendszerről elvesszük a bemeneteket, akkor az pontosan egy, megadott pontba áll be, azaz egy attraktora van. Nemlineáris rendszerek esetében több attraktort is találhatunk, illetve az attraktorok új fajtái jelenhetnek meg. Lehet az attraktor egy zárt görbe (**határciklus**, limit cycle), ami a rendszer nyugalmi állapotban való periodikus mozgását



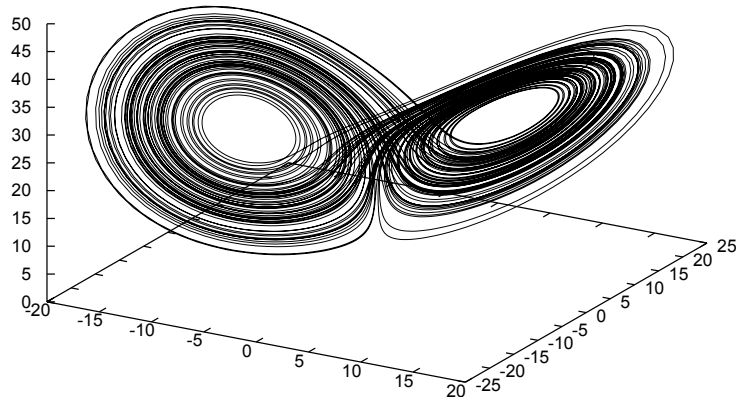
7. ábra. Fázisterek különféle ábrázolásai. Az A, B, C, D táblák egy csillapított inga mozgását írják le. Az A táblán látható a fázisgörbe, egy adott helyről indított inga mozgásának kétdimenziós vetülete. A B tábla ugyanennek a folyamatnak a hagyományos időfüggvényét mutatja. A C táblán a fázistér erővonalas ábrázolása látható, melyről leolvasható, hogy egy adott pontban milyen irányba ható és mekkora "erő" határozza meg a rendszer mozgását. A D táblán egy hasonló rendszer 3 dimenziós, "terep" (landscape) ábrázolása látható, mely rendkívül intuitív, de az erővonalas ábrázolás expresszív, a C táblának megfelelő rendszer nem is ábrázolható terepformában (nem tudjuk minden pontban az óramutató irányába dönteni). Az E tábla Waddington 1957-es elképzelését mutatja a sejtek differenciációjának folyamatáról.

A, B: <http://www.kitnarf.cz/publications/2003/2003.08.esa/2003.08.esa.html>

C: <http://www.entropy.info/scholar/node/dynamical-systems-maps>

D: <http://www.newswise.com/articles/many-paths-few-destinations-how-stem-cells-decide-what-theyll-be>

E: [40].



8. ábra. A Lorenz-attraktor. A Lorenz-attraktor egy "különös" (strange) attraktor, korlátos, de nem periodikus, soha nem ismétli önmagát. Az ilyen attraktorok a kaotikus rendszerek jellemzői. Ábra: <http://www.mizuno.org/c/la/>

mutatja. Amennyiben több periodikus mozgás van a rendszerben, melyek aránya irracionális, akkor az attraktor **határtórusz** lesz. Lehet a rendszernek végül **különös** (strange) attraktora, mely egy korlátos térrészen belül soha nem ismétli önmagát, amire a távolról nézve pillangót formázó Lorenz-attraktor a legnépszerűbb példa (8. ábra)². A rendszer akkor is végül egy attraktorba áll be, ha a bemeneteket nem vesszük teljesen el, csak a bemenetek *transziens* elemeit, azaz meghagyhatjuk az állandó vagy periodikusan változó forrásokat. Az így létrejövő attraktorok viszont nem ugyanazok lesznek, mint amiket a bemenetek teljes elvételével találnánk. Az attraktorok létének fontos tanulsága, hogy egy magára hagyott rendszer nem lehet bárhol a fázistérben, ha magára hagyjuk, akkor legközelebb annak csak kevés, az attraktorok által meghatározott pontjaiban találhatjuk meg.

²Ezek a különös attraktorok **kaotikus** viselkedést jelentenek, azaz az ilyen attraktorokkal rendelkező rendszerek mozgása hosszútávon elvileg sem jósolható. Sok ilyen találunk a természetben, erre a legjobb példa az időjárás viselkedése, ezért van kevés értelme egy hétnél távolabbi jóslásokat tenni. Várható, hogy élő rendszerekben is fogunk ilyen találni, ezért a "tökéletes gyógyszer" valószínűleg nem lehet *előre* megtalálni; elkerülhetetlenné válhat, hogy bizonyos időnként újramérjük a rendszer pozícióját egy beavatkozás után.

2.3.3. Biológiai rendszerek attraktorai

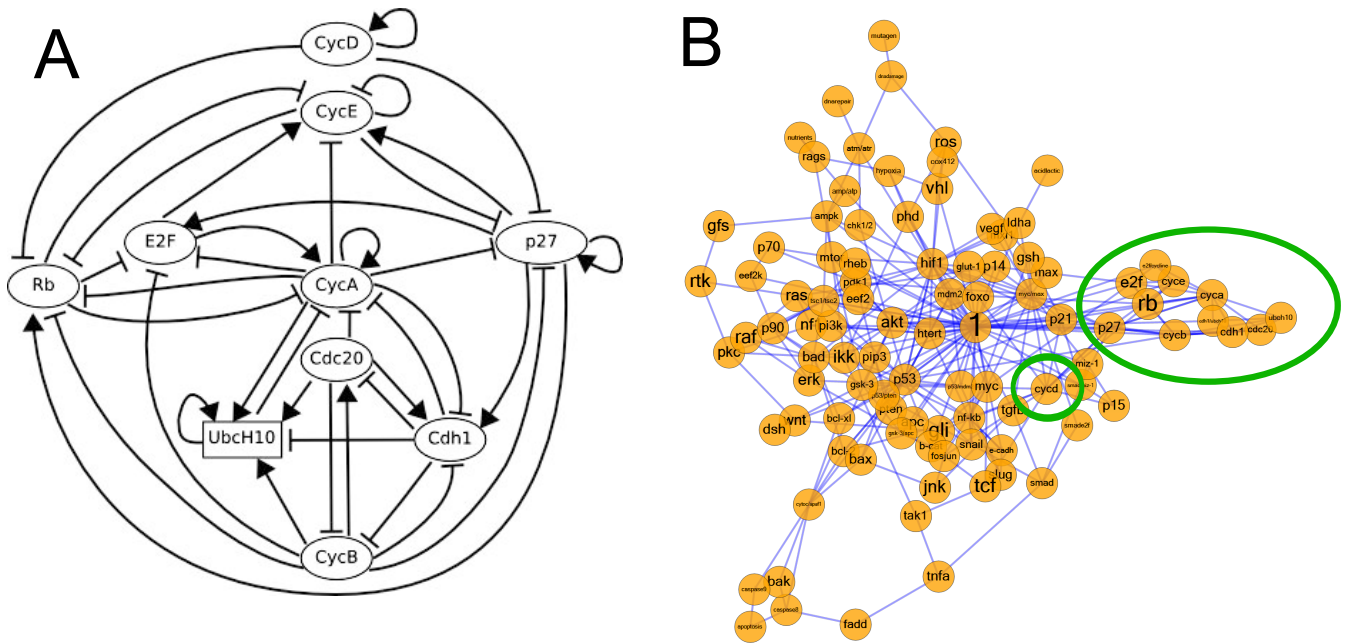
A biológiai rendszerekben az első, attraktoroknak megfeleltethető leírás Waddington 1957-es *The Nature of the Genes* című könyvében [40] található. A biológiai rendszerek ilyen leírását sokáig csak elméleti érdekességként tartották számon, míg Kauffman 1969-ben és 1971-ben [41, 42] meg nem mutatta, hogy azok *felépítésüknél fogva* viszonylag kevés attraktorról kell, hogy rendelkezzenek.

Az idézett cikkekben Kauffman véletlen, Boole-dinamikájú hálózatokban bemutatta, hogy olyan esetekben, ahol minden elem minden elemre hatással van vagy ahol minden elem csak egy másik elemre van hatással, ott az attraktorok száma igen magas is lehet. Amennyiben viszont egy elemre átlagosan kettő-négy másik elem hatott – és a biológiai rendszerek ilyenek – ott az attraktorok száma \sqrt{N} körül szóródott, ahol N a hálózat pontjainak száma. Tehát ezen cikk alapján egy 100 pontos hálózatban 10, egy 1000 pontos hálózatban 31,6, de egy proteom méretű, 100.000 elemű hálózatban is mindössze 316 attraktorra számíthatunk.

Ebben az időben azonban még igen kevés információ állt rendelkezésre a biológiai rendszerekről, Kauffman maga 1.000.000 különböző emberi gén létezését jósolta az idézett cikkben. A P53 fehérjét hat különböző kutatócsoport egymástól függetlenül 1979-ben fedezte fel, a BCL-2-t csak 1984-ben [43], a BAX fehérjét pedig már csak 1993-ban [44] találták meg. Az emberi genom nagy részének 2001-es közzétételével [45, 46], majd a szekvencia 2003-as teljes feldolgozásával, valamint az adatfeldolgozási képességünk exponenciális növekedésével csak a 2000-es évek közepétől állt rendelkezésre elégséges mennyiségű adat jelátviteli rendszerek hálózatos vizsgálatához.

2006-ban Fauré és mtsai. mutatták be az első, valódi modellen végrehajtott attraktor-elemzést [47]. Az általuk használt 10 elemű hálózat a sejtciklus működését modellezte. Ez a 10 pont az általam a későbbiekben használt általános daganatos jelátvitelt leíró 97 pontból álló hálózatnak [24] a jobb "csücskében" helyezkedik el, a 9. ábra szerint. A cikkben szereplő 10 elemű hálózatban összesen két állandósult állapotot találtak, egy nyugalmi és egy proliferáló attraktort.

Ezt követően több cikk is született, ahol különféle rendszerek attraktorait vizsgálták [23, 24, 48, 49]. Ezek közül kiemelkedően részletesen, két teljes cikkben keresztül is



9. ábra. A Fauré és mtsai. (2006) és a Fumiã és Martins (2013) által összeállított dinamikus jelátviteli hálózat összehasonlítása. Az A táblán látható a Fauré és mtsai. által épített hálózat, az eredeti cikk [47] ábrája szerint, míg a B táblán a Fumiã és Martins [24] által épített hálózat Turbine által generált megjelenítése. A B táblán zöld körrel jelölt részek a Fauré hálózat összes pontját tartalmazzák. Az ábra jól mutatja a biológiai információk mennyiségének növekedését is ebben a mindössze 7 éves periódusban.

elemzésre került a Zhang és mtsai. által 2008-ban leírt 60 elemű T-LGL leukémiás hálózat [23, 49]. Ezt a hálózatot alkalmaztam az értekezésemben is verifikációs célból.

Mindennek köszönhetően egyre erősebbé válik az a feltételezés, hogy a jelátviteli rendszerek attraktorai megfeleltethetők lesznek azok különböző lehetséges fiziológiás vagy patológiás állapotainak. Így az a feltételezés sem alaptalan, hogy a sejten belüli hálózatok általános esetben stabil állapotban lévőnek vagy ahhoz közelinek tekinthetők.

Attraktorkereséssel tehát feltérképezhetővé válnak a rendszer lehetséges stabilabb állapotai, ami betekintést nyújthat a vizsgált rendszerünk által elérhető lehetőségekbe, az az által ellátott funkciókba. Ezzel együtt az attraktorok analitikai, diagnosztikai eszközként is használhatóak lehetnek: ha ismerjük a lehetséges attraktorokat, akkor elég lehet a rendszer egy-két elemének állapotát megmérni, ami által egyértelműen meghatározhatóvá válik, hogy melyik attraktorban van a rendszer, majd ezután az ismert attraktor segítségével az *összes többi elem* állapotára tudunk következtetni.

2.4. Mesterséges intelligencia

A mesterséges intelligencia az informatikát kevésbé ismerők részére még ma is kissé misztikus fogalom, pedig az alapelvek kifejezetten könnyen érthetőek.

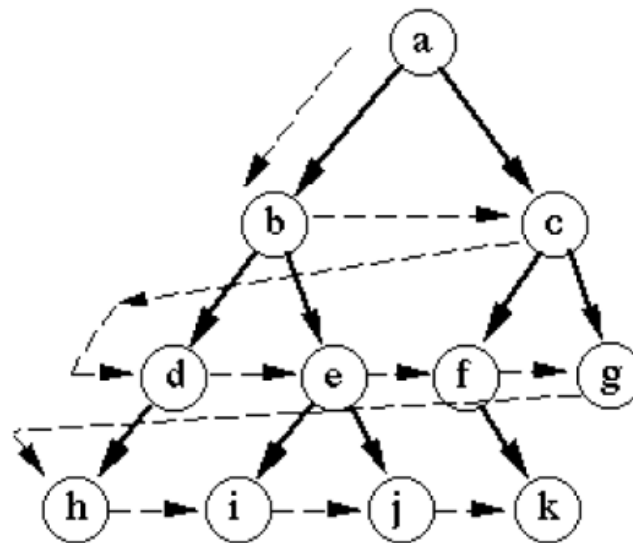
A mi szempontukból a mesterséges intelligencia (MI) egy olyan **ágens** (agent, de a magyar szakirodalom nem az egyszerű "ügynök" fordítást használja), amely tartalmaz egy saját, egyszerűsített belső elképzelést, modellt a körülötte levő világról, és ezen *modell* segítségével képes *megtervezni* a célhoz vezető következő lépést, majd a végrehajtás közben érkező új információkat esetlegesen beépíteni a modelljébe, a "világképebe", és ily módon *tanulni*.

Nem mesterséges intelligencia tehát egy olyan program, mely egyszerű HA X AK-KOR Y utasítások sokaságát tartalmazza, az ilyen rendszerek amúgy is igen könnyen produkálnak hibás viselkedést a bemenetek bármilyen kiszámíthatatlansága mellett, és nagyon hamar nehezen kezelhetővé válnak. A mesterséges intelligenciák "fejében" van egy "szimulátor", azaz képes arra, hogy valamilyen modell szerint "elképzelve", hogy egy adott döntés hatására a környezetének milyen változása várható. Az emberi agyban is pontosan ugyanez a folyamat játszódik le, csak a belső "szimulátorunk" nagyságrendekkel kifinomultabb.

A mesterséges intelligenciák tudománya is egy hatalmas terület, ezért az előzőekhez hasonlóan csak a dolgozatban kifejezetten szerepet játszó megoldásokat ismertetem, azaz a fontosabb tervezési eljárásokat [50].

2.4.1. Keresés gráfokban

A mesterséges intelligencia által történő tervezéshez el kell készítenünk néhány alapelemet: a külső világ egy *ábrázolását*, azaz egy objektumot, amely a világ az ágens számára lényeges állapotát különféle változó állapotának leírásával tartalmazza (pl. sakktábla figuráinak helyzete). Ismernünk kell egy adott pontban *lehetséges döntéseket* (érvényes sakklépések), és tudnunk kell, hogy egy adott döntés (lépés), hogyan hat a modellezett "világ" alakulására (*állapotátmeneti függvény*, pl. lépés utáni új helyzet), és fel kell ismernünk, ha elértük a célt (*célállapotok*, pl. sakk-matt).



10. ábra. Szélességi gráfbejárás. Az állapotgráf ilyen módon történő bejárásakor azon a kiindulási pontból az ábrán látható módon "rétegenként" haladunk végig. Belátható, hogy amennyiben bármely két pont közötti átmenet költsége azonos, a módszer a legrövidebb utat fogja megtalálni a célig. Ábra: <http://www.cse.unsw.edu.au/billw/Justsearch.html>

Amennyiben ezen információkat megadjuk a rendszernek, az különböző algoritmusok segítségével tud egy végrehajtási tervet készíteni, ami a kiinduló állapotunkból a célállapothoz vezet. Ezt a tervet fogjuk felhasználni a későbbiekben.

Könnyen észrevehető, érdekes szinergia, hogy a probléma valójában megfelel egy gráfbejárási problémának. Amennyiben a rendszer minden állapotát egy pontnak tekintjük, és a lehetséges állapotátmenetek a pontokat összekötő élek, úgy a célunk nem más, mint találni egy (optimális esetben a legrövidebb) *irányított* utat a gráfban, mely a kezdőpontból indulva egy célállapotban végződik.

Az így megkonstruált **állapotgráf** hatalmas, és általában teljességében nem bejárható. Sakk esetén az átlagos **elágazási tényező** 35, azaz minden lépésben 35 különféle cselekvési lehetőségünk van. Két lépés esetén $35^2 = 1225$, három lépés esetén $35^3 = 42.875$, azaz a gráf mérete exponenciálisan nő. Nem véletlen, hogy a számítógépek egészen mostanáig igen rosszul teljesítettek a gó nevű játékban, ahol az átlagos elágazási tényező 250. Ezért az algoritmust irányítanunk kell valamilyen módon, hogy ne kelljen a teljes gráfot, vagy akár csak egy mérhető töredékét bejárni.

A legegyszerűbb ilyen algoritmus a **szélességi bejárás** (BFS, breadth-first search), ami a 10. ábrán látható módon "rétegenként" halad előre az állapotgráfban, tehát biztosítani tudja, hogy a "legmagasabban fekvő" megoldást találja meg, azaz azt, ahova a legkevesebb lépésben lehet eljutni. Amennyiben a különböző lépések költsége egyforma, az így talált megoldás optimális.

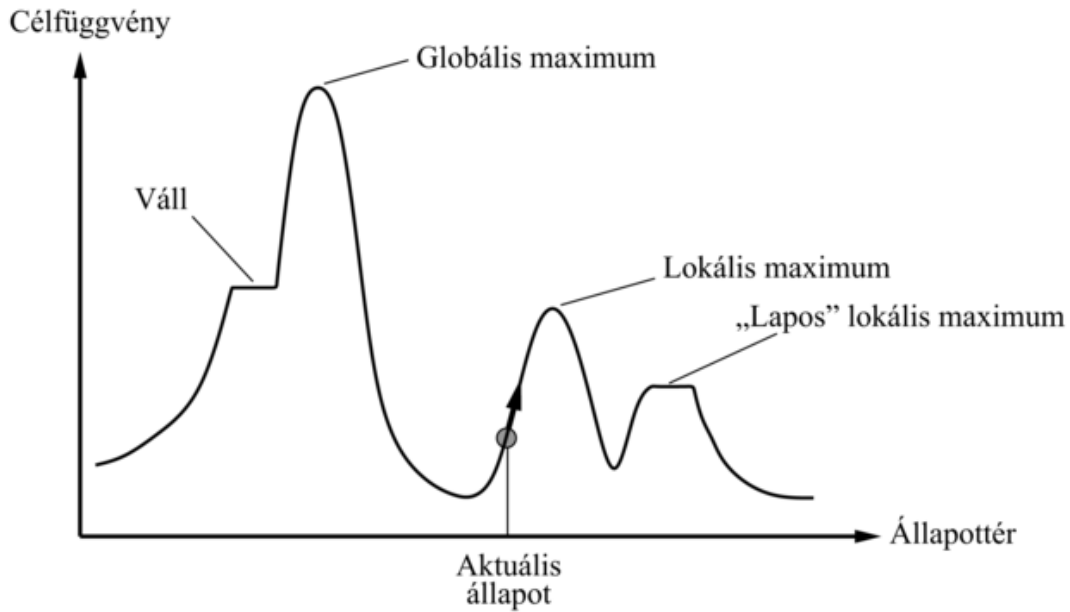
Ennél kicsit bonyolultabbak a **vezérelt algoritmusok**, melyek egy **heurisztika** segítségével megpróbálják "megtippelni", hogy merre érdemes haladni a gráfban, melyik pontok kifejtésétől várható a leggyorsabb előrejutás. Egy igen sokszor használt vezérelt algoritmus az **A***, melynek legismertebb alkalmazása talán a különféle GPS-navigátorok útvonaltervező algoritmusai. Az A* algoritmus képes az optimális megoldás megtalálására változó állapotátmenet-költségek, azaz élsúlyok mellett is, ráadásul a vezéreltség okán jóval gyorsabb, mint az egyszerű szélességi bejárás.

2.4.2. Lokális keresés

Mit lehet tenni abban az esetben, ha már maga az elágazási tényező is rendkívül nagy, vagy akár végtelen? Ebben az esetben nem tudjuk felépíteni az állapotátmeneti gráfot, nem tudunk szisztematikusan haladni benne. Ezért az optimális megoldás megtalálásáról le kell mondanunk. Ilyen esetekben a keresési feltételünk legfeljebb az lehet, hogy egy "elég jó" megoldást találjunk a problémára. Szerencsére ezek az "elég jó" megoldások sok gyakorlati esetben igen közel vannak az optimálishoz, vagy meg is találják azt, de ez nem garantálható, illetve nem tudjuk bizonyítani, hogy a megtalált megoldás optimális.

Ilyen esetben a tervezés szélsőérték-keresési problémára vezethető vissza, azaz egy függvény egy pontjából kiindulva, lépések sorozatával kellene megtalálnunk a függvény globális minimumát vagy maximumát, úgy, hogy a még nem bejárt értékeket nem ismerjük.

A legkézenfekvőbb módszer a **hegymászó algoritmus** (hill climbing) [50]. Ez úgy működik, hogy a vizsgált függvény egy pontjából teszünk egy lépést abba az irányba, amerre a "legmeredekebbnek" tűnik a hegy (vagyis a gradiens irányába), és ezt addig folytatjuk, amíg a hegy csúcsára nem érünk (11. ábra). Mivel azonban a csúcsról a mi esetünkben nem tudunk "körülnézni", hogy látunk-e nagyobbat valamerre, meg kell elégednünk a legközelebbi csúcs megtalálásával. Ez azt jelenti, hogy az algoritmus csak a



11. ábra. A lokális keresés problémája és a hegyászó algoritmus. Az ábrán látható függvény globális maximumát kellene megtalálni. A hegyászó algoritmus a kezdőpontból a legígéretesebb irányba halad, ami ebben az esetben egy lokális maximum megtalálását fogja eredményezni. Az ábra mutat egyéb nehezen kezelhető eseteket is, a "váll" és a "lapos" maximum nem megkülönböztethető, amíg át nem érünk a másik oldalára. Ábra: http://project.mit.bme.hu/mi_almanach/books/aima/ch04s03

legelső *lokális maximumig* jut el, ami a problémát ábrázoló függvény formájától függően lehet egy elég gyenge megoldás is, cserébe ez egy igen gyors algoritmus.

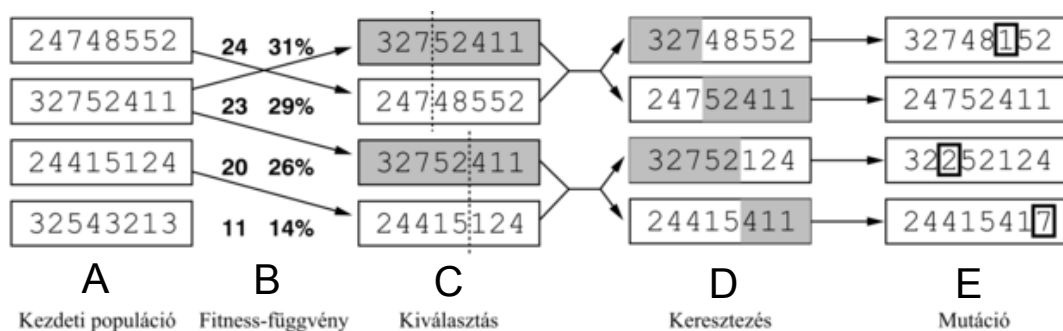
2.4.3. Nyaláb keresés

A hegyászó algoritmus lokális maximumokba való "beragadása" sok esetben komoly problémát jelent. A lokális keresésben a gráfkereső algoritmusokhoz képest komoly visszalépés, hogy egyszerre csak egy pontot vizsgálunk, nincs egy adott "réteg", amelyen belül mérlegelni tudjuk a lehetőségeinket. Ez a limitáció a mai számítógépek teljesítménye mellett teljesen szükségtelen. Ebből a megfontolásból született a **nyaláb keresés** (beam search) algoritmus [50], melynek során tetszőleges számú kezdőpontból több szálon indítunk hegyászó kereséseket, és bizonyos időnként "összerendezzük" a nyalábot, azaz a gyengébben teljesítő szálakat áthelyezzük a legjobban teljesítő szálak környezetébe. Ez az algoritmus már lényegesen jobb megoldásokat eredményez, de az alapvető problémát csak elodázni tudja: előbb-utóbb a nyaláb összes szála egy pont környékére fog sűrűsödni, ahonnan a hegyászó kereséshez hasonlóan nincs továbblépés. Kiutat

szolgáltathat, ha néha egy-egy szálat véletlenszerű helyre pozícionálunk, de nagyon kicsi az esély, hogy az így elhelyezett szál az egészen idáig optimalizált nyálábnál jobb helyre kerül.

2.4.4. Genetikus algoritmus

A biológiában jártas olvasónak feltűnhetett, hogy a nyálábok "átrendezése" hasonlít az evolúció során a természetes szelekció folyamatához. Ezt a hasonlóságot már mások is észrevették, az első teljes **genetikus algoritmust** (GA) Fraser írta le 1957-ben [51]. A keresési folyamatot tovább gondolva, az tovább fejleszthető, ha beletesszük azokat az egyéb mechanizmusokat is, melyeket a természet dolgozott ki az élőlények környezetbe való beleilleszkedésének segítésére. Ilyen szempontból minden bizonnyal mondhatjuk, hogy ez a legrégebben használt, és jól bevált algoritmus.



12. ábra. A genetikus algoritmus működése. A populációból (A) a problémánkat meghatározó fitness-függvény segítségével (B) kiválasztjuk azokat a szülőket, akik a "génjeiket" átörökíthetik (C). Ebből a készletből az utódokat átkeresztezéssel (D) hozzuk létre, majd azok esetleges mutálódása (E) után alakul ki véglegesen a következő generáció. Ábra: http://project.mit.bme.hu/mi_almanach/books/aima/ch04s03

Az algoritmusnak három kulcslépése van: a szelekció, a reprodukció és a mutáció. Készítenünk kell először is egy **kromoszómát**, melyben az adott egyed tulajdonságait írjuk le. Ezután véletlenszerűen kitöltött kromoszómákból létrehozunk egy **populációt**, melyeknek ezután a feladattól függő módon kiszámítjuk a **fitness** értékét (azaz gyakorlatilag megnézzük az adott ponthoz tartozó függvényértéket), majd meghatározzuk a fitness értékeknek megfelelően (**szelekció**, 12. ábra A-C panel), hogy a következő **generáció** elemeit melyik **szülőkből** származtatjuk le. A természethez hasonló szexuális jellegű, két-szülős leszármaztatás újabb optimalizációs lehetőség, hiszen ha mindkét szülő nagy-

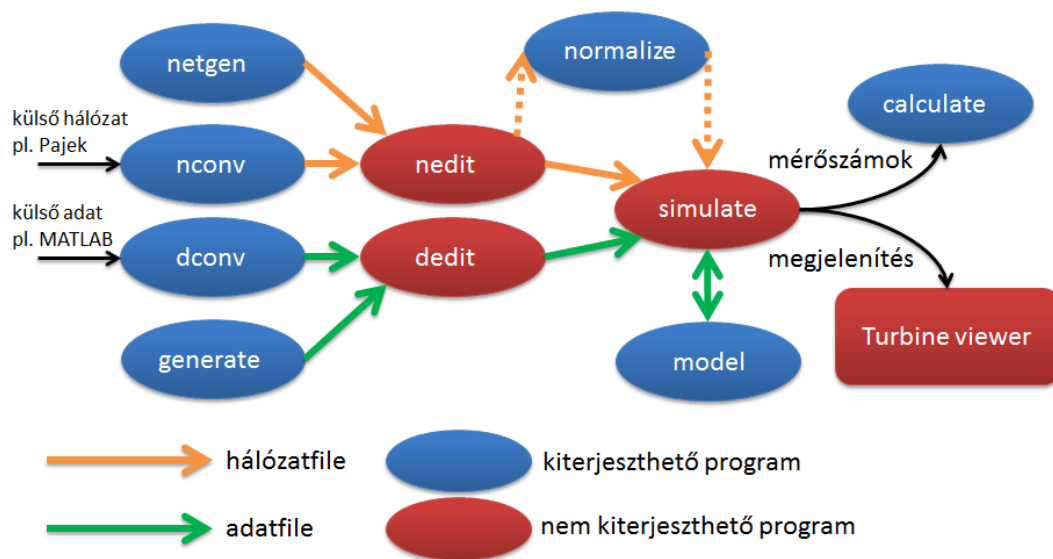
jából a kromoszóma ideális értékeinek felét találta meg, az **átkereszteződés** (12. ábra D panel) során esélyt kapunk az ideális megoldás megtalálására. Átkereszteződés során az új generáció egyedei a kromoszómájuk felét kapják az egyik szülőjüktől, másik felét a másik szülőjüktől. Végül az új populáció néhány egyedének kromoszómáit különböző helyeken véletlenszerűen megváltoztatjuk (**mutáció**, 12. ábra E panel), ezzel olyan új megoldásokat is meg tudunk vizsgálni, melyek az eredeti populációban nem szerepeltek. Ezek után az új generációnak is megvizsgáljuk a fitness-értékeit (melyek remélhetőleg már az előzőeknél jobbak), és ezt az eljárást iteráljuk, amíg egy kielégítő megoldást nem kapunk, vagy már annyi ideje fut a program, hogy úgy döntünk, hogy nem érdemes folytatni. Tapasztalatom szerint tipikusan 1.000 generáció szükséges egy jó eredmény megtalálásához, de ez persze igen sok tényező függvénye, melyek közül a legfontosabb a lehetséges állapotok terének mérete.

3. Célkitűzések

1. Összhangban a hálózat kutatás jelenleg a világ élvonalában lévő egyik legfontosabb célkitűzésével, munkám első célja biológiai hálózatok dinamikájának elemzése volt. Ehhez egy igen gyors, általános hálózatdinamikai elemző programot készítettem el.
2. Doktori munkám második célkitűzése az volt, hogy a hálózatdinamikai elemző programot a dinamikus rendszerek attraktorainak megkeresésére alkalmas modullal bővítssem. Az attraktorkereső modult több hálózaton, köztük biológiai rendszerek jelátviteli hálózatain sikerrel teszteltem zajmentes és zajjal rendelkező körülmények között.
3. Munkám harmadik célkitűzése a hálózatdinamikai programcsomag egy olyan tervező moduljának az előállítás volt, mellyel olyan beavatkozásokat lehet tervezni, amelyek a komplex rendszereket egyik attraktorukból egy általunk meghatározott másik attraktorba (pl. a beteg állapotra jellemző attraktorból az egészséges állapotra jellemző attraktorba) viszik át. Ezt a programcsomag modult rákos sejtekre jellemző jelátviteli hálózatokon teszteltem. A tervező programmal olyan egy-, vagy többcentrumú beavatkozásokat lehet azonosítani, melyek akár személyre szabott gyógyszeres kezelések majdani tervezéséhez is elvezethetnek.

4. Módszerek

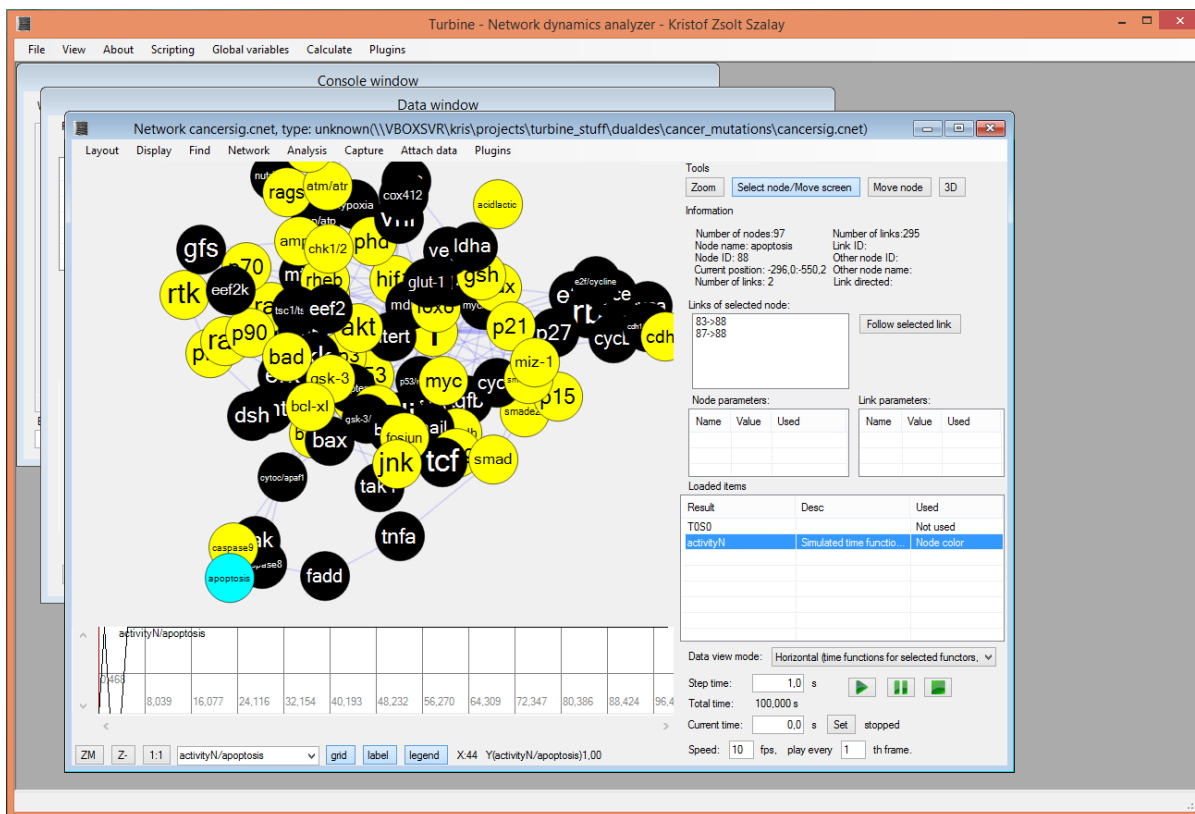
4.1. A Turbine program



13. ábra. A Turbine használatának folyamatábrája. Mind adat, mind hálózatfile-okat készíthetünk procedurálisan, megadott szabályok segítségével (*netgen* és *generate* parancsok), vagy használhatunk külső file-okat (*nconv* és *dconv* parancsok). Ezek további szerkesztését és átalakítását az *netit* és *dedit* programok segítségével végezhetjük. Ezután a *model* megjelöléssel futtatható a szimuláció, mely visszaadja a hálózat szimulált idősorait. Ebből különféle mérőszámok készíthetők (*calculate* program), vagy megtekinthetők vizuálisan (*viewer* program).

A vizsgálatokat a saját fejlesztésű, Turbine-nak nevezett programmal készítettem (a név a perTURBatIoN Analysis szókapcsolatból származik, a kiemelt „turbina” szó angolításával). Jelenleg nincs elérhető hasonló tudású program, ezzel kapcsolatban P1300737 számon tettünk is egy magyar szabadalmi bejelentést, amelyet a közeljövőben tervezünk nemzetközi (PCT) szintre emelni.

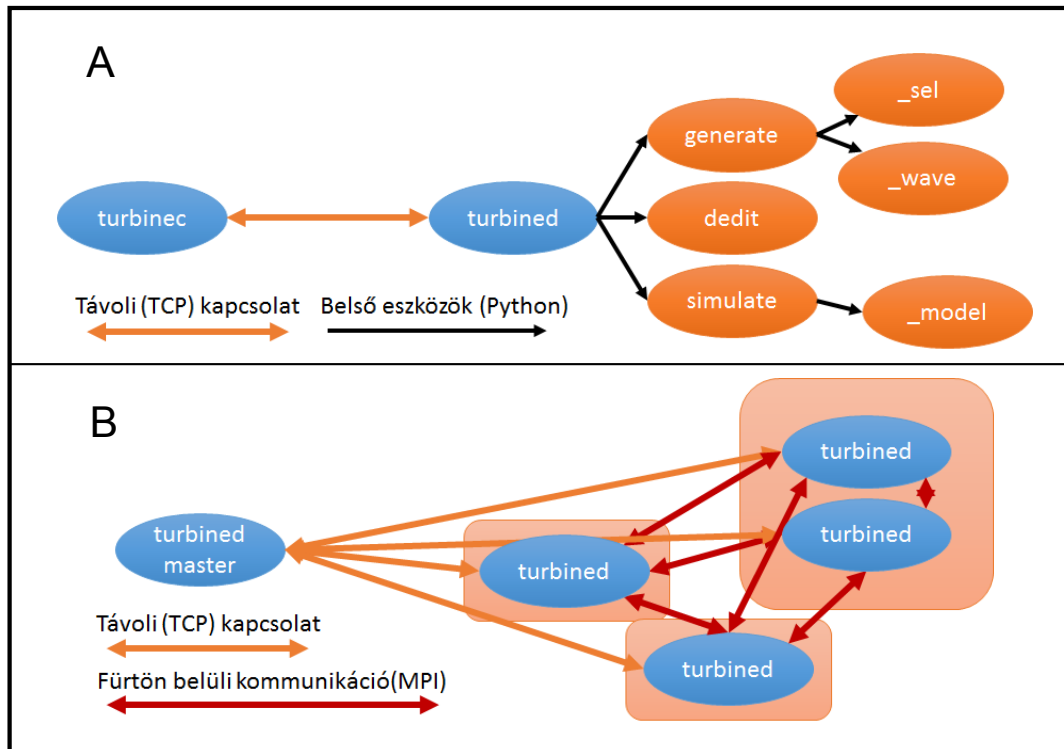
A fejlesztés során a program sebessége és rugalmassága volt a legfontosabb szempont, a program különleges lehetőségei is ebből a két tulajdonságából adódnak. Maga a program a sebességre való optimalizáció okán C++ nyelven készült, és igen moduláris, a különféle alapeladatok, mint adatok/hálózatok átalakítása, a szimulátor, a külön-



14. ábra. A Turbine megjelenítő alkalmazása. Az ábrán A Fumiã és Martins-féle jelátviteli hálózat [24] egy szimulációs idősora van megjelenítve, a dolgozatban szereplő képekben alkalmazott megjelenítési beállítás szerint (sárga: aktív fehérje, fekete: inaktív fehérje). Egy hálózaton több adatsor is megjeleníthető egyszerre, különböző színcsatornákon (piros, zöld, kék), de a pontra írt szöveget, a pont méretét, sőt az elhelyezését (X, Y, Z tengely) is köthetjük tetszőleges adatsorhoz, így egyszerre 8 különböző dimenziót tudunk ábrázolni. A hálózat 3 dimenziós, forgatható megjelenítésre is van lehetőség, valamint az idősorok a zöld "lejátszás" gomb segítségével filmszerűen is megtekinthetők. Az ábra alsó részén a ciánkéssel kijelölt "apoptosis" pont időfüggvénye látható.

féle mérőszámok számítása és maga a szimulátor által használt dinamika is külön-külön DLL file-okban található. Ennek előnye, hogy a program az aktuális vizsgálat igényeivel könnyen kiegészíthető, új formátumok kezelése, új mérőszámok, vagy eddig nem használt dinamikai leírás gyorsan hozzáadható. Az egyes eszközök használatának folyamatábráját mutatja a 13. ábra.

Az iterált szimulációk során további rugalmasságot, a vizsgálatok teljes újrendezésének lehetőségét adja a beépített Python értelmező, melynek segítségével az Eredmények fejezetben leírt bonyolultabb vizsgálatok, mint a kettős tervezés vagy az attraktortervezés is kivitelezhetővé váltak. netgen A program mellé készítettem továbbá egy C[#] nyelven írt grafikus felületet, mely a hálózatok és az adatok vizualizációjára szolgál, és képen



15. ábra. A Turbine kommunikációs mechanizmusai. A: A Turbine szerver (*turbined*) elindítása után azzal TCP kapcsolaton keresztül tudunk kommunikálni, és utasítani azt egyes eszközök futtatására, vagy a Python értelmező segítségével az egyes eszközöket egymás után láncolva bonyolultabb munkafolyamatokat is kialakíthatunk, mint a dolgozatban szereplő attraktorkeresés és beavatkozás-tervezés. Az alsó vonallal kezdődő névvel jelölt pontok Turbine pluginek. B: A Turbine szerverek egymással is képesek kommunikálni, ezért az elvégzendő feladatok (pl. egy populáció egyes szimulációinak futtatása) az elérhető számítógépek között szétszathatóak, ez közvetlenül a mester és a szolgagépek közötti TCP kapcsolaton történik (narancssárga nyilak). Amennyiben egy nagy (milliárdos vagy nagyobb elemszámú) hálózat szimulációját szeretnénk futtatni, a hálózat széttördelhető kisebb elemekre, és a szimulációt a kijelölt számítógépfürt gépei a fürtön belüli, gyors (pl. Infiniband) hálózaton keresztül, MPI protokoll használatával valósítják meg. A narancssárga dobozok mutatják, hogy egy gépen akár több Turbine szerver is futhat, amennyiben az jobb erőforrás-kihasználást eredményez.

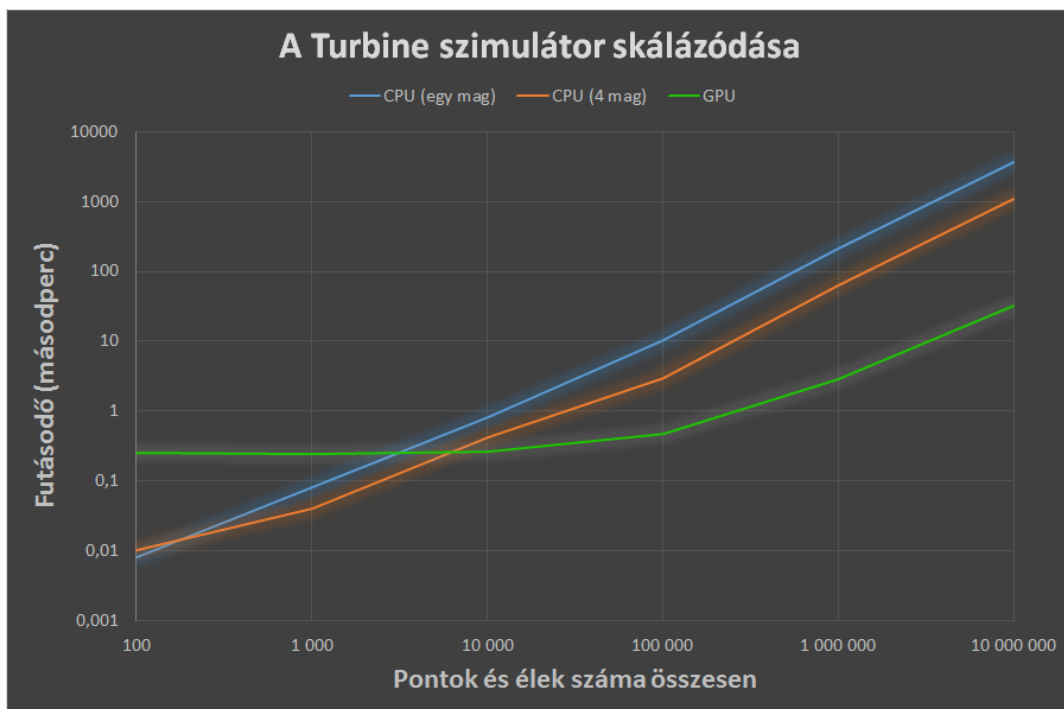
vagy filmszerűen sokrétűen állítható megjelenítési sablon szerint képes egy vagy több szimulációs adatsor visszajátszására, azok összehasonlítására, elemzésére, és exportálására többfajta kép-, és videoformátumban (14. ábra).

A program jelenlegi változata kliens-szerver formátumú, azaz egy **turbined** (turbine daemon, Turbine szerver) nevű folyamat futtatja az egyes alprogramokat, mellyel a beépített Python értelmezőn vagy közvetlen TCP kapcsolaton keresztül vezérelhetünk (15. ábra A panel). A TCP kapcsolatnak köszönhetően a Turbine szerver az Interneten keresztül is

hozzáférhetővé, és más Turbine szerverek által vezérelhetővé tehető, így hosszabb vizsgálsorozatokot szétszthatunk tetszőleges mennyiségű és képességű számítógép között (15. ábra B panel). A Turbine pillanatnyilag körülbelül 60.000 sornyi kódból áll.

4.1.1. Szimulátor

A program lelke egy gyors hálózatszimulátor, mely kevesebb, mint 30 másodperc alatt képes egy összesen 10 millió pontból és élből álló hálózaton 1000 lépés kiszámítására (16. ábra). Kisebb, pár száz vagy ezer elemű hálózatokon egy hasonló szimuláció a másodperc töredékét veszi csak igénybe. Mindez igen gyors működést tesz lehetővé annak tekintetében, hogy a teljes humán proteom (az alternatív splicinggal létrejövő különféle fehérjéket külön számolva) becslések szerint [52] 40.000-70.000 elemből áll.



16. ábra. A Turbine szimulátor skálázódása. Az ábrán kék vonal mutatja a mért futásidőt egy processzormag, a narancssárga a négy processzormag, végül a zöld a program videokártyán történő futtatása esetén mért sebességét különböző méretű hálózatoknál. Látható, hogy a program igen nagy hálózatok szimulációjára is képes egy asztali gépen is.

Egy szimuláció futtatásához meg kell adni a szimulálandó hálózatot, a pontok állapotváltozóinak **kezdőértékeit** (start vector), a **lépéshosszt**, a szimulációs lépések számát,

a használt **dinamikai modellt**, és az esetleges, szimuláció közben alkalmazandó külső hatásokat, **perturbációkat** (perturbation function).

A hálózatot a program egy saját, bináris, sebességre optimalizált CNET formátumban tárolja, természetesen több gyakran használt formátumba (Pajek, GML, adjacencia mátrix, GEXF és egyszerű szöveg) történő konverziós lehetőség biztosításával. A perturbációkat, a kezdőértékeket és az eredményként létrejövő idősorokat hasonlóképp egy CDAT formátumban tárolja a program, melyet Matlab és szöveges formátumra tud alakítani szükség esetén.

A program e tekintetben is igen rugalmas: a szimuláció során nem csak az állapotváltozók, hanem maga a hálózat struktúrája is változhat, új pontok, élek hozhatóak létre, illetve távolíthatóak el a hálózatból; a perturbációk bármely időpillanatban, bármilyen hullámformával hozzáadhatóak, külön-külön, vagy egyszerre tetszőleges számú, és a szimuláció eredményeképpen az összes változó összes idősora rögzítésre kerül. Nagyobb hálózat esetében könnyen elképzelhető, hogy az idősorok nem férnek bele a számítógép memóriájába, így a Turbine további speciális képessége, hogy az idősorokat közvetlenül a merevlemezről képes kezelni, a memóriába történő beolvasás nélkül. Mivel az idősorok szimuláció közben sorfolytonosan íródnak, az ez által eredményezett sebességcsökkenés nem észrevehető.

A program segítségével szükség esetén megoldható rendkívül nagy (milliárd pont) méretű hálózatok elemzése is szuperszámítógép segítségével. Ebben az esetben a hálózat a felhasználó által megadott csoportok szerint (melyek célszerűen a hálózat egyes moduljai) szétdarabolásra kerül, és a szuperszámítógép egyes számítási egységei a rájuk jutó hálózatdarabot szimulálják, majd a más modulban található szomszédjaik (halo, glória) állapotváltozóit MPI protokollon keresztül frissítik minden lépésben.

4.1.2. Iterált szimulációk

Annak felismerésével, hogy ezres-tízezres elemszámú hálózatokon egy másodperc vagy kevesebb idő alatt elvégezhető egy szimuláció, elérhetővé vált szimulációk sokaságának futtatása egy adott hálózaton, amely rengeteg új és hasznos felhasználási lehetőséget nyitott meg számunkra, melyeket a következőkben ismertetek.

4.1.2.1. Attraktorkeresés

Amennyiben úgy futtatunk egy szimulációsorozatot, hogy a pontok kezdőértékeit véletlenszerűen változtatjuk, feltérképezhetjük a rendszer attraktorait, azaz állandósult állapotait. Ezek ismerete több szempontból is igen hasznos lehet. Az irodalomban már több helyen megmutatták [24, 49, 53], hogy a jelátviteli rendszerek attraktorai megfeleltethetők azok különböző lehetséges fiziológias vagy patológias állapotainak. A daganatos elváltozásokat, és az egyéb komplex betegségeket is hasonlóképp többször kötötték már a rendszer attraktoraihoz [53, 54].

A Turbine támogat attraktorkeresést, és képes felismerni nem csak a fixpont-attraktorokat, hanem a határciklusokat és egyes határtóruszokat is, tetszőleges folytonos vagy diszkrét dinamika és külső állandó vagy periodikus gerjesztések használatával is.

4.1.2.2. Beavatkozás-tervezés

Amennyiben nem csak "vakon", véletlenszerűen indítva futtatunk kísérleteket, hanem a szimulációs feltételek körönkénti beállítását egy mesterségesintelligencia-algoritmusra bízunk, úgy további rendkívül hasznos vizsgálati lehetőségek válnak elérhetővé. Jó példa erre a Turbine beavatkozás-tervezési algoritmus. A beavatkozás-tervező algoritmus képes arra, hogy egy adott "beteg" kezdőállapotot és egy adott "egészséges" célállapotot megadva megállapítsa, hogy mi az a minimális (illetve közel-minimális) beavatkozás-sorozat, melynek hatására a rendszer a beteg állapotból az egészséges állapotba vihető át. A szimulációkat vezérlő mesterséges intelligencia ilyenkor egy genetikus algoritmus, mely az alkalmazott beavatkozásokat természetes kiválasztódás, átkereszteződés és mutáció útján optimalizálva keres olyan beavatkozás-sorozatokat, mely a szimulációk során a kezdőállapotot a lehető legközelebb viszik a célállapothoz.

A Turbine beavatkozás-tervező algoritmusának paraméterei igen rugalmasan állíthatók a rendszer típusának és a felhasználó igényeinek megfelelően, és természetesen itt is bármilyen dinamika használható.

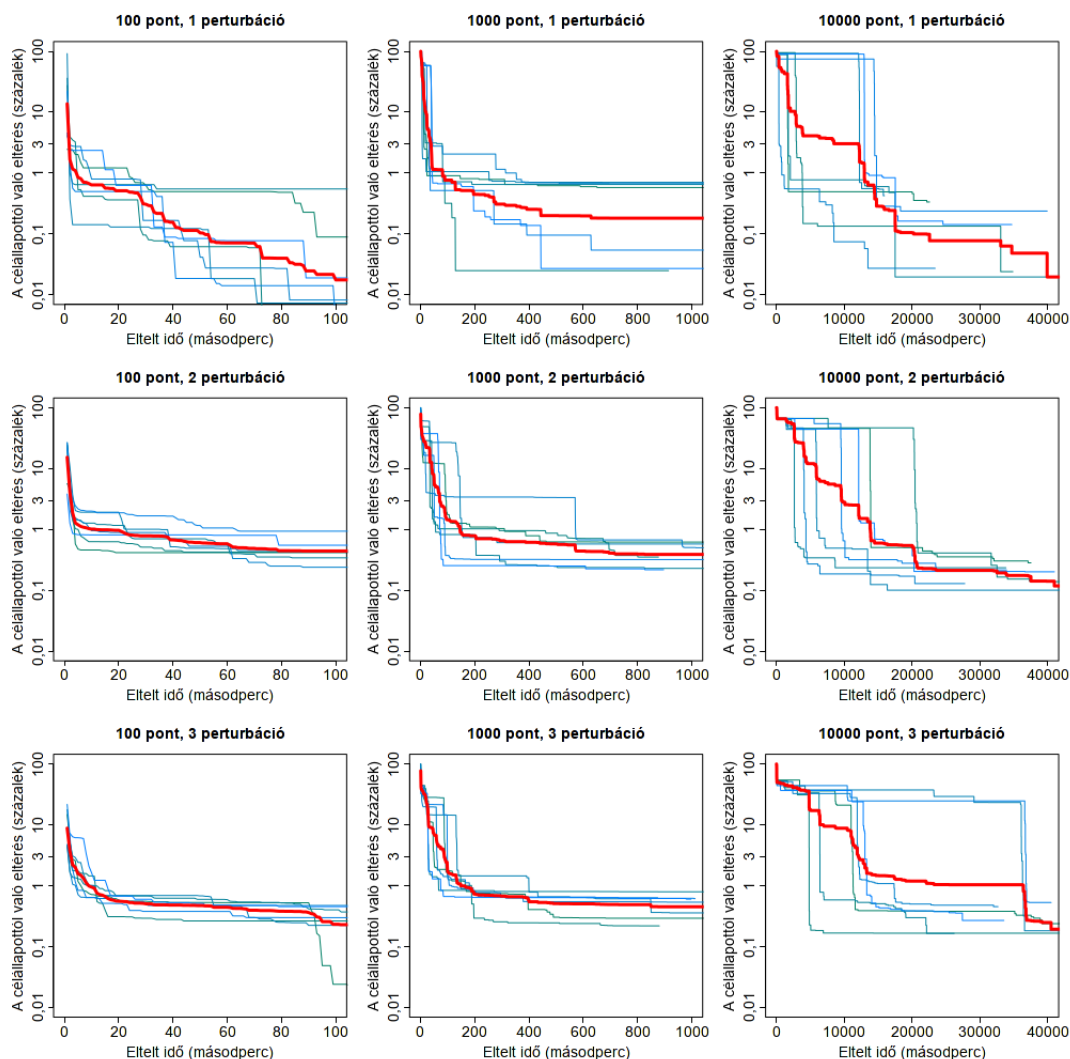
4.1.2.3. Attraktortervezés

Mivel az attraktorkereső és beavatkozás-tervező algoritmusok a Turbine tervezésének megfelelően az egyes alapelemek futását a beépített Python értelmező segítségével hangolják össze, ezért a paraméterezésen túl maguk a folyamatok is könnyen átalakíthatóak, módosíthatóak az igényeknek megfelelően. Megoldható például a két algoritmus összevonásával egy attraktortervező munkafolyamat elkészítése, ahol a rendszer egyes, mesterséges intelligencia által vezérelt módosításai után nem egy szimuláció, hanem egy teljes attraktorkeresés fut le, így meghatározható, hogy milyen módon kell módosítanunk egy rendszert, hogy egy számunkra érdekes attraktor abban megjelenjen, vagy abból egy vagy több káros attraktor eltűnjön.

4.1.2.4. Sebesség és pontosság

A beavatkozás-tervezés új algoritmus lévén nem érdektelen annak a meghatározása, hogy általános esetben milyen körülmények között, azaz mekkora hálózatok esetében alkalmazható, milyen futásidőre és találati pontosságra számíthatunk.

A vizsgálatot Barabási—Albert-féle skálafüggetlen modell-hálózatokon végeztem. Három különböző méretű – 100, 1000 és 10.000 pont – véletlen hálózatot készítettem $k = 6$ csatlakozási értékkel, mindegyikből 21 darabot. A 21-ből 7 hálózatra egy, 7-re kettő, a maradék 7-re pedig három véletlenszerűen kiválasztott pontban Dirac-delta (egységugrás) típusú perturbációt tettem, minden esetben véletlenszerűen $50 \cdot n$ és $100 \cdot n$ közötti méretben (ahol n a hálózat pontjainak száma), és a 0. és a 20. lépés közötti véletlenszerű időpontban, majd futtattam 30 lépés szimulációt az így a hálózathoz adott energia terjedésének kiszámítására. A szimulációhoz a következőkben ismertetett közlekedőedénydinamikát alkalmaztam. A hálózat állapotát a 30. lépésben kimentettem egy file-ba, és ezt adtam meg a beavatkozás-tervező algoritmus célállapotának. A kérdés tehát az volt, hogy az algoritmus különböző méretű hálózatokban, különböző bonyolultságú zavarok esetén milyen gyorsan és milyen pontossággal tudja megtalálni az eredeti, célállapotnak megadott állapotot előidéző gerjesztéseket. Az algoritmus egy gépen, egy Intel Core i7-3630QM processzor egy szálán futott. Az eredményeket a 17. ábra mutatja. Az ábra segítségével megállapítható, hogy az algoritmus mindig 99% feletti pontossággal



17. ábra. A Turbine beavatkozás-tervező skálázódása. Három különböző méretű hálózaton egy, kettő és három teszt-perturbáció megtalálásának sebességét mértük, minden esetben 7 különböző alkalommal. A kék vonalak mutatják az eredeti futásidőket, a piros vonal a csoportos átlagot (logaritmikus átlag). Az ábra alapján gyakorlatban is négyzetes skálázódás várható a programtól, így egy komolyabb számítógépen a teljes proteom esetében történő beavatkozás-tervezés is valószínűleg működőképes lesz.

teljesített, átlagosan 99,5% körül, illetve az algoritmus sebessége körülbelül négyzetesen skálázódik, azaz ha a hálózat pontjainak számát kétszeresére emeljük, az algoritmus futásideje négyszeresére nő. Megnyugtató, hogy a program minden esetben megtalálta az összes perturbációhoz a megfelelő pontot, és igen jól megközelítette annak méretét és a megfelelő időlépést is. Az algoritmus sebessége is kielégítő, egy számítógép egy szálán egy 10.000 pontból álló hálózatba történő beavatkozás tervezése körülbelül 6 órát vesz igénybe (a nagyobb elemszámnál való szórás az együttes futtatás által a merevle-

mez megnövekedett várakozási idejéből adódik). Ezért egy mai, többmagos processzoron akár proteom-méretű hálózatok beavatkozásai is reális idő alatt számíthatók (50.000 pont-ra $\frac{6 \cdot (\frac{50.000}{10.000})^2}{8} = 18,75$ óra nyolcmagos gép esetében), szuperszámítógépen pedig milliós nagyságrendű hálózatok esetében is működik majd.

4.2. Alkalmazott hálózatok

4.2.1. Metionil-tRNS szintetáz fehérje szerkezet

A dolgozatban bemutatott *E. coli* metionil-tRNS szintetáz (MetRS) fehérjeszerkezet-hálózat Amit Ghosh és Saraswathi Vishveshwara 2007-ben publikált munkája [18]. A fehérje szabad, csak Met-et kötő formájának kristályszerkezete ismert, az több PDB-kód (Protein Data Bank) alatt is elérhető [55, 56]. A tRNS-t kötő (zárt) forma kristályszerkezete ellenben ismeretlen *E. coli*-ban, azt a kutatók az *Aquifex aeolicus* ismert tRNS-t kötő MetRS szerkezete alapján (PDB: 2CSX) számítógépes szimulációval alkották meg. A használt nyílt és szimulált zárt konformációk PDB filejait a kutatóktól közvetlenül kaptuk meg. Ezekből a file-okból a fehérjeszerkezet-hálózatokat a RINerator program segítségével készítettem, az alapbeállításokon nem változtatva. A kapott kölcsönhatások erejének mindenhol az abszolútértékét vettem (mivel számunkra, azaz az információterjesztés szempontjából egyformán kezelhető a vonzás és a taszítás). Az így létrejött hálózatok 547 pontot, illetve a nyílt konformáció esetén 15.439, a zárt konformáció esetén pedig 15.301 irányítatlan élt tartalmaznak.

4.2.2. Citokróm P450 2B4 fehérje szerkezet

A másik fehérjeszerkezet-hálózat-pár, amelyet vizsgáltunk, az a házinyúl citokróm P450 2B4 enzimének szabad (PDB: 1PO5) és imidazolt kötő (PDB: 1SUO) formája volt. A letöltött PDB file-okat itt is hasonlóképp, RINerator segítségével, alapbeállításokkal konvertáltam fehérjeszerkezet-hálózattá, majd a kölcsönhatások erejének az abszolútértékét vettem. Az eredményül kapott hálózatok 466 illetve 467 pontból állnak (a zárt konformációban a hem-kötés is szerepelt), valamint 12.746 és 13.971 irányítatlan élből.

4.2.3. Élesztő fehérje-fehérje kölcsönhatás

Az információterjesztési kísérletekben használt élesztő (*S. cerevisiae*) fehérje-fehérje interakciós hálózat Mihalik Ágoston és Csermely Péter [57] munkája. A hálózat alapja a Database of Interacting Proteins 2005-ös változata, melynek különböző stresszeknek kitett, élsúlyaikban különböző változatait a nyugalmi mRNS szintek [58] és azok adott típusú stresszekben tapasztalható megváltozásainak [59] felhasználásával állították össze. A nyugalmi hálózaton kívül egy hő sokknak, egy ozmotikus stressznek illetve egy oxidatív stressznek kitett hálózattal dolgoztunk, mindegyik 2444 pontból és 6271 irányítatlan, súlyozott élből állt.

4.2.4. T-LGL jelátvitel

Az első használt jelátviteli hálózatot Zhang és mtsai. [23] állították össze. A hálózat 60 pontból áll, és eredeti változatában nem tartalmaz éleket, helyettük az aktiváció pontos módját leíró Boole-szabályokat adták meg. A szabályokból a pont-pont kölcsönhatásokat visszaalakítva a hálózat 196 irányított éllel rendelkezett.

4.2.5. Humán daganatos jelátvitel

A daganatos folyamatokban érintett fehérjéket tartalmazó jelátviteli folyamatokat leíró hálózatot Fumiã és Martins [24] készítette. A hálózat eredetileg 96 pontból és 249 irányított, súlyozott élből állt. A kutatók egy küszöbölt folytonos, de lényegében Boole-típusú dinamikát használtak, melyet a Turbine folytonos jelátviteli modelljének segítségével oly módon reprodukáltam, hogy a hálózatba betettem "1" megnevezésű pontot, mely minden kísérletben, folyamatosan, végig aktív volt, és a hálózatban csatlakoztattam a nem 0 küszöbértékkel rendelkező pontokhoz. Ezen élek súlyát az eredeti küszöbérték mínusz egyszeresére állítottam. Ez a változtatás teljesen reprodukálta az eredeti dinamikát a folytonos modell segítségével. Az így kiegészített hálózat 97 ponttal és 295 éllel rendelkezett.

A hálózat továbbá rendelkezett öt bemeneti ponttal, melyekkel a sejt mikrokozonyzetét lehetett beállítani. A következő öt bemeneti pontja van a hálózatnak: "mutagens" (mutagének, genotoxikus stressz), "GFS" (growth factors, növekedési faktorok), "nutrients" (tápanyagellátás), "TNF- α " és "hypoxia". Valójában egy hatodik bemenetet is

találtunk, "GLI" (glioblasztóma-asszociált fehérje, a Hedgehog jelátviteli útvonal effektor [60, 61]) néven, de azt a kutatók az eredeti cikkben nem használták, következésképpen mi is figyelmen kívül hagytuk. A mikrokörnyezet beállítását a kísérletben egy "vonalkód-szerű" jelöléssel mutatjuk be az eredeti cikkhez hasonlóan, mely a bemerített pontok állapotát mutatja a fenti felsorolásnak megfelelő rendezett sorrendben. Így az 11001 érték egy genotoxikus, növekedési faktorok jelenlétében levő, tápanyag és TNF- α nélküli, hipoxiás környezetet ír le. A proliferáló sejtek leírásához az 11100 meghatározású környezetet használtuk, azaz DNS károsodás, növekedési faktorok és tápanyagok jelenlétében, viszont TNF- α nélkül és normoxiában. Ebben a környezetben ugyanis a rendszernek két attraktora is van, egy apoptotikus (ilyenkor ez a normális viselkedés) és egy proliferáló (ami hibás válasz az adott környezetre). Az egészséges sejtekre gyakorolt hatást ellenőrző vizsgálatot a 00000 natív környezetben végeztük, ekkor sem apoptózist, sem proliferációt nem tapasztaltunk. Ha a kezelés hatására mégis tapasztaltuk valamelyiket, azt az adott kezelés toxikus hatására utaló jelzésnek vettük.

4.3. Dinamikus modellek

4.3.1. Közlekedőedények-modell és terjesztési központiság

A közlekedőedények-modell irányított és irányítatlan hálózatban is működik, arra az esetre dolgoztam ki, amikor a rendszerhez (adatok hiányában, vagy a rendszer fajtájából adódóan) nem tudunk egyértelmű, specifikus dinamikai működést rendelni, de mégis szeretnénk annak perturbációk hatására történő viselkedéséről egy durva közelítést mondani. A modell egy viszonylag általános, információ/energiaterjesztést végrehajtó modell, melynek alapötlete a közlekedőedények mozgásából származik. A hálózatban minden pontot egy tartálynak veszünk, melyeket az élek mint csövek kötnek össze, az élsúlytól függő átmérővel. A rendszer dinamikája a tartályokba öntött víz kiegyenlítődési folyamatának egy egyszerűsített (lineáris) változata, azzal a plusz tulajdonsággal, hogy az energia disszipálódik, tehát minden "tartályból" minden lépésben egy konstans mennyiségű víz "elpárolog" – amennyiben van benne. Ennek megfelelően az energia a hálózatból való teljes disszipációjának ideje (**elcsendesedési idő**) a terjedési folyamat során érintett pontok

számától függ. Ezen kívül mind időben, mind térben a lecsengési sebesség (ha a pontok egymás áramlását nem akadályozzák) exponenciális, ami a fehérjék szabadkoncentráció-perturbációjának [62], illetve az elektronikai rendszerekben történő jelek lecsengésének [63] is megfelelő függvényalak. Gátolt terjedés esetén – amikor a pontokban közel hasonló mennyiségű víz van, mert valahol, pl. modulhatáron elakad a terjedés – a lecsengés jóval lassabb, lineárisabbá válik, ezért ily módon a csoportképződés által létrejött hatás is észlelhető. A modell egyenlete a következő:

$$\frac{dS}{dt} = - \sum_{i=0}^l \left(\frac{S - S_i}{2} \omega_i \right) - D_0$$

ahol S a vizsgált pont energiája, l a pont kapcsolatai, S_i az i él túloldalán található pont energiája, ω_i az i él súlya, D_0 pedig a konstans disszipációs tényező.

Fontos szempont, hogy szimuláció során az élsúlyokra a

$$0 \leq \sum_{i=0}^l \Delta t \cdot \omega_i < 1$$

egyenlőtlenségnek teljesülnie kell, különben a rendszer destabilizálódhat, ahol Δt a választott lépésidő. A programban található egy eszköz (*normalize mflow*), mely az összes élsúlyt azonos értékkel leosztja úgy, hogy a feltétel teljesüljön a választott Δt -re, így kvantitatív módon ugyanaz marad a dinamika, de a rendszer nem destabilizálódik.

Egyes pontok energiaterjesztésének relatív hatékonyságának mérésére dolgoztam ki a **terjesztési központosság** (perturbation centrality, PC) mérőszámot, mely az említett elcsendesedési idő reciproka, azaz annál nagyobb, minél kisebb az elcsendesedési idő, azaz minél több pontot érintett az adott pontból kiinduló perturbáció. Ahhoz, hogy különböző kísérletekben, különböző hálózatokon összehasonlítható eredményeket kapjunk, a terjesztési központosság méréséhez használt perturbáció paramétereit fixálni kell. Ezért egy pont terjesztési központosságát (PC értékét) egy darab, az adott pontból induló, $10 \cdot n$ méretű (ahol n a hálózat pontjainak száma) Dirac-delta típusú gerjesztéssel mérjük a közlekedődény-modell segítségével úgy, hogy a D_0 értéket 1-re állítjuk.

4.3.2. Folytonos jelátviteli modell

Ez a modell kifejezetten irányított hálózatokra, a jelátvitel folyamatának egy egyszerűsített, de nem Boole-típusú modellezésére lett kidolgozva, így a fehérjék aktivitásának értéke 0 és 1 közötti értékeket is felvehet. A modell egyenlete a következő:

$$\frac{dA}{dt} = \sum_{i=0}^{l_i} A_i \cdot \omega_i - A \cdot D$$

$$0 \leq A \leq 1$$

ahol A a vizsgált pont aktivitása, l_i az adott pontba *bemenő* élek, A_i a vizsgált upstream node aktivitása, ω_i a vizsgált él súlya és D az aktivált fehérje inaktiválódásának (vagy lebomlásának) mértéke. Amennyiben A 0 alá csökkenne, vagy 1 fölé nőne, az érték megfelelően 0-ra vagy 1-re kerül beállításra, ezek a minimális illetve a maximális aktivitásértékek. Az egyenlet lényegében megegyezik a szakirodalomban található jelátviteli rendszerekben alkalmazott modellekkel [24, 64, 65], gátló kapcsolatok modellezésére is alkalmas (negatív élsúlyok segítségével), és ilyen formában a gátlás-gátlása alakú aktivációs mechanizmusok modellezésére is képes.

4.3.3. Boole-modell

A Boole-modell rendszerek csak egyszerű BE-KI állapotok megkülönböztetésére képes, ám azokhoz igen kifinomult szabályok rendelkeznek a következő formában:

$$(CD45 \mid ((TCR \mid IL2RB) \& ! ZAP70)) \& ! Apoptosis$$

amely kifejezés kifejtve (a "!" jel logikai VAGY-nak, a "&" jel logikai ÉS-nek, a "!" pedig logikai NEM-nek felel meg) akkor igaz, ha az "Apoptosis" pont nem aktív ÉS a "CD45" aktív VAGY az "Apoptosis" pont nem aktív ÉS a "ZAP70" pont nem aktív ÉS a "TCR" vagy az "IL2RB" pontok közül legalább az egyik aktív. Egy Boole-modellben minden ponthoz tartozik egy ilyen kifejezés, amely minden körben kiértékelésre kerül, és a kifejezés igazságtartalmától függően aktiválódik az adott fehérje a szimuláció következő körében.

4.3.4. Statisztikai módszerek

A statisztikai elemzéseket az R programcsomag segítségével [66], RStudio környezetben végeztem. A szignifikanciákat normális eloszlás esetén Student-féle kétmintás t-próba, nem normális eloszlás esetén Mann-Whitney U-próba segítségével mértem. Az adatok normalitását Shapiro teszttel vizsgáltam, a korrelációs számításoknál Spearman-féle r értéket adtam meg. A hatás méretét Cohen-féle D érték segítségével mértem. Az élesztő különböző fehérjecsoportjaiban tapasztalható GO funkció dúsulásokat szintén az R programmal, a g:Profiler könyvtár segítségével végeztem, a szignifikanciaszintek meghatározása során Bonferroni korrekció használatával. Az adott kísérletekben elvégzett statisztikai próbákat és eredményeiket a kísérletek mellett jelöltem.

5. Eredmények

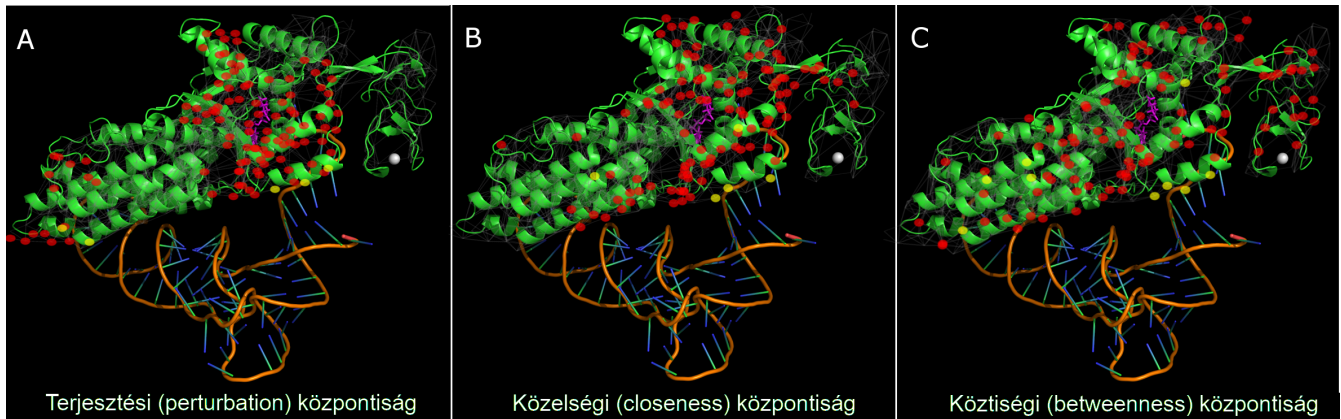
5.1. Általános modellezés

A program fejlesztése során először a közlekedőedények-modell készült el, mely általános esetben is tud egy közelítést adni a rendszer átlagos, perturbációkkal szemben tanúsított viselkedésére. Ezt a modellt, és a hálózat egyes pontjainak perturbációval szemben tanúsított viselkedését kvantitatívan mérni képes terjesztési központiság (perturbation centrality, PC) mérőszámot tehát olyan hálózatokon vizsgáltam, melyekre jelenleg nincs jó dinamikai közelítés. Ilyenek a fehérjeszerkezet-hálózatok és a fehérje-fehérje kölcsönhatás-hálózatok.

5.1.1. tRNS kötőhelyek meghatározása

Először a Ghosh és Vishveshwara által 2007-ben leírt metionil-tRNS szintetáz (MetRS) hálózatok alapján kezdtem részletes vizsgálatba. Az eredeti PDB fileok használatával elkészített fehérjeszerkezet-hálózatban kiszámoltam minden pont terjesztési központiság értékét, majd a legnagyobb központiságú aminosavakat a fehérje PyMOL [67] által készített képén ábrázoltam. Ez a vizsgálat önmagában sem a nyílt, sem a zárt konformáció esetében nem mutatott érdekesnek tűnő eredményt. Következésképp a zárt és a nyílt konformáció terjesztési központiság értékeit kivontam egymásból, és az így kapott értékek felső 20%-át helyeztem rá a fehérje képére. Az eredményt a 18. ábra A táblája mutatja.

Mindhárom táblán piros vagy sárga szín jelöli a felső 20%-ba tartozó pontokat, ezek közül a sárga azokat, amelyek a tRNS-sel fizikailag is kötésben vannak (az aminosav valamely nem-hidrogén atomja és a tRNS valamely nem-hidrogén atomja közötti távolság kisebb, mint 4,5 Å.) Észrevehető, hogy a legfontosabb pontok két csoportban találhatóak, a nagyobb csoport a Met-kötő aktív centrum és a Met-tRNS kötőhely körül, a kisebb csoport pedig a tRNS antikodon hurok környékén. Ezzel a terjesztési központiságok különbségei a két konformáció között valóban sikeresen kijelölték a fehérje működésében



18. ábra. Különbféle központisági mérőszámok által fontosnak jelölt aminosavak az *E. coli* MetRS hálózatában. A vizsgálatokat a Ghosh és Vishveshwara által szimulációval létrehozott [18] zárt, és az ismert nyílt konformációjú MetRS hálózatán végeztem, a zárt konformáció központisági mérőszámaiból kivonva a nyílt konformáció hasonló mérőszámaikat. Piros és sárga pontok ábrázolják a legnagyobb különbségű (felső 20%) központiságokkal rendelkező pontokat. A sárgával jelölt aminosavak ezen kívül a tRNS-el is kötésben vannak, azaz annak valamelyik nem-hidrogén atomja a tRNS valamelyik nem-hidrogén atomjától kevesebb, mint 4,5 Å távolságra található. A terjesztési központiság mérőszám sikeresen azonosította mindkét tRNS-kötőhelyet, a másik két mérőszámmal ellentétben.

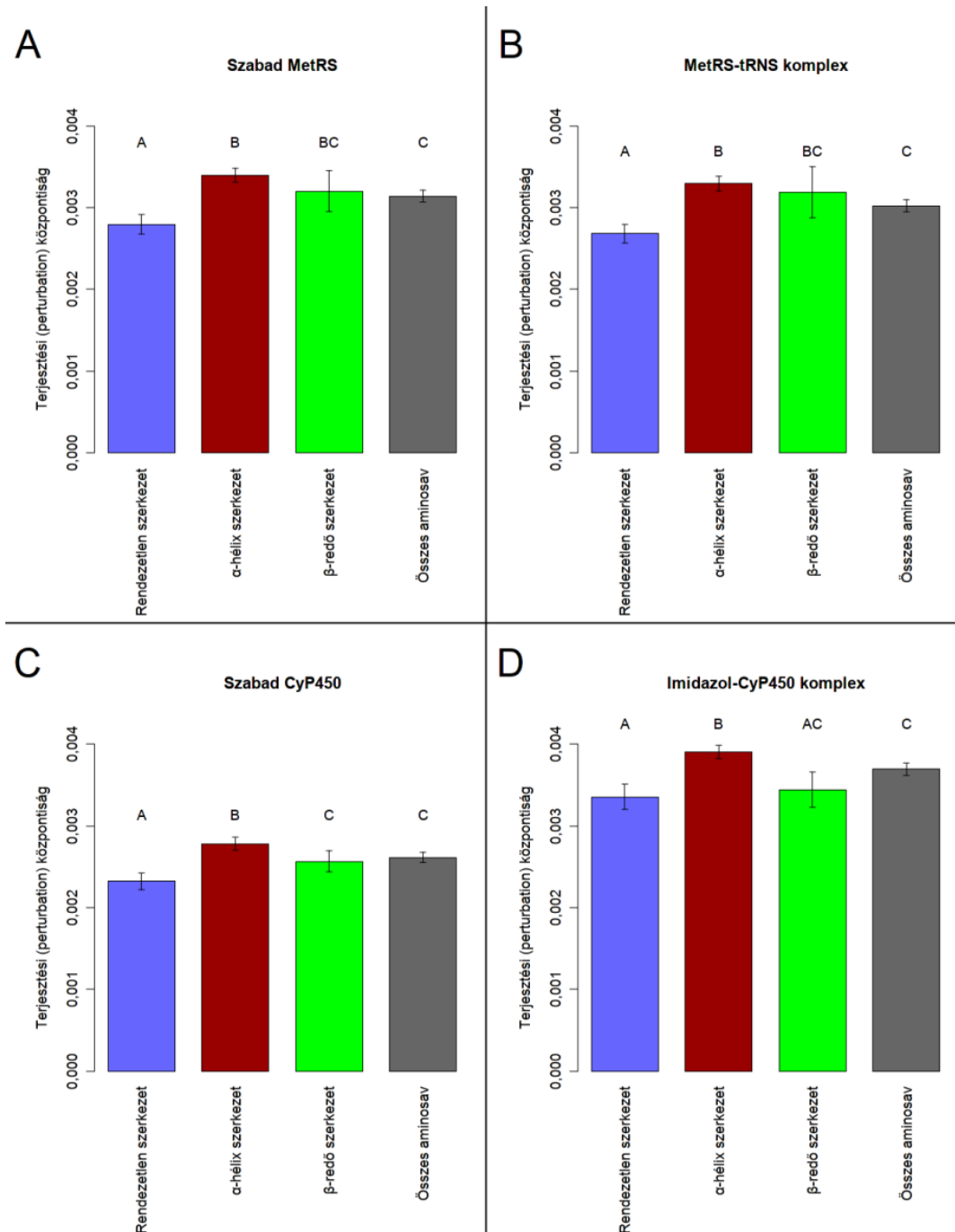
szerepet játszó legfontosabb helyeket. Fontos megemlíteni, hogy a szimuláció során a tRNS maga nem szerepelt a hálózatban, így ezt a program a tRNS kötőhelyek ismerete nélkül, kizárólag a fehérje aminosav-hálózatából számította ki. Tehát a tRNS kötőhely körül oly módon változott meg a fehérje aminosavainak konformációja, hogy a zárt konformációban a fehérje fontos pontjaiban jelentősen könnyebbé vált az információ terjesztése. Annak kiderítéséhez, hogy ehhez mennyire szükséges az általunk bevezetett terjesztési központiság mérőszám, hasonló módon megvizsgáltam a fehérjét a closeness (közelség) és a betweenness (köztiség) mérőszámok használatával is. A closeness értékek változása a B táblán látható, ezen értékek változása a fehérje közepére koncentráldott, illetve a cinket kötő régióra mutat (fehér gömb az ábrán), a betweenness értékek pedig egyenletesen változnak a teljes fehérjében. Ezért kijelenthetjük, hogy ezen fehérje fontos pontjainak kijelölésében az újonnan megalkotott terjesztési központiság mérőszám volt a legsikeresebb.

5.1.2. Másodlagos struktúrák megkülönböztetése

A MetRS vizsgálatainak végzése közben eszünkbe jutott, hogy a terjesztési központiság abszolút értékének meghatározásában (tehát nem a különbségben) a másodlagos szerkezetnek szerepe lehet, hiszen a merevebb, szorosabban kötött szerkezetek elvileg jobban terjesztik a különféle zavarokat. Ehhez a vizsgálathoz hozzávettem még egy fehérjét, a házinyúl citokróm P450 2B4 enzimét (CyP450), melynek nyílt és imidazolt kötő (zárt) konformációjának krisztallográfiás képe ismert. Ezen fehérjék fehérjeszerkezet-hálózatát is elkészítettem, és kiszámítottam mindkét konformáció terjesztési központiság értékeit. A fehérje aktív centruma itt többé-kevésbé a fehérje közepén található, így azt egyszerűbb elemzés is kimutatja, ezért a különbségekben itt nem mutatkozott a dinamikus terjesztési mérőszám előnye. A másodlagos szerkezet és a terjesztési központiság összefüggését mutatja a 19. ábra. Az A tábla a MetRS nyílt konformációján mért adatokat mutatja, a B tábla a MetRS zárt konformációján mérteket; a C tábla a citokróm P450 nyílt, míg a D tábla a citokróm P450 zárt konformációjának adatait. A rendezetlen fehérjeszerkezetek minden alkalommal szignifikánsan³ alacsonyabb átlagos terjesztési központiság értéket mutattak az átlagos értéknél, míg az α -hélix szerkezetben található aminosavak mind a négy vizsgált esetben szignifikánsan¹ magasabbat. A β -redők átlagos terjesztési központiságai a MetRS-ben az átlag és az α -hélixek között helyezkedtek el, egyiktől sem szignifikáns távolságban, míg a citokróm P450-ben átlagos vagy átlag alatti terjesztési központiság értékeket mutattak. Mivel a hatás mérete (Cohen-féle D érték) erre a két megfigyelésre 0,4, illetve 0,285 (ebben a sorrendben), azaz kis-közepes hatást jelent, ezért valószínűleg a másodlagos szerkezetek változása önmagában nem magyarázza meg a MetRS terjesztési központiságainak változását a tRNS kötés során, az egész fehérje megfelelő átrendeződése továbbra is szükségesnek tűnik hozzá.

Ezek a megfigyelések kicsit gyengébb szignifikanciaszinttel reprodukálhatóak voltak closeness értékek segítségével is (C. függelék), de betweenness segítségével csak az α -hélixek átlagnál magasabb értéke volt kimutatható (D. függelék).

³A statisztikai próbák pontos adatai a képaláírásban találhatóak.

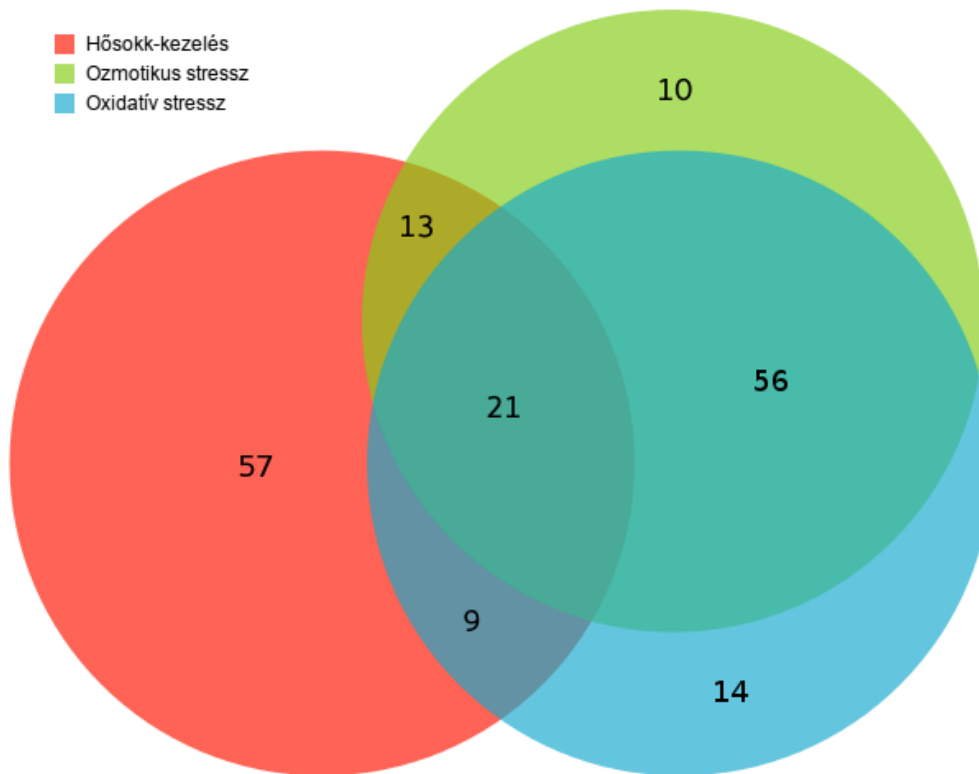


19. ábra. A MetRS és a CyP450 aminosavainak terjesztési központosságai másodlagos szerkezet szerint csoportosítva. Az oszlopok felett található betűk különböző szignifikancia-csoportokat jelölnek, azaz a különböző betűvel jelzett oszlopok közt szignifikáns eltérés volt észlelhető. A rendezetlen szerkezetek terjesztési központosságai az átlagnál szignifikánsan ($p = 3,2 \cdot 10^{-6}$; $9,5 \cdot 10^{-6}$; $2,3 \cdot 10^{-6}$; $0,0001$, a táblák sorrendjében, Mann-Whitney U teszt, $\alpha = 0,00625$ Bonferroni korrekció használatával) alacsonyabbnak bizonyultak, míg az α -hélix szerkezetben levő aminosavak terjesztési központosságai az átlagnál szignifikánsan ($p = 0,00023$; $0,00015$; $0,00083$; $0,0014$, $\alpha = 0,00625$) magasabbnak. Betweenness esetén az α -hélixek nem különböztek szignifikánsan az átlagtól (D. függelék), closeness használatával igen, de a szignifikanciaszintek sokkal gyengébbek voltak (C. függelék). A Cohen-féle D értékek alapján a másodlagos szerkezet által kiváltott hatás kicsi-közepesnek mondható (a D értékek átlaga $0,4$ a rendezetlen szerkezet átlagtól való eltérésére, $0,285$ az α -hélixek átlagtól való eltérésére).

5.1.3. Élesztő stresszben tapasztalható hálózatos funkcióváltozásai

Mivel a terjesztési központiság, mint mérőszám, illetve elsősorban két hasonló hálózat terjesztési központiságainak különbsége jól használható eszköznek bizonyult a fontos változások kimutatásában, megvizsgáltam a módszert egy másik típusú irányítatlan hálózaton, élesztő *S. cerevisiae* fehérje-fehérje kölcsönhatás-hálózatán is. Mihalik és Csermely megfigyelése az volt, hogy az élesztő PPI hálózatainak csoportjai stressz hatására jobban szétválnak, a modulok közötti kapcsolatok gyengülnek [57]. Ezen megfigyelés, egybevetve azzal, hogy a terjesztési központiság értékekre a modulszerkezet hatással lehet, érdekes vizsgálati célponttá tette ezt a hálózats csoportot. A vizsgálatához a forráscikkben ([57]) közölt, különféle típusú (hősokk, oxidatív stressz, ozmotikus stressz) stresszeknek kitett élesztők PPI-jét hasonlítottam össze terjesztési központiság, illetve pontosabban a pontok terjesztési központiság értékeinek nyugalmi állapottól való eltérése szempontjából. Először mind a négy hálózat (a három különböző módon stresszelt és a nyugalmi) minden pontjának terjesztési központiság értékét kiszámítottam. A fehérje összes terjesztési központiságának átlaga a hősokknak kitett hálózatban 40,5%-al alacsonyabbnak adódott ($p = 2,2 \cdot 10^{-16}$, Mann-Whitney U teszt), mint a nyugalmi hálózat átlagos terjesztési központiság-értéke, ami összhangban van az eredeti cikkben szereplő elmélettel, miszerint hősokkban a fehérje-fehérje kölcsönhatás-hálózat moduljai erősebben szétválnak.

Ezután a különféle módon stressznek kitett hálózatok 100 legmagasabb abszolút terjesztési központiság értékű pontját összehasonlítottam (ez a 2444 pontból álló hálózat pontjainak a felső 4%-a). Ezen eredményeket Venn-diagram formában a 20. ábrán ábrázoltam. Az ábra adataiból az látszik, hogy a hősokkban terjesztési központiság érték szerint fontossá váló fehérjék jelentősen különböznek mind az oxidatív, mind az ozmotikus stresszben terjesztési központiság érték szerint legfontosabbá váló fehérjéktől, miközben e két utóbbi stresszfajta jelentős átfedést mutat egymással. Ezen terjesztési központiság szerint fontos fehérjék funkciójának meghatározásához megvizsgáltam, hogy a hálózatban található összes fehérjéhez képest a fontosnak jelölt fehérjékben milyen GO funkciók dúsulását figyelhetjük meg. Az eredményeket az E. függelékben található táblázat mutatja. Mind a nyugalmi, mind a stressznek kitett hálózatokban terjesztési központiság szerint fontos pontok esetében a sejtciklus ("cell cycle", "cell cycle phase", "cell



20. ábra. Fehérjék hálózatos fontosságának eltérő megváltozása különféle típusú stresszekben. Az ábra a különféle típusú stresszeknek kitett élesztő fehérje-fehérje kölcsönhatási hálózatban a terjesztési központosság mérőszám szerint leginkább kiemelkedő 100 pont átfedését, illetve különbségét mutatja. A hősokkban kiemelkedő fehérjék látványosan különböznek a másik két stressztípustól, a 100-ból 21 fehérje azonban mindhárom esetben azonos.

cycle process”), illetve DNS javítás (“DNA repair”, “double-strand break repair”, “response to DNA damage stimulus”), mint funkciók feldúsulása figyelhető meg, azaz az esszenciális funkciók hálózatos fontossága minden esetben olyan erős, hogy az esetleges különbségeket elnyomja. Így megvizsgáltam azon 100 fehérjéből álló csoportokat is, melyek fontossága jelentősen *megnő* vagy *lecsökken* stressz esetén, azaz ismét különbségi vizsgálatot végeztem. Az ezen csoportokban jelentősen dúsuló GO funkciók a különféle stresszek szerint az F. (legnagyobb növekedés stresszben) és a G. (legnagyobb csökkenés stresszben) függelékekben található táblázatokban láthatóak. Ezek a funkciók már sokkal jobban megfelelnek a stresszre jellemző, ismert válaszoknak. Mindhárom fajta stresszben a (terjesztési központosság szerint) fontosabbá váló fehérjék között jelentős dúsulást találtam a “response to stimulus”, “carbohydrate metabolism”, “trehalose metabolic process”,

”glycogen metabolic process” kulcsszavakra, illetve hősokkban és ozmotikus stresszben a ”response to stress” kulcsszó is a legfontosabbak között volt megtalálható. A szénhidrát-metabolizmus ilyen változása élesztőben a szakirodalomban ismert [68]. A stresszben a hálózatos fontosság jelentős csökkenését mutató fehérjék között messze egyértelműen a riboszóma-szintézissel és a transzlációval kapcsolatos kulcsszavak vezetnek, minden stresszben egyaránt. A fehérje-transzláció leállítása stresszben szintén egy jól ismert válaszreakció [69]. Ezek alapján a terjesztési központiság mérőszám ebben az esetben is sikeresen kiemelte az ismert, jelentős stresszválaszokat, tehát a hálózatos központisági mérőszámok által mért fontosság, és annak megváltozása jó összefüggésben állhat az egyes elemek hagyományos értelemben vett fontosságával.

5.2. A jelátvitel specifikus modellezése

Természetesen várható, hogy sokkal pontosabb és jelentősebb predikciókat tudunk adni egy hálózatra, ha arra nem általános, hanem a hálózat specifikus működését leíró modellt tudunk alkalmazni. Ehhez a fehérjék közötti jelátvitel folyamatát választottam, mivel az irodalomban több példát találtunk rá [49, 70], hogy a jelátvitel dinamikájának igen egyszerű, absztrakt leírásával is kísérletileg igazolt predikciókat sikerült tenni.

5.2.1. A jelátviteli modell

A további vizsgálatokhoz kétféle, a szakirodalomban leírt jelátviteli modellt készítettem el a Turbine-hoz: a Zhang és mtsai által leírt T-LGL leukémia modellezésére használt Boole-modellt [23, 49], illetve a Fumiã és Martins által leírt humán daganatos jelátvitel modellezésében használt egyenlet szerint működő modellt [24]. Mindkét modellt megvizsgáltam az eredeti cikkekben szereplő hálózatokon, és úgy találtam, hogy azok az eredeti dinamikai működést az eredeti cikkekben szereplő leírással és eredményekkel összehasonlítva teljes mértékben reprodukálták.

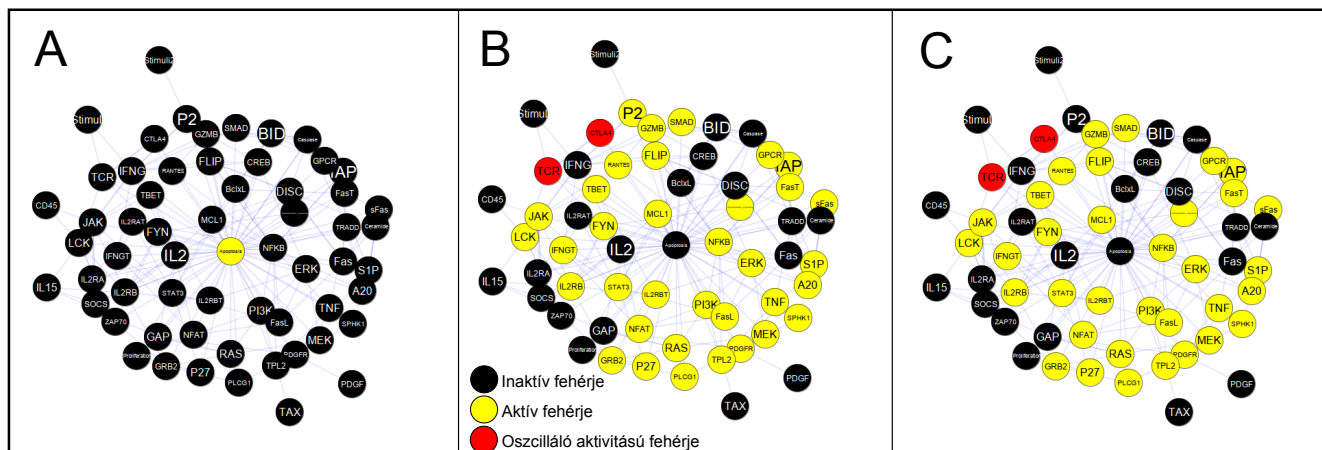
A modellen kívül ahhoz, hogy egy rendszerbe való beavatkozást szimulálni tudjunk, tudnunk kell, hogy a rendszer milyen állapotban van, azaz ismernünk kell az összes pont állapotváltozóinak értékét. Ezen jelátviteli modellekben az állapotváltozók nem az egyes fehérjék koncentrációja, hanem azok *aktivitása*, így a legegyszerűbb megoldás az len-

ne, ha létezne egy olyan kísérletes adatbázis, mely különféle nyugalomban levő sejtek fehérjéinek aktiváltsági állapotát írja le. Ilyen adatbázis azonban nem áll jelenleg rendelkezésre, így azt tudjuk tenni, hogy megvizsgáljuk, hogy a sejt – a hálózat élei által meghatározott kényszerfeltételek hatására – milyen nyugalmi állapotokat vehet fel egyáltalán. Mind a forráscikkekben, mind a mi kísérleteinkben azt tapasztaltuk, hogy ezen lehetséges nyugalmi állapotok – amely nyugalmi állapotok megfelelnek a rendszerelméletben alkalmazott *attraktorok* definíciójának – száma igen korlátozott, ezen hálózatokban jellemzően 3 vagy kevesebb ilyen állapot létezik.

Elviekben azonban igen sok attraktor is lehet, és ahhoz, hogy egyértelműen kimondhassuk, hogy az összes attraktort megtaláltuk, a teljes állapotteret át kell vizsgálni, ami Boole-modellek esetén 2^n (minden pont egymástól függetlenül felvehet BE vagy KI értéket) különféle lehetséges állapotot, folytonos modellek esetén pedig megszámlálhatatlanul végtelen állapotot jelent. Ez 100 pont esetén $1,2 \cdot 10^{30}$ vizsgálatot jelent, tehát Boole-modell esetében "csak" fizikai, folytonos esetben pedig elvi képtelenség is. Minderre azért van szükség, mert az a tény, hogy egy adott állapot attraktor, nem jelenti automatikusan azt is – bár jellemző –, hogy bármely más állapotból elérhető. Elképzelhető, hogy csak akkor találjuk meg az adott nyugalmi állapotot, ha pontosan az a kezdőállapotunk. A szakirodalomban ezért vagy leegyszerűsítik a hálózatot néhány, 10 vagy annál kevesebb pontra, és ezen rendszer lehetséges attraktorait vizsgálják át teljes mértékben [49], vagy lefuttatnak bizonyos mennyiségű, véletlenszerűen indított szimulációt, és felteszik, hogy a szimulációk végeredményei, azaz a szimulációk végeredményeként elért nyugalmi állapotok tartalmazzák az összes jelentős attraktort (a Fumia és Martins által leírt hálózaton [24] az alkotók mind a 32 lehetséges mikrokörnyezet attraktorait mikrokörnyezetenként 1 millió szimuláció lefuttatásával számították ki)⁴.

Mi ez utóbbi módszert használjuk, mivel a program pontosan szimulációk sokaságának futtatására lett kifejlesztve. Nem elhanyagolható kérdés azonban, hogy valóban megtaláljuk, megtalálhatjuk-e így a működés szempontjából fontos attraktorokat.

⁴Amennyiben az egyes differenciálegyenletek integrálhatóak, a rendszer fixpontjai (nem ciklikus attraktorai) megkereshetőek a $\frac{dS_i}{dt} = 0$ feltétellel minden i -re, de ez a módszer a határciklusokat nem találja meg, és a stabilitásról sem ad plusz információt.



21. ábra. A T-LGL hálózat attraktorai. A T-LGL jelátvitelének modellhálózatában az eredeti Boole-szabályok használatával a fent látható három attraktort találtam. Az A táblán látható a legnagyobb, apoptotikus attraktor, a B és C táblákon a proliferatív, LGL leukémiás állapotnak megfelelő attraktorok. Sárga szín jelzi az aktív fehérjéket, fekete az inaktívakat, valamint piros az oszcilláló aktivitású fehérjéket.

5.2.2. Verifikáció a T-LGL modellen

A T-LGL hálózat citotoxikus T-limfociták aktivációjában, és az aktiváció által okozott apoptózisában szerepet játszó fehérjék által alkotott jelátviteli hálózat [23], melyhez az alkotók az összes fehérje aktivációjának feltételeit Boole-szabályként mellékeltek. A rendszer jelentősége, hogy a későbbiekben a hálózat egyszerűsítésével, majd az egyszerűsített hálózat teljes, összes lehetőséget figyelembe vevő elemzésével mind attraktorkeresést, mind beavatkozás-tervezést végeztek rajta [49], ezért a készített mesterségesintelligencia-módszereim működése ezen a rendszeren ellenőrizhető volt.

A nyugalmi T-LGL hálózatban a Turbine programmal 100.000 futtatás során mindössze egyetlen, apoptotikus attraktort találtam. Az eredeti cikknek megfelelően az IL15, Stimuli (antigén stimulus) és PDGFR bemenetek bekapcsolása esetén⁵ két másik, a leukémiás állapotnak megfeleltethető, kóros proliferációt eredményező attraktor is megjelent a rendszerben. Az így meghatározott három attraktort mutatja a 21. ábra. Az ábráról leolvasható, hogy a két proliferatív attraktor mindössze egy pont, a P2⁶ aktivitási szintjében különbözik, ezen kívül teljes mértékben azonos. A proliferatív attraktorokban továbbá az

⁵Ezen választás oka az volt, hogy Zhang és mtsai. [23] megmutatták, hogy ez az a minimális diszreguláció a hálózatban, ami a T-LGL állapot megjelenéséhez vezet.

⁶A P2 egy virtuális pont a hálózatban, mely az interferon- γ transzlációjának interferon- γ függő gátlását mediálja.

1. táblázat. A Turbine által talált beavatkozások T-LGL leukémia ellen. A *gtl.* gátlást, az *akt.* aktivációt jelent. Az eredeti cikkben talált 19 egyedi célpontot a Turbine mind megtalálta, ezen kívül egy olyan egyedi célpontot is (CREB aktiváció), melyet a [49] tanulmány csak valószínűnek jelez. Ezen kívül több kombinációs célpontot is találtunk, melyek gyakran előforduló elemeit (TBET, Fas, sFas, FLIP, IL2) a forráscikkben szintén nagy valószínűségű célpontnak jelölték.

ERK <i>gtl.</i>	TBET <i>gtl.</i> , sFas <i>gtl.</i>
MCL1 <i>gtl.</i>	TBET <i>gtl.</i> , GZMB <i>akt.</i>
NFκB <i>gtl.</i>	sFas <i>gtl.</i> , IL2 <i>akt.</i>
MEK <i>gtl.</i>	FLIP <i>gtl.</i> , sFas <i>gtl.</i>
DISC <i>akt.</i>	Fas <i>akt.</i> , FLIP <i>gtl.</i>
PI3K <i>gtl.</i>	Fas <i>akt.</i> , IL2 <i>akt.</i>
BID <i>akt.</i>	IFNG <i>akt.</i> , TBET <i>gtl.</i>
STAT3 <i>gtl.</i>	PLCG1 <i>gtl.</i> , GRB2 <i>gtl.</i>
Ceramide <i>akt.</i>	TBET <i>gtl.</i> , Fas <i>akt.</i> , LCK <i>akt.</i>
S1P <i>gtl.</i>	TBET <i>gtl.</i> , Fas <i>akt.</i> , SPHK1 <i>akt.</i>
PDGFR <i>gtl.</i>	PLCG1 <i>gtl.</i> , TBET <i>gtl.</i> , Fas <i>akt.</i>
SPHK1 <i>gtl.</i>	TBET <i>gtl.</i> , TNF <i>gtl.</i> , NFAT <i>akt.</i> , Fas <i>akt.</i>
IL2RBT <i>gtl.</i>	FYN <i>akt.</i> , Fas <i>akt.</i> , P2 <i>akt.</i> , TBET <i>gtl.</i>
RAS <i>gtl.</i>	FYN <i>akt.</i> , TCR <i>gtl.</i> , TBET <i>gtl.</i> , Fas <i>akt.</i>
IL2RB <i>gtl.</i>	FYN <i>gtl.</i> , NFAT <i>gtl.</i> , Fas <i>akt.</i> , TBET <i>gtl.</i>
CREB <i>akt.</i>	MEK <i>akt.</i> , TBET <i>gtl.</i> , Fas <i>akt.</i> , Bclx1 <i>akt.</i> ,
JAK <i>gtl.</i>	TBET <i>gtl.</i> , Fas <i>akt.</i> , sFas <i>akt.</i> , IL2RAT <i>gtl.</i> , IAP <i>akt.</i> , TNF <i>gtl.</i>
Caspase <i>akt.</i>	PDGFR <i>akt.</i> , TRADD <i>gtl.</i> , FYN <i>gtl.</i> , RANTES <i>akt.</i> , TBET <i>gtl.</i> , Fas <i>akt.</i> , GPCR <i>akt.</i>
GAP <i>akt.</i>	
Apoptosis <i>akt.</i>	
SOCS <i>akt.</i>	

eredeti cikk 3. ábrájának megfelelően [49] két fehérje aktivációja oszcillál, a TCR és a CTLA4 ponté, ezek piros színnel vannak jelölve. A három attraktorba a szimulációk 99,55%, 0,44% és 0,007%-a futott, ezek közül a legkevesebb abba a proliferatív állapotba érkezett, melyben a P2 pont inaktív.

Mindez teljes mértékben megfelel az eredeti cikkben leírt, az összes lehetséges állapotot vizsgáló eredménynek.

Ezután további vizsgálatként a Turbine beavatkozás-tervező eszközével olyan beavatkozásokat terveztem a rendszerbe, melyek segítségével a leukémiás állapotnak megfelelő attraktorból az egészséges attraktorba lehet mozdítani a rendszert. Ezeket a beavatkozásokat mutatja az 1. táblázat.

A táblázat első oszlopában láthatóak az egyedi célpontok, a másodikban a kombinációs célpontok. Mivel a [49] cikkben csak egyedi beavatkozásokat vizsgáltak, ezért a kombinációs célpontok mind újonnan talált beavatkozási lehetőségek. Az első oszlop 20 beavatkozásából 19 megegyezik a forráscikkben leírt 19 konszenzus beavatkozással, a huszadik (CREB aktiváció) hatásosságát csak valószínűsítik. A kombinációs beavatkozások gyakran előforduló elemei a forrásban használt elemzésben szintén nagy valószínűségű beavatkozási célpontok (sFas, TBET, Fas, FLIP).

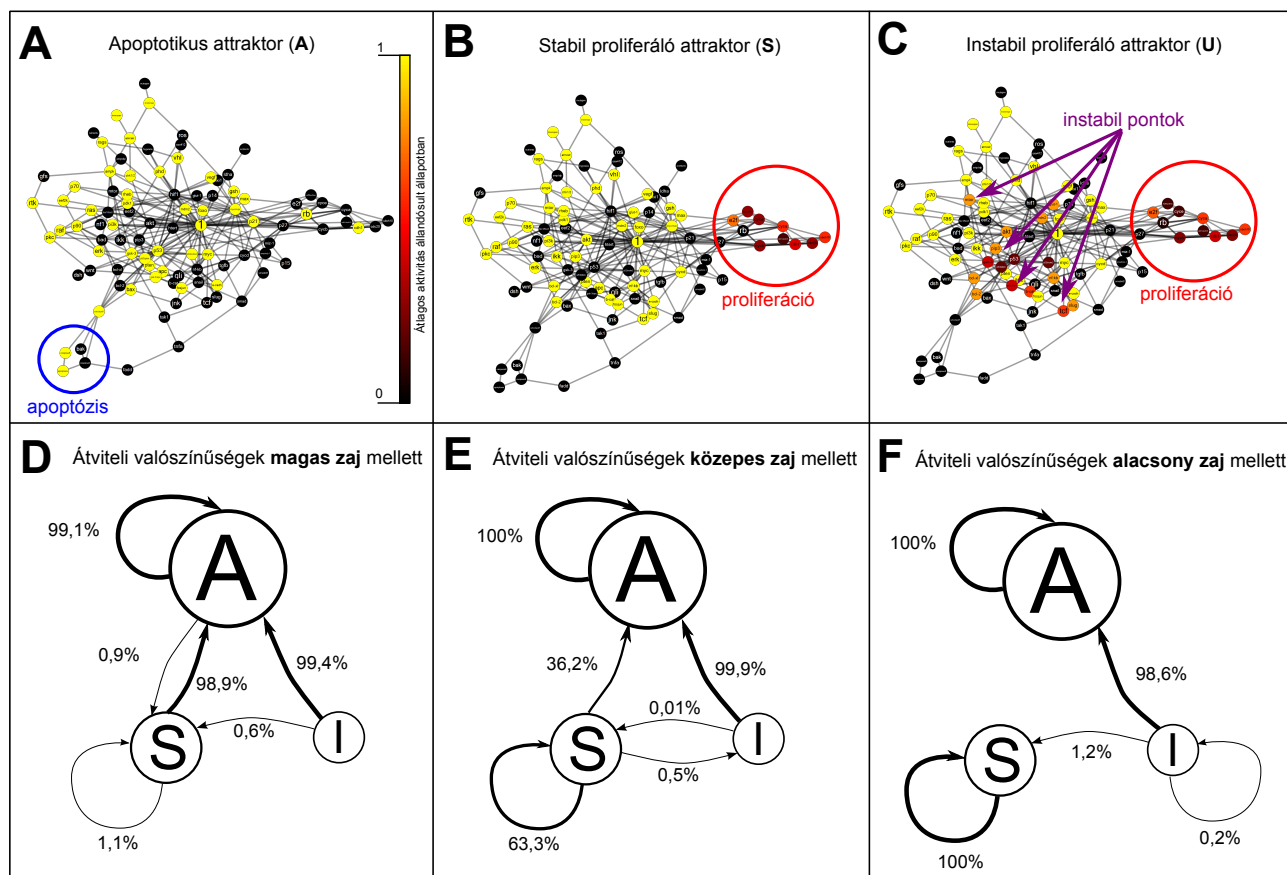
Ezek alapján mondhatjuk, hogy mind a Turbine attraktorkereső, mind a beavatkozástervező algoritmus képes volt reprodukálni az eredeti eredményeket, illetve azokon felül új kombinációs beavatkozásokat javasolni, ezzel igazolva az eszköz működőképességét.

5.2.3. Humán daganatos jelátviteli modell

Az előző eredményeket figyelembe véve arra jutottam, hogy a T-LGL leukémia az előzőleg vizsgált modellhálózatban a valós T-LGL leukémiához képest túl "könnyű" célpont lehet, azaz bár a proliferatív attraktorok jelenléte a standard hálózatban elméleti szempontból érdekes, annak mérete más mutációk nélkül olyan kicsi, hogy egyéb diszregulációk nélkül abba valós körülmények között várhatóan nem is kerül bele a sejt, illetve onnan könnyen kimozdul természetes úton is. Ezt támasztja alá a [71] cikk is, melyben a T-LGL leukémiát a Fas útvonalban található patológias változásokhoz kötik.

Tudjuk, hogy egyetlen kommunikációs módszer sem működik tökéletes pontossággal, bizonyos hibaszázalékkal mindig számolnunk kell. Hipotézisem ezért az volt, hogy a rendszer lényeges stabil állapotainak jó zajtűréssel kell rendelkezniük, azaz az attraktor **medencéjének** akkorának kell lennie, hogy a rendszert érő véletlen zaj ne vihesse át egy másik, az adott körülményeknek nem megfelelő működést eredményező attraktorba a sejt működését. Ilyen apró, véletlen zajok Boole-modellben nem alkalmazhatóak, mivel egy pont hirtelen teljes be-, vagy kikapcsolása igen komoly változás (egy fehérjetípus összes fejrjéjének aktiválása vagy inaktiválása), ennyire erős zaj talán már nem jellemző élő szervezetekben.

Így vizsgálataimat a folytonos modellel vizsgálható, humán daganatokban érintett fehérjék jelátviteli útvonalait ábrázoló hálózattal végeztem, melynek elkészítése Fumiã és Martins munkája [24]. Beállítottam ezért az egyetlen olyan mikrokörnyezetet, ahol a



22. ábra. A Fumiä és Martins által leírt [24] hálózat attraktorai és viselkedésük rózsaszín zaj hatására. A felső sorban látható a három talált attraktor, az A táblán az apoptotikus (kékekkel jelölve a kaszpázokat), a B és C táblán a két proliferáló (pirossal jelölve a ciklineket), illetve a C táblán lilával jelölve az egyéb oszcilláló pontokat, amelyekről az attraktor a nevét kapta. Az alsó sorban láthatóak az átviteli valószínűségek különböző erősségű rózsaszín zaj hozzáadása esetén. Közepes mértékű zajtól már csak az apoptotikus attraktor stabil, alacsony zaj esetén ugyanez még a stabil proliferáló attraktorról is elmondható.

szerzők egy fiziológiás és egy patológiás attraktort is találtak, az 11100 mikrokörnyezetet. Ebben a környezetben a sejt genotoxikus stressz alatt van, növekedési faktorok és tápanyagok jelenlétében, TNF- α jelenléte nélkül, normoxiában. Ilyen körülmények között a sejt fiziológiás válasza az apoptózis, a másik megfigyelhető válasz – a proliferáció – pedig patológiás.

Az attraktor medencéjének – azaz a fázistér egy attraktorba futó állapotainak halmozásának – méretét egyszerűen becsülhetjük szimulációsorozat segítségével. Ehhez véletlenszerűen, egyenletes eloszlással kell kezdőállapotokat generálni, lefuttatni a szimulációkat, végül összeszámolni, hogy melyik attraktorba mennyi kezdőállapot futott be. 10.000 szimuláció futtatásával három attraktort találtunk, egy apoptotikust (22. ábra A táblája),

mely attraktorba a kezdőállapotok 98,3%-a futott be, és két proliferálót (22. ábra B és C táblája), melyekbe a pontok 0,5%-a, és 0,012%-a érkezett. Az attraktor viselkedésének megfelelően az előbbi stabil proliferáló attraktornak, míg a másodikat instabil proliferáló attraktornak neveztük el, mivel az instabil proliferáló attraktorban a proliferációhoz szükséges ciklikus viselkedés mellett máshol is ciklikusan változik a fehérjék aktivitási állapota. Az ábrán sárga szín jelöli az aktív pontokat, fekete az inaktívakat, míg köztes (narancssárga, piros) a ciklikusan változó aktivitású pontokat egy periódus átlagos aktivitási szintje alapján.

Ezután a három különböző attraktort kezdőállapotként beállítva egy olyan kísérletsozort indítottam, melyben mindhárom attraktorra 1000 alkalommal kis (5 lépés ideig $\pm 0,001$ egység), közepes (10 lépés, $\pm 0,1$ egység) és nagy (30 lépés, $\pm 0,4$ egység) rőzsaszín zaj alakú ($\frac{1}{f}$ teljesítménysűrűség-függvényű, ahol f a frekvencia) beavatkozást tettem, és vizsgáltam az attraktor viselkedését. Az eredményeket nagy, közepes és alacsony zaj esetében a 22. ábra D, E és F táblája mutatja, ahol **A** jelöli az apoptotikus attraktort, **S** a stabil proliferálót, **I** pedig az instabil proliferálót. A fiziológias apoptotikus attraktort a méretének megfelelően igen nehéz volt elmozdítani, csak a legnagyobb mértékű zaj esetében sikerült 0,9% eséllyel átvinni a stabil proliferáló attraktorba. Az instabil proliferáló attraktorban levő rendszereket ezzel szemben már a legkisebb zaj is képes volt elmozdítani, nagyrészt (98,6%) az apoptotikus attraktor felé, kisebb részben (1,2%) a stabil proliferáló állapotba. A stabil proliferáló állapot kis zaj esetén még stabilnak bizonyult, közepes és nagyméretű zaj esetében tapasztaltunk csak lehetőséget az attraktor elmozdulására, főként (36,2%) az apoptotikus attraktor felé. Mindebből azt a következtetést vonhatjuk le, hogy folyamatosan fennálló alacsony zaj mellett mind az apoptotikus, mind a stabil proliferatív állapot stabilnak tekinthető, ennél nagyobb, folyamatosan fennálló zaj esetében csak az apoptotikus állapot marad stabil. Ez azt jelenti, hogy – amennyiben a fiziológias attraktorok a patológias attraktoroknál általános esetben is nagyobbak – a természetes zaj *segíti* a jelátvitel szabályozását, annak megfelelő keretek között tartását. Irodalmi kutatásom alapján a sejtben levő transzlációs zaj mértékére 10-15%-ot találtam [72], a fehérjék aktivitásának fiziológias zajára azonban nem találtam adatot. Ezen eredmények segítenek annak a hipotézisnek az alátámasztásában, mely szerint az állapottér méretétől függően a legnagyobb attraktorok azonosítása elég lehet a fő

működési folyamatok megértéséhez, mivel természetes zaj mellett a kisebb attraktorok hamar destabilizálódnak. Nem kizárólag az attraktor mérete számított azonban, hiszen az instabil proliferáló attraktor a stabil proliferálónál nagyobb medencéjűnek bizonyult, itt az instabilitást az a plusz ciklus okozhatta, ami alapján – az eredmények alapján joggal – instabilnak neveztük el az attraktort. Érdekes további megfigyelés volt az idősorok vizuális szemrevételezése alapján, hogy a három attraktor három különböző viselkedést mutatott a zajjal szemben. Az apoptotikus attraktor *elnyelte* a zajt, az onnan kiinduló, zajjal terhelt rendszerek a kísérlet végén változatlanok maradtak, az instabil attraktor *fel-erősítette* a zajt, amíg az olyan nagyra nem nőtt, hogy a rendszer pályája valamelyik másik attraktor felé módosult, míg a stabil attraktor *megtartotta* azt, abban sem elnyelést, sem disszipációt nem tapasztaltam. A stabil proliferáló attraktorba érkezett zaj ugyanazon a pályán mozgott ciklikusan, mint amit az instabil proliferáló attraktornál észleltem; kellő mértékű zaj esetén hasonlóképpen át is billent egy apoptotikus állapotba [73].

5.2.3.1. Daganatos mutációs profilok

Ugyan az attraktorok azonosítása, és a belőlük kinyerhető adatok vizsgálata önmagában is érdekes kutatási munkára ad lehetőséget, esetemben az attraktortervezésnek kifejezett célja az volt, hogy azonosítani tudjuk a vizsgált hálózatunk olyan állapotait, melyek a beavatkozás-tervező algoritmus logikus kezdő-, és célállapotai lehetnek, és így mód nyílik adott modelleken való konkrét beavatkozások megtervezésére.

A T-LGL leukémiás eredmények azonban bizonytalanságot ébresztettek bennem az egy állapottérben történő elemzéssel kapcsolatban. Huang és mtsai. [53] is feltételezik, hogy a daganatos állapotokhoz tartozó attraktorok normál állapotban nem elérhetőek, és a mutációk hatására változik meg úgy az állapottér alakja, hogy azok immár könnyen elérhetővé válnak.

Ez a feltételezés nem alaptalan. Egy kódoló szakaszban történt mutáció egy adott sejt genomjában, amennyiben az befolyásolja a pont aktivitását, egy *folyamatosan fennálló* beavatkozásnak tekinthető, ami ezért képes módosítani a sejt attraktor-térképét (állapottérét).

A daganatos sejtekben tapasztalt jelentős rezisztencia alapján következő hipotézisem az volt, hogy a daganatok kialakulásában fontos "driver"(vezérlő) mutációk (ezt a kifeje-

2. táblázat. A proliferáló attraktor dominanciájához vezető mutációsorozatok. A *ki* funkcióvesztő, a *be* konstitutív aktivációhoz vezető mutáció(ka)t jelent egy adott fehérje neve mögött. A P53, az ATM/ATR és az NF- κ B funkcióvesztése önmagában is elég erős hatás volt ahhoz, hogy egyeduralkodóvá tegye a daganatos attraktort. Az NF- κ B azért nincs feltüntetve, mivel az általa okozott "kvázi-proliferatív" állapotban a ciklinek aktiválásának módja nem egyezik teljesen meg az eredeti közleményben [24] bemutatott proliferációt eredményező aktivációs mintával.

Sorszám	Mutációk
1.	P53 <i>ki</i>
2.	BCL-2 <i>be</i> , PTEN <i>ki</i>
3.	S-NAIL <i>be</i> , CHK1/2 <i>ki</i>
4.	ATM/ATR <i>ki</i>
5.	CHK1/2 <i>ki</i> , PTEN <i>ki</i>
6.	BCL-2 <i>be</i> , AKT <i>be</i>
7.	E-CADH <i>ki</i> , CHK1/2 <i>ki</i>
8.	PTEN <i>ki</i> , IKK <i>be</i>
9.	AKT <i>be</i> , CHK1/2 <i>ki</i>
10.	PTEN <i>ki</i> , BAX <i>ki</i>
11.	CHK1/2 <i>ki</i> , AKT <i>be</i>
12.	AKT <i>be</i> , E-CADH <i>ki</i>
13.	AKT <i>be</i> , BAX <i>ki</i>

zést Rodin és mtsai. használták először 2002-ben [74]) a sejt állapotterét úgy változtatják meg, hogy a patológiás attraktor mérete a fiziológiásnál jelentősen nagyobbra nő, amely ponton a természetes zaj segítő szerepe hátráltatóvá válik, mivel a patológiás attraktorból való kilépést ugyanúgy megnehezíti, mint mutációmentes esetben az abba történő belépést.

Következő lépésként ezért megpróbáltam a Fumiã és Martins által leírt a [24] hálózatban azonosítani olyan mutációkat, illetve mutációk olyan kombinációit, melyek alkalmazásával az 11100 mikrokörnyezetben a patológiás, proliferatív attraktor mérete jelentősen nagyobbra nő, mint a fiziológiás, apoptotikus attraktor, esetleg ez utóbbi teljesen meg is szűnik. Ezt az attraktorkereső és a beavatkozás-tervező program kombinációjával hajtottam végre oly módon, hogy a beavatkozás-tervező genetikus algoritmus a futás során egyes szimulációk helyett az adott mutációkkal érintett hálózat attraktorait térképezi fel, és ezen eredményekből számítja ki a következő körben tesztelendő mutációsorozatokat. 13 ilyen mutációsorozatot azonosítottam, ezeket mutatja be a 2. táblázat. A lista nem teljes leírása az összes elképzelhető vezérlőmutáció-kombinációnak, valószínűleg jóval több

kombináció is lehetséges, ezek csak az általam végrehajtott futtatások eredményei. Összesen 25 kombinációból indultam ki, ezek proliferatív voltát egyesével visszaellenőriztem, és amelyekben a ciklinek aktivációs mintázata nem egyezett meg teljesen az eredeti cikkben szereplő proliferatív mintázattal (illetve az ismert aktivációs sorrenddel [75]), azokat nem vizsgáltam tovább, mivel a modell alapján a mutációsorozat proliferatív voltát nem láttam biztosítottak.

Az eredmények igen biztatóak. A P53 természetesen igen ismert tumor-szuppresszor, emellett mind az ATM/ATR [76], a PTEN [77, 78] az AKT [79–81], a BCL-2 [82, 83], a BAX [84, 85], a CHK1/2 [86] és az IKK [87–90] fehérjék táblázatban szereplő megváltozásainak daganatképződésre illetve sejtproliferációra gyakorolt hatása megtalálható az irodalomban. Ezek alapján valószínűsíthető, hogy a valós daganathoz vezető patológias folyamatok kialakulásának mechanizmusa az állapotter alakjának megváltozásával jól modellezhető.

5.2.3.2. Specifikus célponttervezés

A vezérlő mutációk által okozott állapotter-változások feltérképezésével adódik a kérdés, hogy lehetséges-e egy adott mutációs profilhoz specifikusan megtervezni egy olyan beavatkozást, mely a sejt állapotterét úgy változtatja meg, hogy az a beavatkozás alatt apoptotikus állapotba tud kerülni.

Ezt a vizsgálatot úgy oldottam meg, hogy kezdőállapotnak minden esetben az adott mutációs profilhoz tartozó legnagyobb attraktor egyik pontját jelöltem meg, végállapotnak pedig bármilyen apoptotikus állapotot. Mivel itt nem csak a rendszer pályáját, hanem az *állapotterét* kellett módosítani, így az alkalmazott beavatkozások szintén valamelyik fehérje aktivitásának állandó megváltoztatását jelentették. Ezen felül három további megkötést adtam a tervezőalgorithmusnak, hogy releváns eredményeket kaphassak:

- A citokróm *c-re*, és a tőle downstream pontokra (kaspázok, apoptózis), valamint a FADD-ra nem tehet beavatkozást, mivel ezek azonnal apoptózishoz vezetnek, de nem alkalmazható eredmények.
- A ciklinekre sem lehet beavatkozást tenni, mivel ez a proliferativitás azonosítását gátolja.

- Végül nem tehet beavatkozást azon fehérjékre, melyek az adott profilban mutánsnak voltak jelölve, mivel a.) triviális beavatkozás "visszafordítani" az adott mutációt, b.) a mutáns fehérje a megváltozott konformáció miatt nem biztos, hogy gyógyszer-rezhető.

3. táblázat. Mutáció-specifikus beavatkozások. A *gtl.* gátlást, az *akt.* aktiválást jelent, a *be* konstitutív aktivációt, míg a *ki* konstitutív inaktivációt (funkcióvesztést). A dőlt betűkkel jelölt beavatkozások egészséges sejten futtatott szimuláció alapján várhatóan toxikusak, a csillaggal jelöltek pedig nem vezettek teljes apoptózishoz (a hálózat "apoptózis" pontja ciklikusan kapcsolts csak be, nem volt folyton aktív).

Sorszám	Profil	Beavatkozás-sorozat
1.	P53 <i>ki</i>	<i>RTK gtl. ÉS (JNK akt. vagy TGF-β akt. vagy BAX akt.)</i>
2.	BCL-2 <i>be</i> , PTEN <i>ki</i>	<i>AKT gtl. ÉS (TGF-β akt. vagy BAX akt. vagy JNK akt.) *</i>
3.	S-NAIL <i>be</i> , CHK1/2 <i>ki</i>	<i>RTK gtl.* vagy RAS gtl.* vagy PI3K gtl.* vagy P53 akt.</i>
4.	ATM/ATR <i>ki</i>	<i>P53 akt.</i>
5.	CHK1/2 <i>ki</i> , PTEN <i>ki</i>	<i>PDK1 gtl. * vagy AKT gtl. *</i>
6.	BCL-2 <i>be</i> , AKT <i>be</i>	N/A
7.	E-CADH <i>ki</i> , CHK1/2 <i>ki</i>	<i>AKT gtl. * vagy PDK1 gtl. * vagy RTK gtl. * vagy RAS gtl. * vagy PI3K gtl. *</i>
8.	PTEN <i>ki</i> , IKK <i>be</i>	<i>NF-κB gtl. vagy BCL-2 gtl.</i>
9.	AKT <i>be</i> , CHK1/2 <i>ki</i>	<i>P53 akt. vagy (IKK gtl. ÉS (VHL gtl. vagy P14 akt. vagy ROS akt.))*</i>
10.	PTEN <i>ki</i> , BAX <i>ki</i>	<i>AKT gtl. vagy PDK1 gtl.</i>
11.	CHK1/2 <i>ki</i> , AKT <i>be</i>	<i>P53 akt.</i>
12.	AKT <i>be</i> , E-CADH <i>ki</i>	<i>IKK gtl. vagy NF-κB gtl. vagy BCL-2 gtl.</i>
13.	AKT <i>be</i> , BAX <i>ki</i>	N/A

Az ilyen megkötések mellett az egyes profilokhoz tervezett beavatkozásokat mutatja a 3. táblázat. A gátló beavatkozásokat *gtl.*-lel, az aktiválókat *akt.*-tal jelöltem. A csillaggal jelölt sorok azt mutatják, hogy az adott beavatkozás nem okozott teljes apoptózist, azaz az apoptózist jelző pont állapota ciklikusan változott a teljes aktivitás helyett. A dőlt betűs sorok, illetve fehérjék azt mutatják, hogy az adott beavatkozás várhatóan toxikus, egészséges sejten végzett ellenőrző vizsgálat alapján⁷. Megnyugtató, hogy az eredmények között több ízben előkerült a receptor tirozin kináz gátlás, mint beavatkozás (*RTK gtl.*), mivel a

⁷Az 5.2.3.3. fejezetnek megfelelően 00000 (natív) mikrokoznyezetben vizsgáltam az egészséges sejtre gyakorolt hatást. Amennyiben az így vizsgált sejt a tervezett beavatkozás hatására apoptotikus vagy proliferatív állapotba került, a beavatkozást toxikusnak tekintetem.

legelső és a mai napig eredményesen alkalmazott célzott daganatterápiás beavatkozások receptor tirozin kináz gátlók (cetuximab (Erbix), panitumumab (Vectibix): EGFR, sunitinib (Sutent): széles spektrumú RTK inhibitor, sorafenib (Nexavar): VEGFR, PDGFR). Mindemellett a többi gyakran előforduló talált beavatkozás is megjelenik az irodalomban, mint ígéretes kísérleti célpont, így mind a BCL-2 [91], a PDK1 [92], az IKK [93], az AKT [94]⁸, a PI3K [95]⁹ és az NF- κ B [97] gátlószerei.

Két mutációs profilhoz nem találtam beavatkozás-sorozatot ezen a hálózaton, a többi 11 esetében sikerült a tervezés. A két sikertelen tervezéshez tartozó profil nagyon hasonló: mindkét esetben konstitutívan aktív AKT mellett egy apoptotikus fehérje meghibásodását fedezhetjük fel. Elképzelhető, hogy ez a használt hálózat okozta korlátozás, azaz egy több fehérjét leíró, illetve a fehérjék egyes izoformáit külön-külön ábrázoló hálózatban ebben az esetben is sikerrel tervezhetünk terápiás célpontokat.

5.2.3.3. Tervezés mellékhatások minimalizálásával

A beavatkozások egyes elemei ugyan a szakirodalom által igazoltan is helyesnek tűnnek, de magukat a teljes beavatkozásokat sok esetben (1., ami, mint a P53 elvesztése, kiemelkedően fontos, ezen kívül a 2., 4. és a 11. profil, ez utóbbi kettő toxicitása a P53 közvetlen aktivációjának tulajdonítható) toxikusnak jósolhatjuk, ha megvizsgáljuk a beavatkozás hatását a 00000 mikrokörnyezetben, melyben alapesetben egy csendes (nem apoptotikus, nem proliferatív) attraktort találunk. Az említett beavatkozások azonban az ilyen sejtet is apoptózisba viszik, ami arra utal, hogy a beavatkozás nem specifikusan a hibásan proliferatív sejtekre hat, hanem várhatóan a legtöbb sejtet elpusztítja.

Ezért a tervezőalgorithmus további fejlesztésére volt szükség. Az algoritmust tehát úgy alakítottam át, hogy az egyes beavatkozás-jelöltekkel nem egy, hanem két szimulációt is futtattam a fitness-érték kiszámításához, egyet az előző módszernek megfelelően az 11100 mikrokörnyezetű sejtire, amit a beavatkozásnak apoptózisba kell vinnie, egyet pedig a 00000 mikrokörnyezetű sejtire, melyet a beavatkozás sem apoptózisba, sem proliferációba nem vihetett. Az így számolt két fitness-érték összeadásával adódott az új beavatkozás szelekciós értéke. Ezzel létre lehetett hozni "ideális" beavatkozásokat, me-

⁸A perifosine nevű AKT gátlószer pillanatnyilag II. klinikai fázisban van.

⁹A wortmannin nevű PI3K gátlószerrel folytatott kísérletek annak rossz *in vivo* farmakodinamikája miatt sikertelennek bizonyultak, jelenleg nanorészecskék segítségével történő alkalmazásban vizsgálják újra [96].

lyekben a daganatos sejt teljes apoptózisa mellett a natív sejtben egyáltalán nem következett be apoptózis. Ezek a beavatkozás-sorozatok azonban többnyire igen komplexnek bizonyultak, ezért az első vizsgálat által kifejezetten toxikusnak jelölt profilokhoz (1., 2., 4., 11.) készítettem olyan egyszerűbb beavatkozás-kombinációt, melyekben ugyan elképzelhető, hogy nem jut apoptózisba a sejt, de a natív sejtben egy pillanatra sem aktiválódik az apoptotikus útvonal.

Ezeket a beavatkozás-sorozatokat a 4. táblázat tartalmazza, míg az "ideális" beavatkozásokat az 5. táblázat.

4. táblázat. Nem toxikusra tervezett beavatkozások. A *gtl.* gátlást, az *akt.* aktiválást jelent, a *be* konstitutív aktivációt, míg a *ki* konstitutív inaktivációt (funkcióvesztést). Ezek a beavatkozások az egészséges sejtnek a szimulációk alapján nem ártanak, de a proliferáló sejtet apoptózisba viszik, bár többnyire az apoptózist mutató jel nem maximális erősségű.

Sorszám	Mutációk	Nem toxikus beavatkozás
1.	P53 <i>ki</i>	BCL-XL <i>gtl.</i> ÉS VHL <i>gtl.</i> ÉS (PDK <i>gtl.</i> vagy AKT <i>gtl.</i> vagy PI3K <i>gtl.</i>)
2.	BCL-2 <i>ki</i> , PTEN <i>ki</i>	TGF-B <i>akt.</i> ÉS P90 <i>gtl.</i> ÉS PDK1 <i>gtl.</i>
4.	ATM/ATR <i>ki</i>	VHL <i>gtl.</i> ÉS (RAS <i>gtl.</i> vagy PHD <i>gtl.</i>)
11.	CHK1/2 <i>ki</i> , AKT <i>be</i>	BCL-2 <i>gtl.</i> ÉS PHD <i>gtl.</i> ÉS SMADE2F <i>gtl.</i>

Új, és érdekes jelenség, hogy a beavatkozás-sorozatok számos esetben tartalmazták a von Hippel-Lindau fehérje gátlását. (VHL *gtl.*). Ez nem egy természetes beavatkozás, mert a VHL fehérje hibás működése von Hippel-Lindau szindrómában épphogy vesedaganat kialakulásához vezet [98], mindazonáltal közelebbről megvizsgálva ezeket a beavatkozásokat kiderült, hogy a VHL gátlására valóban nem az effektivitáshoz volt szükség, hanem ezen beavatkozás hozzáadása védte meg az egészséges sejteket az apoptózistól. Mindez azt jelentheti, hogy a VHL fehérje rövid idejű gátlása a célzott kezelések toxicitását csökkentheti, illetve eddig toxikusnak tartott beavatkozások válhatnak a fehérje gátlásának szinergiájával alkalmazhatóvá. Egy VHL gátlására alkalmas kismolekulát nemrég találtak Buckley és mtsai. [99], anémiás illetve iszkémiás állapotok kezelésére célozva. Érdeemes megjegyezni, hogy az anémia a célzott daganatterápiás szerek viszonylag gyakori mellékhatása, mely imatinib és sunitinib [100], sorafenib [101] és cetuximab [102] alkalmazása során is ismert.

Az 1. profilhoz nem találtam "ideális" beavatkozást, mely teljes apoptózist okozhatna a sejtnek úgy, hogy az egészségesre nem gyakorol apoptotikus hatást, ennek ellenére

5. táblázat. "Ideális" beavatkozások. A *gtl.* gátlást, az *akt.* aktiválást jelent, a *be* konstitutív aktivációt, míg a *ki* konstitutív inaktivációt (funkcióvesztést). Ezek a beavatkozások teljes erősségű apoptózist eredményeznek az adott mutációsorozattal rendelkező proliferáló sejtek esetében, de egyáltalán nem okoznak apoptózist a nyugalomban levő, mutációmentes sejtek esetében a szimulációk szerint. A beavatkozások viszont az előző két csoportban láthatóaknál többnyire komplexebbek.

Sorszám	Mutációk	"Ideális" beavatkozás
1.	P53 <i>ki</i>	N/A
2.	BCL-2 <i>be</i> , PTEN <i>ki</i>	P14 <i>akt.</i> ÉS BCL-2 ÉS BAX <i>akt.</i> ÉS (AKT <i>gtl.</i> vagy PDK <i>gtl.</i>)
3.	S-NAIL <i>be</i> , CHK1/2 <i>ki</i>	BAD <i>akt.</i> ÉS NF- κ B <i>akt.</i> ÉS HIF1 <i>akt.</i> ÉS (RTK <i>gtl.</i> vagy RAS <i>gtl.</i> vagy PDK <i>gtl.</i> vagy AKT <i>gtl.</i>)
4.	ATM/ATR <i>ki</i>	BCL- <i>gtl.</i> ÉS AKT <i>akt.</i> ÉS VHL <i>gtl.</i> ÉS FOXO <i>gtl.</i>
5.	CHK1/2 <i>ki</i> , PTEN <i>ki</i>	BAD <i>akt.</i> ÉS VHL <i>gtl.</i> ÉS AMPK <i>gtl.</i> ÉS AKT <i>gtl.</i>
6.	BCL-2 <i>be</i> , AKT <i>be</i>	N/A
7.	E-CADH <i>ki</i> , CHK <i>ki</i>	BAD <i>akt.</i> ÉS SMAD <i>gtl.</i> ÉS VHL <i>gtl.</i> ÉS (RAS <i>gtl.</i> vagy PDK <i>gtl.</i> vagy AKT <i>gtl.</i>)
8.	PTEN <i>ki</i> , IKK <i>akt.</i>	NF- κ B <i>gtl.</i> vagy BCL-2 <i>gtl.</i>
9.	AKT <i>be</i> , CHK1/2 <i>ki</i>	BCL-2 <i>gtl.</i> ÉS AKT <i>gtl.</i> ÉS VHL <i>gtl.</i> ÉS FOXO <i>gtl.</i>
10.	PTEN <i>ki</i> , BAX <i>ki</i>	AKT <i>gtl.</i> vagy PDK1 <i>gtl.</i>
11.	CHK1/2 <i>ki</i> , AKT <i>be</i>	MYC/MAX <i>akt.</i> ÉS AKT <i>gtl.</i> ÉS VHL <i>gtl.</i> ÉS (IKK <i>gtl.</i> vagy RAS <i>gtl.</i>)
12.	AKT <i>be</i> , E-CADH <i>ki</i>	IKK <i>gtl.</i> vagy NF- κ B <i>gtl.</i> vagy BCL-2 <i>gtl.</i>
13.	AKT <i>be</i> , BAX <i>ki</i>	N/A

összességében meglepőnek tartottam, hogy már egy ilyen alacsony felbontású hálózat-
tal (csak néhány útvonalat tartalmaz és sok különböző fehérje, illetve izoforma össze lett
vonva egy pontba, pl. RTK, RAS, RAF) is ilyen szofisztikált, és a szakirodalommal jól
egybevágó beavatkozási terveket lehet alkotni, melyek képesek két különböző állapotú
sejtre más-más hatást kifejteni.

5.2.3.4. Zavaró mutációk

Amikor egy rosszindulatú daganat észlelhetővé válik, akkor természetesen már nem csak
a vezérlő mutációkat tartalmazza. A daganat számára az invazivitást, intravazációt és
egyéb tulajdonságokat adó mutációkon kívül a nagymérvű DNS-károsodás során olyan
"potyautas" (passenger) mutációk is bekerülhetnek a daganatos sejtek genomjába, me-
lyek kifejezetten rezisztenssé tehetik azt a célzott terápiára. Ennek legjobb példája a
kRAS gén mutációja, mely esetében a konstitutívan aktív kRAS fehérje a tumort EGFR-

gátló gyógyszerekkel szemben igen ellenállóvá teszi [103]. A kRAS gén vizsgálata az egyik legfontosabb, klinikai gyakorlatban széles körben alkalmazott molekuláris vizsgálat daganatos betegeknél; ennek a vizsgálatnak a szükségessége az EGFR-gátló gyógyszerek (Erbix/cetuximab, Vectibix/panitumumab) leírásában is feltüntetésre került. Ezért megvizsgáltam azon mutációs profilokat, melyekre tervezett beavatkozás receptor tirozin kináz gátlást is tartalmaz (1., 3., 7.), hogy egy újabb, RAS-t aktiváló mutáció esetén a szimuláció szerint is elvesztik-e a hatékonyságukat.

Valóban ez történt. Az összes vizsgált beavatkozás elvesztette az apoptotikus hatását, ráadásul egyetlen eset kivételével, ahol egy TGF- β aktiváló beavatkozás szerepelt (az 1. beavatkozáscsoport egyik megvalósítása), a sejtek proliferációja is változatlan maradt. Ezzel szemben azok a beavatkozások, melyek más útvonalakat céloztak, továbbra is hatásosak maradtak, tehát például a PI3K gátlása továbbra is hatásos terápia volt a konstitutívan aktivált RAS-sal kiterjesztett 3. és 7. profil esetében, továbbá a 7. profil esetében a PDK1 és az AKT gátlása is.

Az 1. profil esetében azonban minden egyszerű beavatkozás-sorozat tartalmazott RTK gátlást. Így, mivel a P53 fehérje funkcióvesztése igen általános különféle daganatok kialakulása során, erre az esetre (P53 *ki*, RAS *be*) külön terveztem egy kezelést. Ebben az esetben azonban a legegyszerűbb kezelés is elég komplex, ami megerősítheti azt a megfigyelést, hogy kRAS mutáns esetben a betegség rossz prognózisa várható.

A legegyszerűbb ilyen beavatkozás-sorozat három elemből áll: RAF gátlás, BAX aktiválás és vagy a PI3K vagy a PDK gátlása. Ennek segítségével a sejt teljesen apoptózisba vihető, cserébe az egészséges sejtekre számolt szimulációk alapján magas toxicitás várható. Érdekes itt megemlíteni, hogy az első, Vemurafenib nevű bRAF gátlószer 2011 augusztusában került piacra, előrehaladott melanóma kezelésére való indikációval. A legegyszerűbb olyan kezelés, ami az egészséges sejteket nem vitte apoptózisba, de a daganatos sejteken kifejtett apoptotikus hatást (habár nem teljeset), megegyezik az 1. profil eredeti változatára számolt nem-toxikus beavatkozással a 4. táblázatban, azaz BCL-XL gátlás, VHL gátlás és például AKT gátlás szükséges hozzá.

6. Megbeszélés

6.1. A Turbine program

Biológiai hálózatok esetében már a hálózat felépítése során is érdekes információkhoz juthatunk. Erre nagyon jó példa a Signalink adatbázis [104], mely három faj (*C. elegans*, *D. melanogaster* és *H. sapiens*) nyolc jelátviteli útvonalát tartalmazza. A három útvonal összehasonlításával több alkalommal is új fehérjékre, illetve fehérjék új funkcióira bukkantak a kutatók egy másik faj azonos jelátviteli útvonalában azonos helyen levő ismert ortológ funkciója vagy szerkezete alapján.

Egy másik érdekes topológiai módszer a Diseasesome [105], melyben különféle betegségek hálózatát építették fel, egy fehérje-fehérje kölcsönhatás-hálózat az adott betegségben érintett fehérjéinek távolsága, illetve átfedése segítségével. Egy ilyen hálózat olyan betegségek hasonlóságára, hívhatja fel a figyelmet, kezeléseket esetlegesen olyan kereszt-hatásait segíthet meghatározni, melyek a hálózat ismerete nélkül messze nem lennének magától értetődőek.

A dinamikai leírás igénye a 2000-es évektől kezdődően jelent meg kísérletes szinten. A Broad Institute egy minden génre kiterjedő perturbációs kísérletsorozat lefolytatását tűzte ki céljául [106–108]. Az ilyen kísérletek elvégzését megkönnyítette a kísérleti technológiák fejlődése, ma már egy sejtpopuláció mRNS-, illetve fehérjekoncentrációprofilját nem csak pillanatképként, hanem több időpontban, idősor formájában is tudjuk mérni [109].

Hálózatok dinamikájával adatok híján a 2000-es évek közepéig igen keveset foglalkoztak. 2005 előtt általános dinamikát főleg sejtautomatákon (cellular automata), azaz gyakorlatilag csak rácshálózatokon vizsgáltak elterjedten. Az ideghálózatok működésének dinamikai leírása ennél régebben, az 1990-es években terjedt el. A Turbine program megírása előtt is rendelkezésre állt néhány, specializált rendszer, melyben volt lehetőség hálózatok modellezésére, így biokémiai rendszerek modellezésére a BIOCHAM [110], ideghálózatok modellezésére a NEURON [35], a véletlen bolyongást modellező ITM-

Probe [111], Boole-dinamikára pedig a CellFrame [112]. A Turbine ezekkel szemben egy általános, bármilyen algoritmizálható dinamikát szimulálni képes, gyorsaságra optimalizált szimulátornak készült, az első verziója 2011-ben került megemlítésre egy Science Signaling cikkben [113], míg a szimulátor teljes, első eredményekkel megerősített leírását 2013-ban publikáltuk [114]. Eközben 2012-ben bemutatták a Conedy nevű programot is [115], mely egy Python kiegészítés hálózatok szimulációjához, ami ily módon hasonlóképp különféle modelleket is tud kezelni, előnye a Python segítségével beépített differenciálegyenlet-megoldó rendszer, és a Python-környezet által biztosított számos csomag elérhetősége. Elsődlegesen nem sebességre, hanem programozhatóságra lett tervezve, azaz arra, hogy kisebb hálózatokon különböző (akár minden egyes pontban más) dinamikákat ki lehessen próbálni. Ennek eredményeképpen a Turbine-nál lassabb; egy magas futtatással mérve a [115] cikk és a dolgozatban bemutatott szimulációs sebességmérések alapján a Turbine 64-szer gyorsabb (a Conedy ismereteim szerint nem támogat több magon, illetve videokártyán való futtatást, a Turbine CUDA (videokártyán futó) verziójával összehasonlítva az a Conedynél körülbelül 1.600-szor gyorsabb), ezen kívül nincs hozzá megjelenítő, illetve egyéb, nagyobb hálózatok esetében lényeges funkciók, mint az adatfile-ok közvetlen merevlemezzel való kezelése is hiányoznak.

A Turbine jelenlegi verziójának tervezése és fejlesztése során azonban a Conedy előnyös tulajdonságait érdemesnek tűnt figyelembe venni, ezért szintén a Python nyelv integrálása mellett döntöttem. A Conedy megközelítésével szemben azonban, mely a Python kiegészítésére épül, azaz a programot a Python értelmező irányítja, a Turbine a Python értelmezőt *beágyazva* tartalmazza. Így a program futása a Turbine szerver irányítása alatt áll, de szükség esetén a Turbine eszközei, illetve a különféle Python által biztosított szolgáltatások Python-programkód segítségével is vezérelhetőek. Ezzel a megoldással a program használati köre, illetve testreszabási lehetőségei jelentősen szélesedtek, miközben a futásidő nem növekedett lényegesen.

2010-től kezdődően egyre több irányban indultak biológiai hálózatok dinamikájával foglalkozó kutatások, például a Boole-, vagy Boole-szerű dinamikai megoldásokat egyre részletesebben, attraktorok illetve lehetséges beavatkozások szintjén is vizsgálták [23, 24, 49]. Zhong és mtsai. [116] megmutatták, hogy a "finomabb", durva fehérjeszerkezet-

változáshoz nem vezető mutációk dinamikai modellekben a biológiai interakciós hálózatok *élsúlyainak* megváltozásával modellezhetőek.

Általánosabb hálózatelméleti megfontolásból indult el Liu és mtsai. [117] cikkje alapján az a kutatási terület, melyben hálózatok irányíthatóságát vizsgálták, azaz megmutatták, hogy pusztán elméleti megfontolások (egészen pontosan a Kálmán-féle irányíthatósági feltétel [118] hálózatokra való felírása) alapján meg lehet egy irányított hálózatról mondani, hogy melyik pontokra kell beavatkozásokat tenni ahhoz, hogy a hálózat bármely tetszőleges állapotból bármelyik másik állapotba átvihető legyen. Megmutatták, hogy ez valóban lehetséges, és egy egyszerű szabály alkalmazásával, a maximális fedés módszer használatával (maximum matching method) ezek a pontok megtalálhatóak. Az eredmények azonban rámutattak arra is, hogy valós hálózatokban az irányítandó pontok száma igen nagy, a hálózat összes pontjának 10 és 90%-a közötti mennyiségű ponton kell beavatkozni, hogy valóban bármelyik állapotból bármely másik állapotba át tudjuk vinni a rendszert.

A Turbine beavatkozás-tervező program ötlete többek között ezen utóbbi eredmények hatására merült fel bennem 2012-ben, azon megfontolásból, hogy amennyiben ismerjük a dinamikai működés pontos leírását, illetve csak egy kezdő-, és egy célállapotot választunk ki, a szükséges pontok száma a fenti esetben számított értéknél jelentősen alacsonyabb is lehet. A dolgozatban bemutatott eredmények alátámasztották ezt a feltevést. Nagyon kevés esetben volt csak szükség egyszerre 3 pontnál több helyen történő beavatkozásra mind a T-LGL, mind a Fumiã és Martins által leírt [24] hálózat esetében.

A Turbine program által megvalósítotthoz hasonló, iterált szimulációkkal végrehajtott beavatkozás-tervezéssel nem találkoztam az irodalomban, de két helyen is találtam utalást arra [119, 120], hogy az a szimuláció lassúsága miatt nem megvalósítható. Szerencsére ez már a Turbine beavatkozás-tervező tesztelése után történt, így kiderült, hogy a program sebessége nem csak gyorsabb az eddigieknél, de a nagyságrendi sebességkülönbség hatására a Turbine használatával valóban új típusú vizsgálatok lehetősége nyílik meg.

A [119] cikkben ezen kívül a Turbine-hoz hasonló, egy állapotból egy állapotba történő beavatkozás-tervezésre is bemutatott egy NECO nevű módszert, melyben a szimulációk helyett egy közelítő mátrix használatával jósolják egy kisméretű beavatkozás hatását, és az így jósolt kisméretű beavatkozások összeadásával, azok egymás után láncolásával

határozzák meg a végső alkalmazandó beavatkozást. A módszer előnye, hogy csak néhány szimuláció lefuttatását igényli, így kis hálózatok esetében a Turbine-nál gyorsabb működést várhatunk tőle. Nagyobb hálózatok esetében ez az előny már nem egyértelmű, mivel a program számításai $O(n^{2.5})$ szerint skálázódnak, míg a Turbine mind elméleti megfontolások, mind mérési eredmények alapján $O(n^2)$ szerint, azaz 10.000 pontnál már egy nagyságrendnyi futásidő-különbséget még a párhuzamosítást nem számolva is "behoz" a Turbine. A két program által talált beavatkozások nagymértékben eltérnek egymástól, a NECO általában a lehetséges beavatkozási pontok nagy részét felhasználja, azok mindegyikére kisebb-nagyobb gerjesztést alkalmazva, a Turbine beavatkozástervező pedig inkább a beavatkozási pontok minimalizálására törekszik, egy-egy pontra nagyobb gerjesztést téve. A NECO alkalmazása ezért olyan esetekben előnyös a Turbine-nal szemben, ahol a rendszer könnyen, bármely pontján módosítható, és fontos, hogy egy gerjesztés ne legyen "túl nagy", mint például az áramellátási hálózatok esetében; olyan esetekben, ahol viszont fontos, hogy egyszerre kevés beavatkozást alkalmazzunk, ott a Turbine-től jobb eredményekre számíthatunk, például a biológiai hálózatok esetében, hiszen nem célszerű 500-féle gyógyszert adni egy betegnek, egy áramellátó-hálózat 500 állomásán egy időben kicsit módosítani viszont célravezetőbb, mint egy nagy "tüskét" küldeni a villamos hálózatba. Kevés egyszerre történő beavatkozás alkalmazásának további előnye, hogy amennyiben még kicsit is bizonytalan a hálózat szerkezete vagy paraméterei, kevés beavatkozás esetében robosztusabb eredményeket kaphatunk, hiszen sokkal kevesebb pontnak kell teljesen a modell szerint elvárt viselkedést produkálnia a valóságban ahhoz, hogy a beavatkozás működőképes legyen.

6.2. Terjesztési központiság

A kidolgozott terjesztési központiság mérőszám tulajdonságait elsősorban az alatta levő dinamikai modell, a közlekedőedények modellje határozza meg. A modell matematikai elemzésével megmutattuk [114], hogy amennyiben egy ponthoz hozzáadott energia a pontot magába foglaló modul méretéhez képest kicsi, a szimuláció során mind térben, mind időben exponenciális lecsengésű lesz a folyamat. Amennyiben az energia a pontot magába foglaló modulhoz mérten nagy, a modulból az energia a modulon belüli terjedés-

nél lassabban jut ki a modulon kívülre, így térben az energiaeloszlás egyenletessé válik (azonos szintre "töltődik" a modul), időben pedig lineáris lecsengésűvé (a pontok nem tudnak egymásnak energiát átadni, mivel azonos energiaszinten vannak). Valós helyzetekben a lecsengési függvény valahol a kettő között helyezkedik el, tapasztalataim szerint az exponenciális formához jóval közelebb. Maslov és Ispolatov mutatta meg 2007-ben, hogy a fehérjék szabadkoncentráció-változásai exponenciális terjedést mutatnak [121], de kvalitatív módon az összes intenzív (kiegyenlítő) fizikai változó értékei a nevükből adódóan hasonló kiegyenlítődést mutatnak. Az állapotváltozók (energia, információ) modulokba való "beragadása" a hálózatos szakirodalomban szintén ismert, a modulhatárokat már többször azonosították információ-átadási szempontból is határvonalként [122–124].

A másodlagos struktúrák terjesztési központiságai között mért különbségekre bár irodalmi adatot nem találtam, az eredmény, miszerint az α -hélix szerkezetben levő aminosavak az átlagosnál jobb, míg a rendezetlen szerkezetben találhatóak az átlagosnál alacsonyabb terjesztési központiságot mutattak, igen jól egybevág azzal a strukturális megfontolással, mely szerint az α -hélixekben található aminosavak a rendezetlenek struktúrák aminosavainál "sűrűbbek", azaz az aminosavak egyes atomjai közt több különféle típusú csatolást találhatunk [125].

A metionil-tRNS szintetáz molekula fehérjeszerkezet-hálózatán bemutatott eredmény, mely szerint a terjesztési központiság segítségével a tRNS kötőhelyek a tRNS jelenléte nélkül is azonosíthatóak, azt mutatják, hogy a tRNS bekötődése hatására valószínűleg olyan formában változik meg a fehérje konformációja, hogy a két kötőhely közötti kommunikációt lehetővé tegye, hiszen a terjesztési központiság mérőszám kiemelkedő növekedése a két tRNS kötőhely környékén a fehérje tRNS-kötő konformációjában pontosan azt jelenti, hogy ezen aminosavakról induló információk gyorsabban el tudnak terjedni a fehérje többi pontján. Ilyen típusú kísérletekre, illetve eredményekre találhatóak irodalmi példák. Zhou és mtsai. [126] több mint 500 különböző fehérje ubikvitinációs helyeinek closeness mérőszámát vizsgálták, és megállapították, hogy a closeness (közelségi) központiság szempontjából ezek a helyek kiemelkedőnek tekinthetők. Tang és mtsai. [127] megmutatták, hogy a fehérjék aktív centrumának meghatározásában az aminosavak konzerváltsága után a closeness központiság a második legjobb prediktor.

Az élesztő fehérje-fehérje kölcsönhatás-hálózatában végzett kísérletünk alátámasztotta az eredeti cikkek [57, 128] eredményét, miszerint az élesztő fehérje-fehérje kölcsönhatás-hálózata hő sokk hatására modulárisabbá válik, hiszen az átlagos terjesztési központosság érték jelentősen, 40,5%-kal csökkent. Ennek a csökkenésnek az oka – azonos gerjesztésméret mellett – az, hogy az egyes pontokra tett perturbáció a két különböző hálózatban kevesebb pontot ér el, amely hatást könnyen okozhat a modulok erősebb szétválása hő sokkban, hiszen így a modulon kívüli pontok csak hosszabb úton érhetőek el, tehát a pontok átlagos távolsága nagyobb lesz, ami végeredményben az exponenciális lecsengés miatt kevesebb pont elérését eredményezi.

A stressznek kitett élesztőben terjesztési központosság szempontjából mind a legnagyobb növekedést, mind a legnagyobb csökkenést mutató fehérjék funkciói az irodalomban is ismertek, mint stresszben kifejezettebbé, illetve erősebben szabályozottá váló funkciók. A legfontosabb stresszben növekvő funkció a szénhidrát-metabolizmus enzimeinek koncentráció növekedése [68], a legfontosabb stresszben csökkentő funkció pedig a riboszóma alkotóelemeinek koncentráció csökkenése [69] voltak a terjesztési központosság eredmények szerint, amelyek egybevágtak a kísérletesen igazolt eredményekkel.

6.3. Attraktorok és beavatkozás-tervezés

Az a gondolat, hogy egy rendszer lehetséges működésének megértéséhez elég az egyes attraktorok funkcióit megismerni, leginkább talán Kauffman két alapvető fontosságú cikkének volt köszönhető [41, 42], melyben megmutatta, hogy biológiai típusú rendszereknek a rendszer strukturális felépítéséből adódóan kis számú attraktornal kell rendelkezniük. Ezzel gyakorlati fontosságot nyert Waddington 1957-es epigenetikai eszmefuttatása, és az ebben a könyvben publikált, azóta híressé vált "tájképét" (epigenetic landscape) [40] azóta igen sokszor mutatták be különféle biológiai rendszerek fejlődésének mintázatát példázandóan.

A Fumiã és Martins által leírt [24] hálózaton végzett vizsgálataim, melyben attraktorokat rózsaszín zajnak tettünk ki, azt az új információt tették ehhez a képhez hozzá, hogy vélhetően nem kell az összes attraktort megismernünk, a rendszer működésének leírásához elég lehet a legnagyobb attraktorok felderítése [73]. Ez nagy könnyedséget jelent

az állapotter leírásában, mert így véletlen kezdőfeltételekkel indított szimulációsorozattal ezek az attraktorok az *állapotter méretétől függetlenül* megtalálhatóak, mivel a szükséges szimulációk számának meghatározásához csak a legkisebb vizsgálandó attraktor és a legnagyobb vizsgálandó attraktor méretének arányát szükséges ismerni, azaz nem az állapotter teljes mérete érdekes, csak az attraktormedencék relatív mérete.

A T-LGL hálózat vizsgálatát inkább az algoritmus működésének igazolása céljából végeztem, itt megmutattam, hogy egy teljes keresést lehetővé tevő méretű hálózaton a Turbine segítségével is hasonlóképp teljes képet kaphatunk a rendszer attraktorairól, mivel ilyen kis méretű hálózatoknál igen sok szimuláció lefuttatására nyílik lehetőség. A proliferáló attraktorok esetében a P2 pont billegése felhívja a figyelmünket arra, hogy minden ilyen "billegő" pont az attraktorok számának duplázódásához vezethet. Amennyiben a teljes hálózatunk több részgráfja is gyakorlatilag külön-külön többféle attraktorral rendelkezhet (mint az várható egy teljes proteom modell esetében, ahol egymástól független funkciók jelennek meg), várhatóan egyszerűbb, érthetőbb képet kapunk – és nem utolsósorban a számítás is könnyebbé válik –, ha ezen részgráfok attraktorait külön-külön határozzuk meg, majd a gráf "összerakása" után megnézzük, hogy ezen attraktorok milyen kombinációi zárják esetleg ki egymást. A nagy könnyítést itt az okozza, hogy a gráf "összerakásakor" csak az attraktor-állapotokat kell átvizsgálni, itt már a teljes állapotterben történő vizsgálatra nem lesz szükség, hiszen ha egy állapot egy adott részgráfban már önmagában sem stabil, akkor a teljes gráfba "visszarakva" sem lesz az. Egy ilyen redukciós módszert közölt Zañudo és Albert 2013-ban [129].

Az attraktor-, vagy részattraktor-alapú leírások működésének szükségessége elméleti és alapvető klinikai megfontolásokból is következik. Elméleti alapon azt mondhatjuk, hogy akkor lehetséges egy hálózatban a pontok állapotváltozóinak bármilyen kombinációja nyugalmi állapotban, ha a hálózat nem tartalmaz éleket (vagy igen furcsa dinamikát használunk). Minden egyes él nem csak egy kölcsönhatásként fogható fel, de egy megszorítását is jelenti a hálózat nyugalmi állapotterének, azaz bármely két, kölcsönhatásban levő pontpárnak csak a kölcsönhatás által megengedett közös nyugalmi állapota lehetséges. Mivel a vizsgált gráfjaink nagyrészt összefüggőek, ezért az élek nagy száma a lehetséges nyugalmi állapotok számát is jelentősen szűkíti. Az attraktorok létjogosultságának klinikai megfontolása egyszerűen a különböző biomarkereink működésének ténye.

Egy közismert példán értelmezve ezt a gondolatot, ahhoz, hogy a szérumban a glükóz szint megváltozásából következtetni tudjunk a sejtek inzulinrezisztenciájára, vagy a β -sejtek pusztulására, ahhoz szükséges, hogy létezzenek olyan biokémiai szabályozórendszerek, azaz reakcióhálózatok, melyek kényszerfeltételei – azaz a hálózat élei – által gyakorolt hatások között megtalálhatóak az inzulin hormon által a glikolízisre illetve a glükoneogenezisre gyakorolt ellentétes értelmű reakciók, oly módon, hogy annak szintje fiziológiás esetben a 3-5 mmol/l érték közé kell esnie. Ilyen formában az éhomi vércukorszint közel állandó értéke egy szabályozott folyamat, és ebből következően egy attraktor jelenlétét mutatja.

Az egyszerű betegségek, melyekhez egyértelmű metabolikus vagy mutációs háttér társítható, ily módon egy attraktorhoz lehetnek köthetőek, hiszen adott biomarkerek egy konkrét állapotával asszociáltak. A komplex betegségek ezzel szemben pont azért lehetnek komplexek, mert több különböző biomarkerrel is gyengén asszociáltak, azaz több különböző attraktorhoz tartoznak. Ily módon a GWAS sikertelensége komplex betegségek esetében tulajdonképpen magától értetődő, de elképzelhető, hogy meggyőző eredményeket tudnánk kapni, ha a vizsgált betegcsoportokat az attraktoroknak – illetve ebből következően a biomarkerek állapotának – megfelelően tudnánk vizsgálni. Nem egyértelmű azonban, hogy ezen betegségekben több, állandó attraktorról van szó, vagy minden komplex betegség egy többé-kevésbé új attraktort alakít ki magának. Baffy és Loscalzo [130] ez utóbbi mellett érvelnek, bemutatva egy folyamatot, melyben az adverz hatásnak (pl. túltáplálás) többször kitett szervezet előre nem látható módon távolodik el folyamatosan a fiziológiás attraktortól az állapotterben, míg egy, a betegség manifeszt formáját eredményező attraktort nem talál magának, mely attraktorban a rendszer ezen adverz hatás mellett is stabil tud maradni.

A különféle daganatok attraktornak való megfeleltetésének lehetőségét 2009-ben írták le Huang és mtsai. [53]. Az attraktoralapú, dinamikus elemzésekre támaszkodó terápia az elmúlt évben még jobban a figyelem középpontjába került [54, 131]. A Turbine beavatkozás-tervezője pontosan a [131] cikkben felvetett attraktorból-attraktorba történő átvitel problémájára kínál megoldást. A publikációban [131] jelzik, hogy valószínűleg egyetlen gyógyszer sokszor nem elég, több gyógyszer racionálisan tervezett kombinációjára lehet szükség. Bár a T-LGL hálózatra számított beavatkozások azt mutatják, hogy

mutációmentes esetben egy gyógyszeres kezelés elég, a Fumiã és Martins által leírt [24] hálózat mutációkkal terhelt változataiban kielégítő eredmény eléréséhez valóban az esetek nagy részében több célpontra egyszerre történő beavatkozás szükséges.

Az igazán hatásos kezelésekhez azonban szükséges lehet a daganatok egyéb tulajdonságait is figyelembe venni. A tumorok már a detekciós limit elérésekor sem homogének, azok több klonális populációból állhatnak [132], különböző genetikai mutációkkal, így a megfelelő kezeléshez a tumor heterogeneitását mindenképpen figyelembe kell venni. Zhao és mtsai. [133] megmutatták, hogy a tumor heterogeneitása a klinikumban pillanatnyilag megtalálható gyógyszerek közötti szinergiát gyengítik, a tumor heterogeneitását figyelembe véve a meglévő gyógyszerek kombinációs kezelése nem nyújtott az egyes kezelések összehatásánál nagyobb előnyt. Amennyiben a tervezési illetve már a gyógyszerfejlesztési fázisban figyelembe vesszük a lehetséges kombinációs terápiákat, a klonális frakciók ismeretében tervezhető lehet egy olyan beavatkozás, mely az összes meglévő frakcióra hatásos lehet, de az ilyen kezelés elképzelhető, hogy jelentősen komplexebb lesz az egyetlen frakcióra hatásos beavatkozásoknál. Elképzelhető ezzel szemben az is, hogy sok célzott gyógyszer a tumor különböző frakcióiban "újra felhasználható" lesz, azaz a különböző frakciók megcélzásához elég lehet akár csak egy-egy új célpontra ható gyógyszert hozzáadni a terápiás keverékhez. A keverék komplexitásának növekedésével az egészséges sejtek állapotát egyre nehezebb lesz a "helyén tartani", azaz a kezelés toxicitása, illetve a mellékhatásspektrum várhatóan nőni fog, ezért a célzott, de szisztémásan adagolt terápiák mellett az adagolásban is fontossá válhat a daganatsejtek minél pontosabb megcélzása. Ilyen "célzott-célzott" terápiákban talán a komplexebb kezelések is elfogadható mellékhatásprofilokat mutatnak majd, különösen, mivel a többkomponensű gyógyszerek egyes komponenseiből kisebb mennyiség is elég [134]. A jelenlegi gyógyszerek mindenesetre nem tűnnek elég precíznek ilyen típusú beavatkozások összeállításához, ezért a gyógyszerfejlesztésben fontossá válhat egy olyan irány felfuttatása, mely egyes fehérjéket minél specifikusabban képes gátolni vagy aktiválni, akár úgy is, hogy önmagában betegség kezelésére nem alkalmas, csak "építőkövek" egy racionálisan tervezett terápiás keveréknek.

Szintén nem hagyható figyelmen kívül a tumorok gyors evolúciójának képessége sem. Bozic és mtsai. [135] megmutatták, hogy a detekciós limit méreténél nagyobb tömör tu-

morokat már nem feltétlenül lehet olyan gyorsan elpusztítani egyszerű diffúziós problémák miatt, hogy ne alakulhasson ki legalább egy új mutáció, illetve rezisztencia. Ezért a beavatkozások tervezésénél annak figyelembevétele is szükségessé válhat, hogy *egy* mutáció ne tehesse rezisztenssé a sejtet a beavatkozással, illetve a beavatkozás-sorozattal szemben. Erre is van mód a Turbine segítségével, de mivel az új mutáció gyakorlatilag bárhol lehet, az ilyen típusú tervezésnek a terápia komplexitására gyakorolt hatását még nem lehet megbecsülni. A komplexitás mérsékelt növekedése esetén az előző bekezdésben vázolt "célzott-célzott" terápiák alkalmazása ebben az esetben is célravezető lehet.

Amennyiben a hálózathoz nem lehet egyértelmű dinamikát társítani, vagy a kinetikai paraméterek nem állnak rendelkezésre, egy egyszerűsített beavatkozás-tervezési lehetőséget nyújt a Bayes-i elemzés. Ha az élek irányai rendelkezésre állnak, az ok-okozati összefüggésekből Bayes-i következtetéssel (adott okozatok által terjesztett valószínűségi sűrűségfüggvények számításával) meghatározhatóak a hálózat két adott állapota közötti eltérés *legvalószínűbb okozói*. Ezek az okozók a "gyanúterjedés" (belief propagation, BP (nincs hivatalos magyar fordítása)) algoritmus, illetve annak "könnyített" változata, a Viterbi-algoritmus segítségével algoritmikusan megtalálhatóak. Ezek az algoritmusok azonban igen lassúak (NP, illetve #P-nehezék). Chindelevitch és mtsai. [136] megmutatták, hogy erre az esetre létezik a problémának analitikus megoldása, mely gyorsabban számolható, de még ez a módszer is jelenleg csak egy "okozat", azaz egy célpont gyors megtalálására képes.

Egy hálózatkutató számára természetesen minden hálózat. Amennyiben azonban a fehérje-fehérje kölcsönhatás-hálózatok, illetve a jelátviteli hálózatok nem bizonyulnak kellő pontosságúaknak, a fehérjék egyedi komplexeinek modellezéséhez a hálózat valószínűleg nem jó absztrakció, ami annak az egyszerű igazságnak, hogy „hiba volna mindent szögnek nézni azért, mert kalapács van a kezünkben” a helyességét bizonyítja ebben az esetben is. Ettől függetlenül, amennyiben a készített szimuláció elég gyors, a Turbine az ilyen nem-hálózatos szimulációkhoz is megfelelő környezetet nyújthat, hiszen a beavatkozás-tervező genetikus algoritmus csak a beavatkozást leíró szövegfájl-okkal dolgozik, számára a szimulációs absztrakció lényegtelen.

Nem csak hálózatos, vagy szimulációs módszerrel lehet azonban beavatkozásokat tervezni, egy "fekete dobozos", modellmentes megközelítés is eredményes lehet, ha ren-

delkezésre áll megfelelő adatmennyiség. Creixell és mtsai. [54] az egyes gyógyszerek hatásosságát közvetlenül a daganatban kifejeződő biomarkerekhez társították, belső, hálózatos modell nélkül, és ezen modellen történő egyszerű lineáris optimalizációval határozták meg az adott heterogén tumor esetében alkalmazandó beavatkozás-sorozatot.

A gyógyszer-célpont-felderítésen kívül más kutatási folyamatokban is segítséget nyújthat a beavatkozás-tervezés. Egy kiindulási pluripotens sejt, és a végeredményként kapott differenciált sejt állapotainak betáplálásával képet kaphatunk a differenciációs folyamat legvalószínűbb útjairól, vagy hasonlóképp egy adott daganattípus kialakulási mechanizmusáról. Amennyiben köztes állapotokat is ismerünk, az útvonalbecslés tovább finomítható. Használhatunk beavatkozás-tervezést ideghálózatokban az emléknymok kialakulásának követésére, vagy fehérjeszerkezet-hálózatokban a várható konformációváltozás becslésére.

A beavatkozás-tervezési folyamat más hálózatok esetében is jól használhatónak bizonyulhat. Ennek segítségével például meghatározható lehet, hogy bankok hálózatában milyen beavatkozással kerülhető el egy válság, vagy milyen hatások válthatják azt ki. Ökológiai, táplálékláncokat leíró hálózatokban megtalálható lehet egy olyan beavatkozás, mely segít egy megbillent egyensúlyú populációt helyrebillenteni, vagy azt a "biológiai növényvédőszer", mely egy mezőgazdasági termék produktumát úgy tudja maximalizálni, hogy az élővilágot nem veszélyezteti [137].

A Turbine iterált szimulációs rendszerével nem csak a bemutatott aktraktorkeresésre, beavatkozástervezésre vagy ezek kombinációjára van lehetőség, egyéb tanuló algoritmusok megvalósítása is könnyen lehetséges. Amennyiben ismerjük a kezdőállapotot, a beavatkozásokat, és az azok hatására létrejött végállapotot, esetleg a kezdő-, és a végállapot közötti átmenet bizonyos pontjait is, akkor iterált szimulációk segítségével lehetséges a hálózat struktúrájának finomítása is (inferencia), azaz például a kinetikai paraméterek behangolása az ismert kísérleti eredmények alapján, vagy a hálózatból hiányzó pontok vagy élek legvalószínűbb helyzetének predikciója. Molinelli és mtsai. [120] megmutatták, hogy szisztematikus perturbációs kísérletek sorozatával és a Bayes-féle belief propagation algoritmus segítségével ilyen inferenciával kisebb hálózatok teljes egészében, üres gráfból kiindulva felépíthetőek.

Amennyiben a tervezett beavatkozásaink illetve kutatásaink nem kizárólag csak elméleti célokat szolgálnak, elengedhetetlen, hogy képet kapjunk az eredmények megbízhatóságáról, stabilitásáról, hiszen egyrészt a valós rendszer nem teljesen egyezik meg a rendszer modelljével, másrészt a mérési pontatlanságok, illetve a rendszer zaja mind szisztematikus, mind véletlen hibákat vihetnek az elemzésbe. A problémát tovább nehezíti, hogy a pontosabb modellek általában már nem diszkrét értékűek (azaz az aktivitás szintje folytonosan változhat), így könnyen találhatunk kaotikus viselkedést a rendszerünkben¹⁰, mely annak reakcióját rendkívül érzékenyvé teszi a kezdőfeltételekre az ismert "pillangó-hatást" kiváltva [138].

A készített beavatkozások megbízhatóságát többféleképpen is biztosíthatjuk. Tervezhetünk az eddig bemutatott módszerekkel determinisztikusan egy beavatkozást, majd az elkészült beavatkozás hatását a kezdőfeltételek, vagy maga a hálózat kisebb-nagyobb mértékű módosításai után, többször szimuláljuk. A célállapotba eljutó módosított kezdőfeltételek száma vagy az összes szimuláció átlagos távolsága a célállapottól is megbízhatósági mérőszám lehet.

Ennél több gépidő feláldozásával tervezhetünk kifejezetten megbízhatóságra is, esetlegesen több beavatkozási pont igénybevételével. Ekkor a kezdőfeltételek vagy a hálózat módosításainak hatását minden jelölt vizsgálata során szimuláljuk, és a végleges fitness értékben az így kiszámolt megbízhatósági mérőszámot is szerepeltethetjük.

Úgy gondolom, hogy a megbízhatósági vizsgálatokkal kapcsolatban hasznos tudást emelhetünk át egy olyan területről, ahol szintén modellezéssel, illetve szimulációkkal foglalkoznak, és a terület legendás pontatlansága okán a megbízhatósági vizsgálatokban komoly tapasztalat gyűlt össze - a meteorológia, azaz az időjárás-előrejelző modellek megbízhatóságával kapcsolatban. Itt ugyan beavatkozásokról szó sincs, de az ismert tény, hogy pár napnál távolabbi előrejelzés legtöbbször csak jelentős pontatlansággal végezhető, ezért a szimulációkat bizonyos időközönként újrafuttatják az új mérési adatok fényében. A legtöbbször egyszerű fáklyadiagramokat alkalmaznak, azaz az első megközelítéshez hasonlóan a kezdőfeltételek kis változtatásaival 50 különböző szimulációt futtatnak¹¹,

¹⁰Diszkrét értékű, korlátos rendszerekben nincs kaotikus viselkedés, mivel a lehetséges állapotok száma véges. Ezért nem tud az a helyzet előállni, mint a Lorenz-attraktor esetében, hogy egy pálya egy másikat tetszőlegesen megközelít, de nem tökéletesen azonos értékeket vesz fel.

¹¹Forrás: http://met.hu/ismertetok/Valoszinusegi_elorejelzes_alapjai.pdf

majd az egyes időjárás-paraméterek szimulált értékeinek különböző kvantiliseit különböző színekkel ábrázolva jelenítik meg az időjárás-előrejelzés várható bizonytalanságát. Nagyon széttartó "fáklya" esetén az előrejelzés biztonsága alacsony, összetartó fáklya esetén magasabbnak mondható. Ezen módszer használatával a tervezett beavatkozások/terápiák bizonytalanságára egy időfüggő becslést kapunk, mely arra is alkalmas lehet, hogy definiáljunk egy "újramérési időt", azaz egy olyan időpontot, amikor a hálózat egyes markereinek állapotát újra megvizsgálva, és a szimulációkat már ebből a kezdőpontból futtatva igazíthatunk a terápián, hogy az bizonyosan a megfelelő irányba vigye a sejteket, és ily módon a kívánt hatást váltsa ki.

7. Következtetések

Munkám legfontosabb eredményei a következő pontokban foglalhatóak össze:

- Elkészítettem a Turbine nevű gyors, általános hálózatszimulációs rendszert, mely rendkívül nagy hálózatok hatékony szimulációjára képes, ezen kívül olyan, magasabb szintű dinamikai vizsgálatok is megvalósíthatóak a program használatával, mint az attraktorkeresés vagy a beavatkozás-tervezés.
- Készítettem a programhoz egy általános dinamikai modellt, melyet közlekedőedény-modellnek neveztem el, illetve egy specifikus, jelátviteli hálózatok pontosabb elemzésére képes jelátvitel-dinamikai modellt.
- Kísérleteim alapján megállapítottam, hogy a rendszerek viselkedésének leírásához a legnagyobb attraktorok ismerete elég lehet azok számától függetlenül, mivel a relatív kis attraktorok a természetes szintnek megfelelő zaj hatására destabilizálódnak.
- Megmutattam, hogy sok, daganatokban jellemző diszreguláció úgy változtatta meg egy sejt jelátviteli hálózatának állapotterét, hogy a patológiás attraktor vált a legnagyobb méretűvé.
- Végül ezen diszregulációkat alkalmazva a használt jelátviteli hálózatra megmutattam, hogy a beavatkozás-tervező eszköz segítségével racionálisan tervezhetőek olyan beavatkozások, melyek a diszregulált sejtet apoptózisba viszik úgy, hogy közben az egészséges sejteket a szimulációk tanúsága szerint nem, vagy kevésbé károsítják.

Eredményeim részletezésének első pontjaként a közlekedőedény-modellnek nevezett általános dinamikai modell segítségével definiáltam egy terjesztési központiság nevű, hálózatos központisági mérőszámot.

A terjesztési központiság mérőszám hatásosságát először két fehérje, az *E. coli* metionil-tRNS szintetáz és a házinyúl citokróm P450 2B4 fehérjeszerkezet-hálózatainak nyílt

és zárt konformációján a különféle másodlagos struktúrák átlagos terjesztési központiságának kiszámításával teszteltem. Az eredmények a vártak megfelelően a rendezettebb szerkezetek esetén átlag felettinek (csak az α -hélixek esetén szignifikánsan) bizonyultak, míg a rendezetlen másodlagos szerkezetek esetében átlag alattinak. További vizsgálatként különféle típusú stresszeknek kitett élesztő fehérje-fehérje kölcsönhatás-hálózatát vizsgáltam, ahol a kiemelkedően változó terjesztési központiságú fehérjék az élesztőben ismert stresszreakcióknak megfelelő választ mutatták.

Megmutattam, hogy ezen terjesztési központiság mérőszám legnagyobb változásai az *E. coli* metionil-tRNS szintetáz fehérjeszerkezet-hálózatában a fehérje aktív centrumával illetve a két tRNS-kötőhellyel estek egybe. Ebből az a következtetés vonható le, hogy a fehérje szerkezete a tRNS bekötődése során úgy módosulhatott, hogy az a fehérje két távoli tRNS-kötőhelye közti energiaátadást megkönnyítse. Ez az eredmény két másik népszerű centralitás-mérőszám használatával nem volt reprodukálható. Ezek alapján az újonnan bevezetett terjesztési központiság mérőszám változásai olyan fontos hatásokat is meg tudnak jeleníteni, melyre a többi vizsgált centralitás-mérőszám nem volt képes.

A jelátviteli hálózatok működésének leírására készített specifikus modell helyességét mások jelátviteli szimulációs eredményeinek reprodukálásával ellenőriztem.

A készített attraktorkereső és beavatkozás-tervező algoritmusok működését egy T-LGL sejt jelátvitelét leíró hálózaton teszteltem a teljes, minden állapotra kiterjedő attraktorkereséssel szemben. Az eredmények megmutatták, hogy a Turbine nem teljes vizsgálata is képes volt megtalálni a rendszer összes attraktorát, illetve a hálózatot leíró cikkben bemutatott beavatkozásokat.

Egy daganatos sejtek jellemzően megváltozó jelátviteli útvonalait leíró hálózatban az attraktorokhoz rózsaszín zaj hozzáadásával megmutattam, hogy mutációk nélkül csak az egészséges, apoptotikus attraktor stabil már 10%-nyi zaj mellett is. Ez azt mutatja, hogy a sejtes folyamatok működésének leírásához elég lehet a legnagyobb attraktorokat megtalálni, mivel a kis attraktorok könnyen destabilizálódnak, illetve azt is, hogy a természetes zaj segíthet az egészséges állapot stabilizálásában.

A beavatkozás-tervező és az attraktorkereső algoritmus összevonásával készítettem egy olyan eszközt, mely adott feltételeknek megfelelő attraktort képes tervezni egy hálózatba, annak állapotterének módosításával.

Megvizsgáltam, hogy milyen állapotér-módosítások vezetnek olyan sejtek kialakulásához, melyekben a legnagyobb attraktor proliferál. Ezek az állapotér-módosítások megfeleltek ismert daganatos mutációknak, illetve azok kombinációjának. Ebből arra következtettem, hogy a daganatos betegségek leírásához az egészséges eset attraktorainak vizsgálata kevés, a mutációk által megváltoztatott állapotér figyelembevétele szükséges.

A beavatkozás-tervező program segítségével olyan beavatkozásokat terveztem, melyek az előzőekben módosított állapotérben is képesek voltak a daganatos sejtek apoptózisát elérni a szimuláció során. Ezek a beavatkozások egybevágtak már alkalmazott vagy ígéretes célzott terápiás célpontokkal.

A beavatkozás-tervező program további módosításával olyan beavatkozásokat is terveztem, melyek úgy képesek a daganatos sejtek apoptózisát okozni, hogy eközben az egészséges sejteket a szimuláció során nem viszik apoptózisba. Ezen beavatkozások közt igen gyakran szerepelt a VHL fehérje, melynek ideiglenes gátlása ezért szinergikus módon esetleg biztosíthatja más célzott terápiák mellékhatásainak csökkentését, illetve eddig toxikus terápiák kezelésben való alkalmazhatóságát.

Végül megmutattam, hogy a kRAS gén mutációja a receptor tirozin kináz gátlás alapú beavatkozások hatékonyságának elvesztését okozta a szimulációk során, összhangban a tapasztalati megfigyeléssel. Terveztem továbbá egy olyan beavatkozást, mely a P53 funkcióvesztő mutációja és konstitutívan aktív kRAS mellett is apoptózisba tudja vinni a szimulált daganatos sejtet.

8. Összefoglalás

A hálózatok igen egyszerű és hatásos absztrakciót szolgáltatnak komplex rendszerek működésének leírására. Biológiai rendszerekből több módon is készíthetünk hálózatokat, melyeket a dolgozatban bemutattam. Ezen hálózatok időbeli, dinamikus viselkedésének számítógépes szimulációja, és a szimulációk segítségével alkotott további vizsgálati módszerek képezték disszertációm fő témáját. Elsőként egy általános hálózat-dinamikai modellt mutattam be, melynek intuíciója a közlekedőedények viselkedéséből származtatható. Az ezen modell segítségével készült terjesztési központiság mérőszámot fehérjeszerkezet, és fehérje-fehérje kölcsönhatás-hálózatokon vizsgáltam, és megmutattam, hogy a terjesztési központiság ezen hálózatokban a kiemelkedő fontosságú elemeket két ismert központisági mérőszámnál pontosabban emelte ki. Ezután részleteztem az iterált szimulációk segítségével kialakított két új módszert, melyek hálózatokban való attraktorkeresésre és beavatkozástervezésre használhatóak. Ezen módszereket egy T-sejtes leukémia jelátviteli hálózatát leíró rendszeren ellenőriztem egy specifikus, jelátvitel szimulációjára kidolgozott modell segítségével, majd áttértem egy általánosabb, daganatokban gyakran érintett jelátviteli útvonalakat leíró hálózatra. Ezen a hálózaton a természetes állapotnak megfelelőnek tekinthető zaj hozzáadásával megmutattam, hogy a legnagyobb attraktorok megtalálása elég lehet a hálózat működésének leírásához, mivel a kis attraktorok természetes szerkezet zajszintjei mellett instabillá váltak. A dolgozat további részében tervezetten készítettem olyan mutációkat, melyek a rendszer állapotterét úgy változtatták meg, hogy a proliferáló állapot vált a domináns attraktorrá, és megmutattam, hogy ezek a változtatások megegyeztek egyes daganatos sejtekben ismert elváltozásokkal. Végül az így megváltozott állapotterű rendszerekhez terveztem olyan beavatkozásokat, melyek azt a szimulációk szerint az egészséges állapotba vissza tudták vezetni, többek közt úgy is, hogy eközben az egyébként egészséges sejteket a szimuláció során nem károsították.

9. Summary

Networks present a simple and useful abstraction to describe the behavior of complex systems. Networks can also be constructed from biological systems in multiple different ways. The topic of my thesis was simulating the time-dependent dynamic behavior of these biologically-inspired complex networks. First, I have presented a general network dynamics model motivated by the behavior of communicating vessels. I have defined a dynamic centrality measure termed "perturbation centrality" to mathematically describe the dynamic response of networks to generic perturbations. I have analyzed the perturbation centrality measure both on protein structure networks and protein-protein interaction networks, observing that this new measure has the potential to highlight important parts of the network which were otherwise undetected by two popular centrality measures, betweenness and closeness centrality. Thereafter, I presented in detail the novel simulation-based tools for attractor search and intervention design. I have validated these methods on a network which has already been analyzed in detail in the literature: the signaling network of T-LGL leukemia, using a dynamic model built specifically for the dynamic analysis of signaling networks. Finally, I have analyzed a network of signaling pathways involved in cell cycle progression. I have observed that only the largest attractors of the network were stable at noise levels comparable to those in the natural system. I have rationally designed mutation sets which changed the state space of the network in a manner that the largest attractor was pathologically proliferating, and pointed out that these dysregulations correspond to known mutations in cancer. I have also shown that it is possible to rationally design interventions, which could guide back these pathologically proliferating systems to an apoptotic state while keeping healthy cells – according to the simulations – unharmed.

10. Irodalomjegyzék

- [1] Chung F, Lu L (2002) Connected components in random graphs with given expected degree sequences. *Ann Comb* 6: 125–145.
- [2] Erdős P, Rényi A (1959) On random graphs. *Publ Math Debrecen* 6: 290–297.
- [3] Travers J, Milgram S (1969) An experimental study of the small world problem. *Sociometry* 32: 425–443.
- [4] Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393: 440–2.
- [5] Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286: 509–12.
- [6] Price DJ (1965) Networks of scientific papers. *Science* 149: 510–5.
- [7] Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL (2000) The large-scale organization of metabolic networks. *Nature* 407: 651–4.
- [8] Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci U S A* 99: 7821–6.
- [9] Szalay-Bekő M (2014) Biológiai hálózatok átfedő modularizálását végző számítógépes programok és azok alkalmazási területei. Ph.D. thesis, Szegedi Tudományegyetem.
- [10] Shen HW, Cheng XQ, Guo JF (2009) Quantifying and identifying the overlapping community structure in networks. *J Stat Mech Theory Exp* 2009: P07042.
- [11] Sabidussi G (1966) The centrality index of a graph. *Psychometrika* 31: 581–603.
- [12] Freeman L (1977) A set of measures of centrality based on betweenness. *Sociometry* 40: 35 – 41.

- [13] Page L, Brin S, Motwani R, Winograd T (1998) The PageRank citation ranking: bringing order to the web. *Stanford Digit Libr Technol Proj* : 1–17.
- [14] Ipsen ICF, Kirkland S (2006) Convergence analysis of a PageRank updating algorithm by Langville and Meyer. *SIAM J* 27: 952–967.
- [15] Kannan N, Vishveshwara S (1999) Identification of side-chain clusters in protein structures by a graph spectral method. *J Mol Biol* 292: 441–64.
- [16] Doncheva NT, Klein K, Domingues FS, Albrecht M (2011) Analyzing and visualizing residue networks of protein structures. *Trends Biochem Sci* 36: 179–82.
- [17] Böde C, Kovács IA, Szalay MS, Palotai R, Korcsmáros T, Csermely P (2007) Network analysis of protein dynamics. *FEBS Lett* 581: 2776–82.
- [18] Ghosh A, Vishveshwara S (2007) A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proc Natl Acad Sci U S A* 104: 15711–6.
- [19] Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42: D199–205.
- [20] Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahrén D, Tsoka S, Darzentas N, Kunin V, López-Bigas N (2005) Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Res* 33: 6083–9.
- [21] Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, Haraldsdottir H, Mo ML, Rolfsson O, Stobbe MD, Thorleifsson SG, Agren R, Bölling C, Bordel S, Chavali AK, Dobson P, Dunn WB, Endler L, Hala D, Hucka M, Hull D, Jameson D, Jamshidi N, Jonsson JJ, Juty N, Keating S, Nookaew I, Le Novère N, Malys N, Mazein A, Papin JA, Price ND, Selkov E, Sigurdsson MI, Simeonidis E, Sonnenschein N, Smallbone K, Sorokin A, van Beek JHGM, Weichart D, Goryanin I, Nielsen J, Westerhoff HV, Kell DB, Mendes P, Palsson BO (2013) A

- community-driven global reconstruction of human metabolism. *Nat Biotechnol* 31: 419–25.
- [22] Prigent S, Collet G, Dittami SM, Delage L, de Corny FE, Dameron O, Eveillard D, Thiele S, Cambefort J, Boyen C, Siegel A, Tonon T (2014) The genome-scale metabolic network of *Ectocarpus siliculosus* (EctoGEM): a resource to study brown algal physiology and beyond. *Plant J* : in press, doi: 10.1111/tpj.12627.
- [23] Zhang R, Shah MV, Yang J, Nyland SB, Liu X, Yun JK, Albert R, Loughran TP (2008) Network model of survival signaling in large granular lymphocyte leukemia. *Proc Natl Acad Sci U S A* 105: 16308–13.
- [24] Fumiã HF, Martins ML (2013) Boolean network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes. *PLoS ONE* 8: e69008.
- [25] Atlas of Cancer Signaling Networks. URL <https://acsn.curie.fr/>.
- [26] Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res* 32: D277–80.
- [27] Genoud T, Trevino Santa Cruz MB, Métraux JP (2001) Numeric simulation of plant signaling networks. *Plant Physiol* 126: 1430–7.
- [28] Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4: 17.
- [29] Chen L, Liu R, Liu ZP, Li M, Aihara K (2012) Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep* 2: 342.
- [30] Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S, Boucher L, Winter A, Stark C, Nixon J, Ramage L, Kolas N, O'Donnell L, Reguly T, Breitkreutz A, Sellam A, Chen D, Chang C, Rust J, Livstone M, Oughtred R, Dolinski K, Tyers M (2013) The BioGRID interaction database: 2013 update. *Nucleic Acids Res* 41: D816–D823.

- [31] Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguéz P, Doerks T, Stark M, Müller J, Bork P, Jensen LJ, von Mering C (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 39: D561–D568.
- [32] Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340: 245–6.
- [33] Young KH (1998) Yeast two-hybrid: so many interactions, (in) so little time... *Biol Reprod* 58: 302–11.
- [34] Abbott LF (1999) Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Res Bull* 50: 303–4.
- [35] Hines ML, Carnevale NT (1997) The NEURON simulation environment. *Neural Comput* 9: 1179–209.
- [36] White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci* 314: 1–340.
- [37] Rankin CH, Beck CD, Chiba CM (1990) *Caenorhabditis elegans*: a new model system for the study of learning and memory. *Behav Brain Res* 37: 89–92.
- [38] Chiang AS, Lin CY, Chuang CC, Chang HM, Hsieh CH, Yeh CW, Shih CT, Wu JJ, Wang GT, Chen YC, Wu CC, Chen GY, Ching YT, Lee PC, Lin CY, Lin HH, Wu CC, Hsu HW, Huang YA, Chen JY, Chiang HJ, Lu CF, Ni RF, Yeh CY, Hwang JK (2011) Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr Biol* 21: 1–11.
- [39] Hintiryan H, Gou L, Zingg B, Yamashita S, Lyden HM, Song MY, Grewal AK, Zhang X, Toga AW, Dong HW (2012) Comprehensive connectivity of the mouse main olfactory bulb: analysis and online digital atlas. *Front Neuroanat* 6: 30.
- [40] Waddington CH (1957) *The strategy of the genes*. London: George Allen & Unwin.

- [41] Kauffman S (1969) Homeostasis and Differentiation in Random Genetic Control Networks. *Nature* 224: 177–178.
- [42] Kauffman S (1971) Differentiation of malignant to benign cells. *J Theor Biol* 31: 429–51.
- [43] Tsujimoto Y, Finger LR, Yunis J, Nowell PC, Croce CM (1984) Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. *Science* 226: 1097–9.
- [44] Oltvai ZN, Milliman CL, Korsmeyer SJ (1993) Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell* 74: 609–19.
- [45] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C,

Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Feder-spiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nord-siek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglu S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowski J (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.

- [46] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA,

Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X (2001) The sequence of the human genome. *Science* 291: 1304–51.

- [47] Fauré A, Naldi A, Chaouiya C, Thieffry D (2006) Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics* 22: e124–31.

- [48] Saadatpour A, Albert I, Albert R (2010) Attractor analysis of asynchronous Boolean models of signal transduction networks. *J Theor Biol* 266: 641–56.
- [49] Saadatpour A, Wang RS, Liao A, Liu X, Loughran TP, Albert I, Albert R (2011) Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia. *PLoS Comput Biol* 7: e1002267.
- [50] Russell SJ, Norvig P (2003) *Artificial intelligence: a modern approach* .
- [51] Fraser A (1957) Simulation of genetic systems by automatic digital computers. I. Introduction. *Aust J Biol Sci* 10: 484 – 491.
- [52] Harrison PM, Kumar A, Lang N, Snyder M, Gerstein M (2002) A question of size: the eukaryotic proteome and the problems in defining it. *Nucleic Acids Res* 30: 1083–90.
- [53] Huang S, Ernberg I, Kauffman S (2009) Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. *Semin Cell Dev Biol* 20: 869–76.
- [54] Creixell P, Schoof EM, Erler JT, Linding R (2012) Navigating cancer network attractors for tumor-specific therapy. *Nat Biotechnol* 30: 842–8.
- [55] Mechulam Y, Schmitt E, Maveyraud L, Zelwer C, Nureki O, Yokoyama S, Konno M, Blanquet S (1999) Crystal structure of *Escherichia coli* methionyl-tRNA synthetase highlights species-specific features. *J Mol Biol* 294: 1287–97.
- [56] Crepin T, Schmitt E, Mechulam Y, Sampson PB, Vaughan MD, Honek JF, Blanquet S (2003) Use of analogues of methionine and methionyl adenylate to sample conformational changes during catalysis in *Escherichia coli* methionyl-tRNA synthetase. *J Mol Biol* 332: 59–72.
- [57] Mihalik A, Csermely P (2011) Heat shock partially dissociates the overlapping modules of the yeast protein-protein interaction network: a systems level model of adaptation. *PLoS Comput Biol* 7: e1002187.

- [58] Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95: 717–28.
- [59] Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 11: 4241–57.
- [60] Kinzler KW, Bigner SH, Bigner DD, Trent JM, Law ML, O’Brien SJ, Wong AJ, Vogelstein B (1987) Identification of an amplified, highly expressed gene in a human glioma. *Science* 236: 70–3.
- [61] Ruiz i Altaba A (1999) Gli proteins encode context-dependent positive and negative functions: implications for development and disease. *Development* 126: 3205–16.
- [62] Antal MA, Böde C, Csermely P (2009) Perturbation waves in proteins and protein networks: applications of percolation and game theories in signaling and drug design. *Curr Protein Pept Sci* 10: 161–72.
- [63] Fodor G (2000) *Hálózatok és rendszerek analízise*. Budapest: Műegyetemi Kiadó.
- [64] McDonald D, Waterbury L, Knight R, Betterton MD (2008) Activating and inhibiting connections in biological network dynamics. *Biol Direct* 3: 49.
- [65] Terfve C, Cokelaer T, Henriques D, MacNamara A, Goncalves E, Morris MK, van Iersel M, Lauffenburger DA, Saez-Rodriguez J (2012) CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Syst Biol* 6: 133.
- [66] R Core Team (2012). *R: A language and environment for statistical computing*. URL www.r-project.org.
- [67] Schrödinger LLC (2010). *The PyMOL molecular graphics system*. URL www.pymol.org.

- [68] Mizock BA (1995) Alterations in carbohydrate metabolism during stress: a review of the literature. *Am J Med* 98: 75–84.
- [69] Lindquist S (1981) Regulation of protein synthesis during heat shock. *Nature* 293: 311–314.
- [70] Wang RS, Albert R (2011) Elementary signaling modes predict the essentiality of signal transduction network components. *BMC Syst Biol* 5: 44.
- [71] Rose MG (2004) T-cell large granular lymphocyte leukemia and related disorders. *Oncologist* 9: 247–258.
- [72] Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A* 99: 12795–800.
- [73] Szalay KZ, Nussinov R, Csermely P (2014) Attractor structures of signaling networks: consequences of different conformational barcode dynamics and their relations to network-based drug design. *Mol Inform* 33: 463–468.
- [74] Rodin SN, Rodin AS, Juhasz A, Holmquist GP (2002) Cancerous hypermutagenesis in p53 genes is possibly associated with transcriptional bypass of DNA lesions. *Mutat Res* 510: 153–68.
- [75] Morgan DO (2006) *Cell cycle: principles of control*, volume 80. London: New Science Press.
- [76] Chiu YT, Liu J, Tang K, Wong YC, Khanna KK, Ling MT (2012) Inactivation of ATM/ATR DNA damage checkpoint promotes androgen induced chromosomal instability in prostate epithelial cells. *PLoS ONE* 7: e51108.
- [77] Tamura M, Gu J, Matsumoto K, Aota S, Parsons R, Yamada KM (1998) Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN. *Science* 280: 1614–7.
- [78] Waite KA, Eng C (2002) Protean PTEN: form and function. *Am J Hum Genet* 70: 829–44.

- [79] Kim MS, Jeong EG, Yoo NJ, Lee SH (2008) Mutational analysis of oncogenic AKT E17K mutation in common solid cancers and acute leukaemias. *Br J Cancer* 98: 1533–5.
- [80] Shoji K, Oda K, Nakagawa S, Hosokawa S, Nagae G, Uehara Y, Sone K, Miyamoto Y, Hiraike H, Hiraike-Wada O, Nei T, Kawana K, Kuramoto H, Aburatani H, Yano T, Taketani Y (2009) The oncogenic mutation in the pleckstrin homology domain of AKT1 in endometrial carcinomas. *Br J Cancer* 101: 145–8.
- [81] Lindhurst MJ, Sapp JC, Teer JK, Johnston JJ, Finn EM, Peters K, Turner J, Cannons JL, Bick D, Blakemore L, Blumhorst C, Brockmann K, Calder P, Cherman N, Dearnorff MA, Everman DB, Golas G, Greenstein RM, Kato BM, Keppler-Noreuil KM, Kuznetsov SA, Miyamoto RT, Newman K, Ng D, O'Brien K, Rothenberg S, Schwartzentruber DJ, Singhal V, Tirabosco R, Upton J, Wientroub S, Zackai EH, Hoag K, Whitewood-Neal T, Robey PG, Schwartzberg PL, Darling TN, Tosi LL, Mullikin JC, Biesecker LG (2011) A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N Engl J Med* 365: 611–9.
- [82] Tsujimoto Y, Cossman J, Jaffe E, Croce CM (1985) Involvement of the bcl-2 gene in human follicular lymphoma. *Science* 228: 1440–3.
- [83] Hooper SD, Jiao X, Sundström E, Rehman FL, Tellgren-Roth C, Sjöblom T, Cavelier L (2012) Sequence based analysis of U-2973, a cell line established from a double-hit B-cell lymphoma with concurrent MYC and BCL2 rearrangements. *BMC Res Notes* 5: 648.
- [84] Meijerink JP, Smetsers TF, Slöetjes AW, Linders EH, Mensink EJ (1995) Bax mutations in cell lines derived from hematological malignancies. *Leukemia* 9: 1828–32.
- [85] Rampino N, Yamamoto H, Ionov Y, Li Y, Sawai H, Reed JC, Perucho M (1997) Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science* 275: 967–9.

- [86] Niida H, Murata K, Shimada M, Ogawa K, Ohta K, Suzuki K, Fujigaki H, Khaw AK, Banerjee B, Hande MP, Miyamoto T, Miyoshi I, Shirai T, Motoyama N, Delhase M, Appella E, Nakanishi M (2010) Cooperative functions of Chk1 and Chk2 reduce tumour susceptibility in vivo. *EMBO J* 29: 3558–70.
- [87] Starczynowski DT, Trautmann H, Pott C, Harder L, Arnold N, Africa JA, Leeman JR, Siebert R, Gilmore TD (2007) Mutation of an IKK phosphorylation site within the transactivation domain of REL in two patients with B-cell lymphoma enhances REL's in vitro transforming activity. *Oncogene* 26: 2685–94.
- [88] Valentine R, Dawson CW, Hu C, Shah KM, Owen TJ, Date KL, Maia SP, Shao J, Arrand JR, Young LS, O'Neil JD (2010) Epstein-Barr virus-encoded EBNA1 inhibits the canonical NF-kappaB pathway in carcinoma cells by inhibiting IKK phosphorylation. *Mol Cancer* 9: 1.
- [89] Vlantis K, Wullaert A, Sasaki Y, Schmidt-Suprian M, Rajewsky K, Roskams T, Pasparakis M (2011) Constitutive IKK2 activation in intestinal epithelial cells induces intestinal tumors in mice. *J Clin Invest* 121: 2781–93.
- [90] Sunami Y, Wirth T (2011) Intestinal carcinogenesis: IKK can go all the way. *J Clin Invest* 121: 2551–3.
- [91] Souers AJ, Levenson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, Dayton BD, Ding H, Enschede SH, Fairbrother WJ, Huang DCS, Hymowitz SG, Jin S, Khaw SL, Kovar PJ, Lam LT, Lee J, Maecker HL, Marsh KC, Mason KD, Mitten MJ, Nimmer PM, Oleksijew A, Park CH, Park CM, Phillips DC, Roberts AW, Sampath D, Seymour JF, Smith ML, Sullivan GM, Tahir SK, Tse C, Wendt MD, Xiao Y, Xue JC, Zhang H, Humerickhouse RA, Rosenberg SH, Elmore SW (2013) ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med* 19: 202–8.
- [92] Raimondi C, Falasca M (2011) Targeting PDK1 in cancer. *Curr Med Chem* 18: 2763–9.

- [93] Lee DF, Hung MC (2008) Advances in targeting IKK and IKK-related kinases for cancer therapy. *Clin Cancer Res* 14: 5656–62.
- [94] Crowell JA, Steele VE, Fay JR (2007) Targeting the AKT protein kinase for cancer chemoprevention. *Mol Cancer Ther* 6: 2139–48.
- [95] Knight ZA, Shokat KM (2007) Chemically targeting the PI3K family. *Biochem Soc Trans* 35: 245–9.
- [96] Karve S, Werner ME, Sukumar R, Cummings ND, Copp JA, Wang EC, Li C, Sethi M, Chen RC, Pacold ME, Wang AZ (2012) Revival of the abandoned therapeutic wortmannin by nanoparticle drug delivery. *Proc Natl Acad Sci U S A* 109: 8230–5.
- [97] Madonna G, Ullman CD, Gentilcore G, Palmieri G, Ascierto PA (2012) NF- κ B as potential target in the treatment of melanoma. *J Transl Med* 10: 53.
- [98] Tory K, Brauch H, Linehan M, Barba D, Oldfield E, Filling-Katz M, Seizinger B, Nakamura Y, White R, Marshall FF (1989) Specific genetic change in tumors associated with von Hippel-Lindau disease. *J Natl Cancer Inst* 81: 1097–101.
- [99] Buckley DL, Van Molle I, Gareiss PC, Tae HS, Michel J, Noblin DJ, Jorgensen WL, Ciulli A, Crews CM (2012) Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1 α interaction. *J Am Chem Soc* 134: 4465–8.
- [100] Brózik A, Hegedűs C, Erdei Z, Hegedűs T, Özvegy Laczka C, Szakács G, Sarkadi B (2011) Tyrosine kinase inhibitors as modulators of ATP binding cassette multi-drug transporters: substrates, chemosensitizers or inducers of acquired multidrug resistance? *Expert Opin Drug Metab Toxicol* 7: 623–42.
- [101] Schutz FAB, Je Y, Choueiri TK (2011) Hematologic toxicities in cancer patients treated with the multi-tyrosine kinase sorafenib: a meta-analysis of clinical trials. *Crit Rev Oncol Hematol* 80: 291–300.
- [102] Cui R, Chu L, Liu ZQ, Xiao YY, Zhu XL, Chen YJ, Xu Q (2014) Hematologic toxicity assessment in solid tumor patients treated with cetuximab: A pooled analysis of 18 randomized controlled trials. *Int J Cancer* : in press, doi: 10.1002/ijc.29045.

- [103] Plesec TP, Hunt JL (2009) KRAS mutation testing in colorectal cancer. *Adv Anat Pathol* 16: 196–203.
- [104] Korcsmáros T, Szalay MS, Rovó P, Palotai R, Fazekas D, Lenti K, Farkas IJ, Csermely P, Vellai T (2011) Signalogs: orthology-based identification of novel signaling pathway components in three metazoans. *PLoS ONE* 6: e19240.
- [105] Zhou X, Menche J, Barabási AL, Sharma A (2014) Human symptoms-disease network. *Nat Commun* 5: 4212.
- [106] Golub TR (2004) Toward a functional taxonomy of cancer. *Cancer Cell* 6: 107–8.
- [107] Wagner BK, Haggarty SJ, Clemons PA (2004) Chemical genomics: probing protein function using small molecules. *Am J Pharmacogenomics* 4: 313–20.
- [108] Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313: 1929–35.
- [109] Tai Y, Speed TP (2004) Statistical analysis of microarray time course data. In: Nuber UA, editor, *DNA Microarrays*, New York: Taylor & Francis Group. 2nd edition, pp. 258–278.
- [110] Calzone L, Fages F, Soliman S (2006) BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge. *Bioinformatics* 22: 1805–7.
- [111] Stojmirovi A, Yu YK, Stojmirović A (2009) ITM Probe: analyzing information flow in protein networks. *Bioinformatics* 25: 2447–9.
- [112] Gong Y, Zhang Z (2007) CellFrame: a data structure for abstraction of cell biology experiments and construction of perturbation networks. *Ann N Y Acad Sci* 1115: 249–66.

- [113] Farkas IJ, Korcsmáros T, Kovács IA, Mihalik A, Palotai R, Simkó GI, Szalay KZ, Szalay-Beko M, Vellai T, Wang S, Csermely P, Szalay-bekő M (2011) Network-based tools for the identification of novel drug targets. *Sci Signal* 4: pt3.
- [114] Szalay K, Csermely P (2013) Perturbation centrality and Turbine: a novel centrality measure obtained using a versatile network dynamics tool. *PLoS ONE* 8: e78059.
- [115] Rothkegel A, Lehnertz K (2012) Conedy: a scientific tool to investigate complex network dynamics. *Chaos* 22: 013125.
- [116] Zhong Q, Simonis N, Li QR, Charloteaux B, Heuze F, Klitgord N, Tam S, Yu H, Venkatesan K, Mou D, Swearingen V, Yildirim MA, Yan H, Dricot A, Szeto D, Lin C, Hao T, Fan C, Milstein S, Dupuy D, Brasseur R, Hill DE, Cusick ME, Vidal M (2009) Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* 5: 321.
- [117] Liu YY, Slotine JJ, Barabási AL (2011) Controllability of complex networks. *Nature* 473: 167–73.
- [118] Kalman RE (1960) On the general theory of control systems. *IRE Trans Autom Control* 4: 100–110.
- [119] Cornelius SP, Kath WL, Motter AE (2013) Realistic control of network dynamics. *Nat Commun* 4: 1942.
- [120] Molinelli EJ, Korkut A, Wang W, Miller ML, Gauthier NP, Jing X, Kaushik P, He Q, Mills G, Solit DB, Pratilas CA, Weigt M, Braunstein A, Pagnani A, Zecchina R, Sander C (2013) Perturbation biology: inferring signaling networks in cellular systems. *PLoS Comput Biol* 9: e1003290.
- [121] Maslov S, Ispolatov I (2007) Propagation of large concentration changes in reversible protein-binding networks. *Proc Natl Acad Sci U S A* 104: 13655–60.
- [122] Danon L, Arenas A, Díaz-Guilera A (2008) Impact of community structure on information transfer. *Phys Rev E* 77: 1–6.

- [123] Miller JC (2009) Percolation and epidemics in random clustered networks. *Phys Rev E* 80: 020901.
- [124] Trpevski D, Tang WKS, Kocarev L (2010) Model for rumor spreading over networks. *Phys Rev E* 81: 056102.
- [125] Cho JH, Meng W, Sato S, Kim EY, Schindelin H, Raleigh DP (2014) Energetically significant networks of coupled interactions within an unfolded protein. *Proc Natl Acad Sci U S A* 111: 12079–12084.
- [126] Zhou Y, Liu S, Song J, Zhang Z (2013) Structural propensities of human ubiquitination sites: accessibility, centrality and local conformation. *PLoS ONE* 8: e83167.
- [127] Tang YR, Sheng ZY, Chen YZ, Zhang Z (2008) An improved prediction of catalytic residues in enzyme structures. *Protein Eng Des Sel* 21: 295–302.
- [128] Palotai R, Szalay MSM, Csermely P (2008) Chaperones as integrators of cellular networks: changes of cellular integrity in stress and diseases. *IUBMB Life* 60: 10–18.
- [129] Zañudo JGT, Albert R (2013) An effective network reduction approach to find the dynamical repertoire of discrete dynamic networks. *Chaos* 23: 025111.
- [130] Baffy G, Loscalzo J (2014) Complexity and network dynamics in physiological adaptation: an integrated view. *Physiol Behav* 131: 49–56.
- [131] Huang S, Kauffman S (2013) How to escape the cancer attractor: rationale and limitations of multi-target drugs. *Semin Cancer Biol* 23: 270–8.
- [132] Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, McLellan MD, McMichael JF, Wallis JW, Lu C, Shen D, Harris CC, Dooling DJ, Fulton RS, Fulton LL, Chen K, Schmidt H, Kalicki-Veizer J, Magrini VJ, Cook L, McGrath SD, Vickery TL, Wendl MC, Heath S, Watson MA, Link DC, Tomasson MH, Shannon WD, Payton JE, Kulkarni S, Westervelt P, Walter MJ, Graubert TA, Mardis ER, Wilson RK, DiPersio JF (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481: 506–10.

- [133] Zhao B, Hemann MT, Lauffenburger DA (2014) Intratumor heterogeneity alters most effective drugs in designed combinations. *Proc Natl Acad Sci U S A* 111: 10773–10778.
- [134] Csermely P, Ágoston V, Pongor S (2005) The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci* 26: 178–82.
- [135] Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, Shah P, Moon YS, Yaquibie A, Kelly N, Le DT, Lipson EJ, Chapman PB, Diaz LA, Vogelstein B, Nowak MA (2013) Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife* 2: e00747.
- [136] Chindelevitch L, Ziemek D, Enayetallah A, Randhawa R, Sidders B, Brockel C, Huang ES (2012) Causal reasoning on biological networks: interpreting transcriptional changes. *Bioinformatics* 28: 1114–21.
- [137] Flint ML, Dreistadt SH (1998) *Natural enemies handbook: the illustrated guide to biological pest control*. Oakland: University of California Press.
- [138] Lorenz EN (1963) Deterministic nonperiodic flow. *J Atmos Sci* 20: 130–141.

Ábrák jegyzéke

1.	Különböző gráf típusok.	8
2.	Watts–Strogatz hálózatok különböző β értékek mellett.	10
3.	A véletlen és a skálafüggetlen hálózatok közötti különbség szemléltetése.	12
4.	Moduláris hálózat szemléltetése.	14
5.	Az élesztő két-hibrid rendszer.	20
6.	Dinamikus rendszer feketedoboz-ábrája.	22
7.	Fázisterek különféle ábrázolásai.	23
8.	A Lorenz-atraktor.	24
9.	A Fauré és mtsai. (2006) és a Fumiã és Martins (2013) által összeállított dinamikus jelátviteli hálózat összehasonlítása.	26
10.	Szélességi gráfbejárás.	28
11.	A lokális keresés problémája és a hegymászó algoritmus.	30
12.	A genetikus algoritmus működése.	31
13.	A Turbine használatának folyamatábrája.	34
14.	A Turbine megjelenítő alkalmazása.	35
15.	A Turbine kommunikációs mechanizmusai.	36
16.	A Turbine szimulátor skálázódása.	37
17.	A Turbine beavatkozás-tervező skálázódása.	41
18.	Különböző központisági mérőszámok által fontosnak jelölt aminosavak az <i>E. coli</i> MetRS hálózatában.	49
19.	A MetRS és a CyP450 aminosavainak terjesztési központiságai másodla- gos szerkezet szerint csoportosítva.	51
20.	Fehérjék hálózatos fontosságának eltérő megváltozása különféle típusú stresszekben.	53
21.	A T-LGL hálózat attraktorai.	56

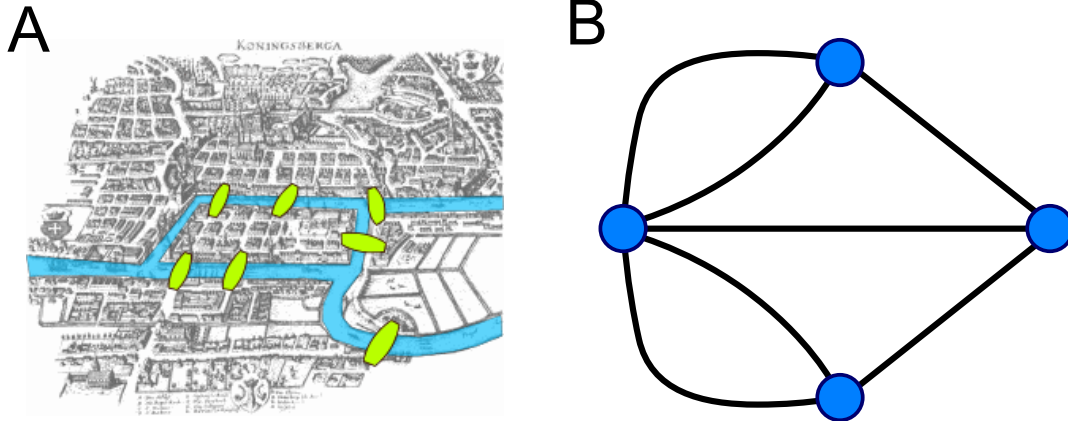
22. A Fumiã és Martins által leírt hálózat attraktorai és viselkedésük rózsaszín zaj hatására.	59
--	----

Táblázatok jegyzéke

1. A Turbine által talált beavatkozások T-LGL leukémia ellen	57
2. A proliferáló attraktor dominanciájához vezető mutációsorozatok	62
3. Mutáció-specifikus beavatkozások	64
4. Nem toxikusra tervezett beavatkozások	66
5. "Ideális" beavatkozások	67

Függelékek

A. Königsberg hídjai



Königsberg (ma Kalinyingrád) a Pregel folyó mentén található város. A XVIII. században hét híd kötötte össze a város két oldalát, és a folyón található két szigetet. Sok akkori polgár töprengett azon, hogy hogy lehet úgy átmenni a város összes hídján, hogy egy hidat se érintsünk kétszer (A tábla). Euler először rájött, hogy maga a szigeteken illetve a partokon megtett út lényegtelen, és ez alapján megrajzolta az első gráfot, a szárazföldeket pontokkal, a hidakat pedig a pontokat összekötő élekkel jelölve (B). Majd úgy gondolkodott, hogy akárhányszor is megyünk be egy pontba, azt – kettő, a kezdeti és a célpont kivételével – el is kell hagynunk valahol, ezért ahhoz, hogy a feladat végrehajtható legyen, a pontok fokszámának – kettő kivételével – párosnak kell lennie. Mivel a fenti gráfban minden pont fokszáma páratlan, a probléma nem megoldható, nincs ilyen út. A gráfok ezen útjait később Euler tiszteletére Euler-útnak nevezték el. A II. világháború során két híd lebombázták a hétből (a középső szigetet és a partot összekötő középső hidakat), így a hidak mai állapotából rajzolt gráfban már van Euler-út. Turisztikai szempontból azonban továbbra sem túl hasznos ez az út, mert annak az egyik szigeten kell kezdődnie, és a másik szigeten kell végződnie, így azoknak, akik nem a szigeten laknak, továbbra is ismételniük kell néhány hidat.

B. A Milgram-kísérlet

Stanley Milgram kisvilág-kísérlete¹² volt az egyik első tudományos bizonyíték a társadalmi hálózatunk kisvilág-tulajdonságára. Milgram Omahában (Nebraska) és Wichitában (Kansas) lakó embereknek küldött egy csomagot, melyben leírja a kísérlet célját, bizonyos alapinformációkat a célszemélyről, és egy papírt, ahova a lánc minden tagja ráírhatta a nevét követési célból. Azért így esett a kezdő-, és célpontokra a választás, mert így mind földrajzilag, mind társadalmilag "messze" levőnek voltak tekinthetők a személyek.

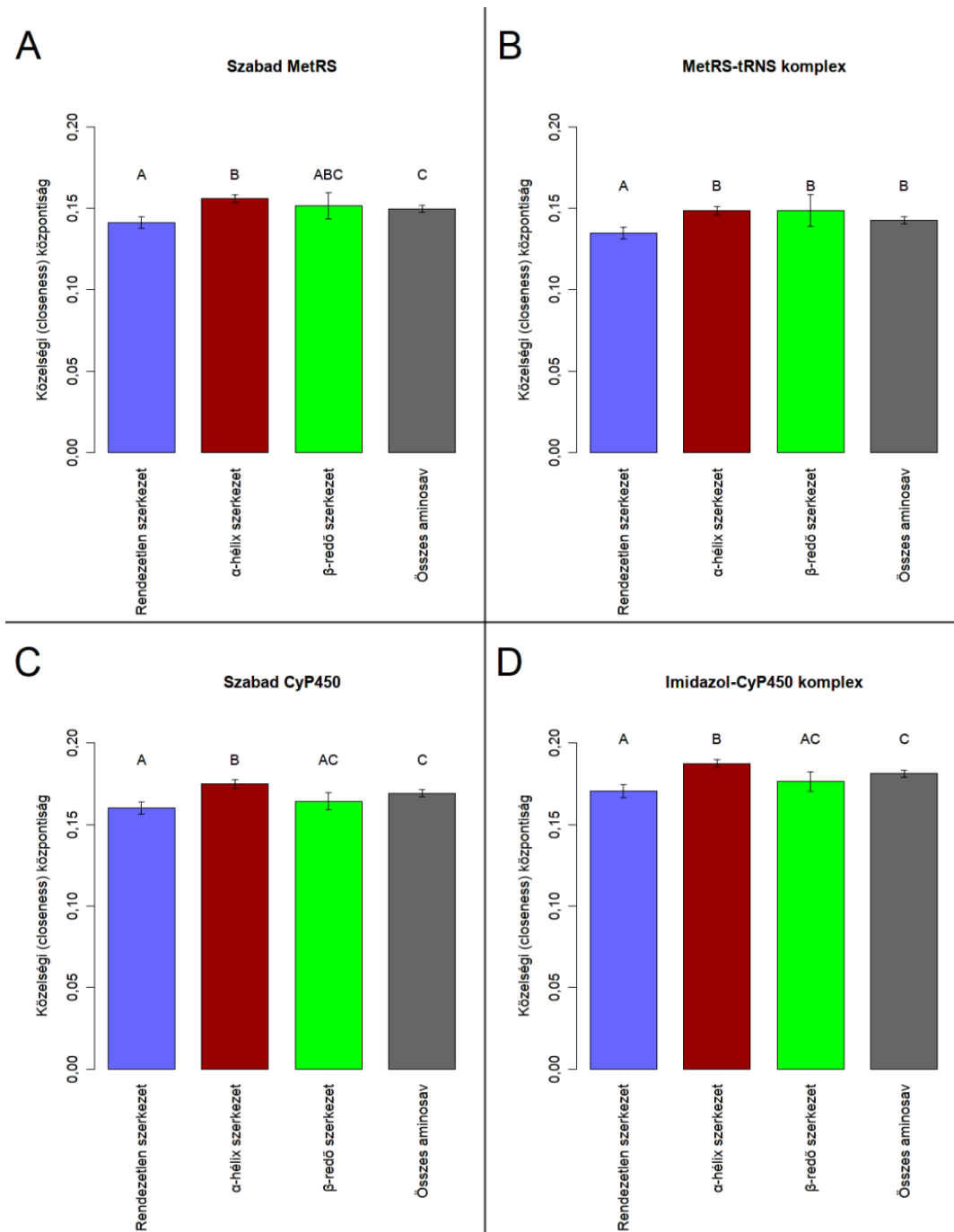
Az első, a csomagban található instrukció az volt, hogy ha személyesen ismeri a célszemélyt (egész pontosan tegeződő viszonyban vannak), akkor küldje el egyből neki a levelet, ráírva a követő lapra saját nevét. Ha nem ismeri a célszemélyt, akkor küldje el az egész csomagot egy ismerősének, akiről úgy gondolja, hogy ő ismerheti. Emellett arra is megkérték, hogy küldjön egy lapot a Harvardra, hogy a lánc megszakadása esetén is tudják követni, hogy hol akadt el.

Sok lánc megszakadt, 296-ból 232 csomag nem jutott el a célig. Az a 64, amely mégis eljutott, azokban a papírokon átlagosan 5 és 6 közötti név szerepelt. Ebből vonták le a kutatók a híres "hat lépés távolság" következtetést.

Napjaink kommunikációs robbanásával ez a távolság számottevően csökkent. A Facebook közel egy milliárd felhasználója közül két tetszőleges felhasználó átlagos távolsága csak 4,74.

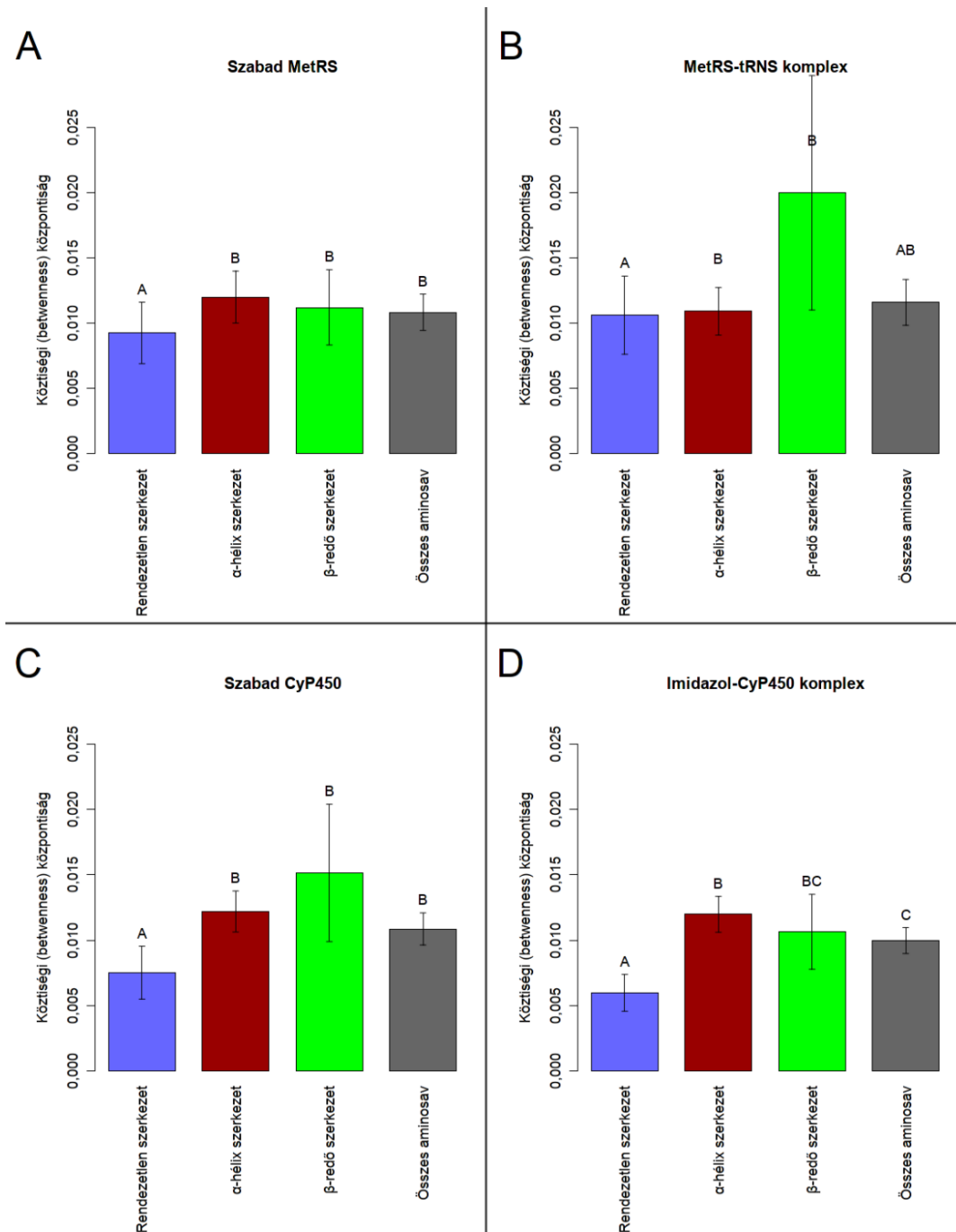
¹²Travers J, Milgram S (1969) An experimental study of the small world problem. *Sociometry* 32: 425–443.

C. A MetRS és a CyP450 aminosavainak closeness központiságai másodlagos szerkezet szerint csoportosítva



A MetRS és a CyP450 aminosavainak **closeness** központiságai másodlagos szerkezet szerint csoportosítva. Az ábra felépítése megegyezik a dolgozat a 19. ábrájával, ebben az esetben closeness értékekre számítva a grafikonokat. A másodlagos szerkezetek közti szórások ebben az esetben voltak a legalacsonyabbak; érdekes módon azonban ezen mérőszám megkülönböztetési képessége is gyengébb volt a terjesztési központiságnál, pl. a B tábla α-hélicei esetében.

D. A MetRS és a Cyp450 aminosavainak betweenness központiságai másodlagos szerkezet szerint csoportosítva



A MetRS és a Cyp450 aminosavainak **betweenness** központiságai másodlagos szerkezet szerint csoportosítva. Az ábra felépítése megegyezik a dolgozat a 19. ábrájával, ebben az esetben betweenness értékekre számítva a grafikonokat. A három vizsgált mérőszám esetében itt tapasztaltuk a legnagyobb szórást a különféle másodlagos szerkezetek közt. A rendezetlen szerkezetek betweenness központiságai az átlagnál többnyire továbbra is szignifikánsan ($0,0016; 0,011; 4,3 \cdot 10^{-5}; 6,8 \cdot 10^{-7}$, a táblák sorrendjében, Mann-Whitney U teszt, $\alpha = 0,00625$ Bonferroni korrekció használatával) alacsonyabbnak bizonyultak, de az α -hélix szerkezetben található aminosavak betweenness központisága már nem mutatott az átlagtól való szignifikáns eltérést.

E. Élesztő fehérje-fehérje kölcsönhatás-hálózatában a legmagasabb terjesztési központiságú pontokban dúsuló GO funkciók nyugalmi és stresszelt állapotokban

A terjesztési központiság értékeit a Turbine program segítségével számítottam, a dúsulási elemzést pedig a g:Profiler program R beépülő moduljának használatával. Szignifikánsnak minősítettem egy GO kulcsszót, ha a dúsulásra kapott p-érték szigorúan kisebb volt, mint 0,05 a Bonferroni korrekció használata után. Az eredmények mind stresszelt, mind stresszeletlen esetben a sejtciklusban és a DNS javításban szerepet játszó fehérjék fontosságát mutatják.

A. GO funkciók dúsulása a nyugalmi állapotú élesztő 100 legközpontibb elemében	
<i>P-érték</i>	<i>GO kulcsszó</i>
$3,98 \cdot 10^{-6}$	non-homologous end-joining
0,000139	cell cycle
0,00189	cellular response to stimulus
0,00484	M phase
0,00637	cell cycle phase
0,00886	response to DNA damage stimulus
0,0165	cell cycle process
0,0212	double-strand break repair
0,023	DNA metabolic process
0,0256	DNA repair
0,048	double-strand break repair via non-homologous end joining
B. GO funkciók dúsulása a hő sokknak kitett élesztő 100 legközpontibb elemében	
<i>P-érték</i>	<i>GO kulcsszó</i>
$2,6 \cdot 10^{-5}$	condensed chromosome
0,000126	double-strand break repair via nonhomologous end joining
0,000414	double-strand break repair
0,000638	non-recombinational repair
0,00304	condensed nuclear chromosome
0,0063	recombinational repair
0,00869	DNA repair
0,0112	cell cycle
0,0127	nuclear part
0,0138	response to DNA damage stimulus

0,0304	M phase
0,0322	cellular response to stimulus
C. GO funkciók dúsulása az oxidatív stressznek kitett élesztő 100 legközpontibb elemében	
<i>P-érték</i>	<i>GO kulcsszó</i>
$1,7 \cdot 10^{-5}$	cell cycle process
$2,94 \cdot 10^{-5}$	cell cycle
0,000225	M phase
0,000284	microtubule motor activity
0,000354	cell cycle phase
0,00104	organelle fission
0,00294	mitosis
0,00331	nuclear division
0,00386	chromosome segregation
0,00531	nucleus
0,00595	mitotic cell cycle
0,00851	M phase of mitotic cell cycle
0,0107	condensed chromosome
0,0125	microtubule-based process
0,0153	cellular response to stimulus
0,0201	motor activity
0,0256	mitotic sister chromatid segregation
0,0257	microtubule-based movement
0,0348	chromosome organization
0,0413	sister chromatid segregation
0,0443	condensed chromosome kinetochore
0,0479	chromosomal part
D. GO funkciók dúsulása a ozmotikus stressznek kitett élesztő 100 legközpontibb elemében	
<i>P-érték</i>	<i>GO kulcsszó</i>
0,000101	cell cycle process
0,000167	cell cycle
0,00027	microtubule motor activity
0,00028	M phase
0,00028	non-homologous end-joining
0,000453	cell cycle phase
0,00389	combined immunodeficiency

0,00769	nuclear part
0,0109	nucleus
0,011	condensed chromosome
0,0128	upslanted palpebral fissure
0,0133	chromosomal part
0,019	motor activity
0,0216	double-strand break repair
0,0252	chromosome segregation
0,0253	macromolecular complex
0,0255	mitotic sister chromatid segregation
0,0256	DNA metabolic process
0,0289	cafe-au-lait spot
0,0322	organelle fission
0,0369	protein complex
0,0395	response to DNA damage stimulus
0,0411	sister chromatid segregation
0,0433	condensed chromosome kinetochore

F. Élesztő fehérje-fehérje kölcsönhatás-hálózatában a stressz hatására legjobban növekedő terjesztési központiságú pontokban dúsuló GO funkciók különféle stresszelt állapotokban

A terjesztési központiság értékeket a Turbine program segítségével számítottam, a dúsulási elemzést pedig a g:Profiler program R beépülő moduljának használatával. Szignifikánsnak minősítettem egy GO kulcsszót, ha a dúsulásra kapott p-érték szigorúan kisebb volt, mint 0,05 a Bonferroni korrekció használata után. Az eredmények egyértelműen összhangban vannak az ismert stresszválaszokkal, amit a "response to stimulus" és a "response to stress" kulcsszavak szignifikáns dúsulása is megerősít.

A. A hőszokkban legjobban növekvő elemekben dúsuló GO funkciók	
<i>P-érték</i>	<i>GO kulcsszó</i>
5,18·10 ⁻⁸	cellular carbohydrate metabolic process
6,15·10 ⁻⁷	response to stimulus
1,25·10 ⁻⁶	cellular carbohydrate biosynthetic process
2,31·10 ⁻⁵	carbohydrate biosynthetic process
0,000323	organic substance catabolic process
0,000354	catalytic activity
0,00046	catabolic process
0,000616	cellular catabolic process
0,00102	energy derivation by oxidation of organic compounds
0,00105	oxidation-reduction process
0,00116	carbohydrate metabolic process
0,00276	glycoside biosynthetic process
0,00276	oligosaccharide biosynthetic process
0,00276	disaccharide biosynthetic process
0,00276	trehalose biosynthetic process
0,00547	alditol biosynthetic process
0,00547	glycerol biosynthetic process
0,00873	generation of precursor metabolites and energy
0,0139	glycogen metabolic process
0,0156	glycogen biosynthetic process
0,0188	glycogen breakdown (glycogenolysis)
0,0191	organic substance metabolic process

0,0199	energy reserve metabolic process
0,0217	alpha,alpha-trehalose-phosphate synthase complex (UDP-forming)
0,0248	trehalose metabolic process
0,0368	fungus-type vacuole lumen
0,0437	single-organism metabolic process
B. Az oxidatív stresszben legjobban növekvő elemekben dúsuló GO funkciók	
<i>P-érték</i>	<i>GO kulcsszó</i>
$4,24 \cdot 10^{-7}$	cellular carbohydrate metabolic process
$2,2 \cdot 10^{-5}$	cellular carbohydrate biosynthetic process
$8,53 \cdot 10^{-5}$	glucan metabolic process
$8,53 \cdot 10^{-5}$	cellular glucan metabolic process
0,000254	carbohydrate biosynthetic process
0,000585	glucan biosynthetic process
0,000596	cellular polysaccharide metabolic process
0,000649	TRAPP complex
0,000928	glycogen metabolic process
0,00102	carbohydrate metabolic process
0,00143	energy reserve metabolic process
0,00194	response to stimulus
0,00249	polysaccharide metabolic process
0,00852	cellular polysaccharide biosynthetic process
0,00993	polysaccharide biosynthetic process
0,0142	oxidation-reduction process
0,0168	acetyl-CoA + H ₂ O + oxaloacetate => citrate + CoA
0,0196	glycogen biosynthetic process
0,0196	cis-Golgi network
0,0289	alpha,alpha-trehalose-phosphate synthase complex (UDP-forming)
0,0321	trehalose metabolic process
0,0477	protein phosphatase type 1 regulator activity
C. Az ozmotikus stresszben legjobban növekvő elemekben dúsuló GO funkciók	
<i>P-érték</i>	<i>GO kulcsszó</i>
$4,42 \cdot 10^{-9}$	cellular carbohydrate biosynthetic process
$2,66 \cdot 10^{-6}$	carbohydrate biosynthetic process
$5,85 \cdot 10^{-6}$	cellular carbohydrate metabolic process
$1,98 \cdot 10^{-5}$	glycogen biosynthetic process

0,000582	glucan biosynthetic process
0,000586	trehalose metabolic process
$6 \cdot 10^{-4}$	protein phosphorylation
0,000619	protein serine/threonine kinase activity
0,000922	glycogen metabolic process
0,00104	protein kinase activity
0,00141	energy reserve metabolic process
0,00306	glycoside biosynthetic process
0,00306	oligosaccharide biosynthetic process
0,00306	disaccharide biosynthetic process
0,00306	trehalose biosynthetic process
0,00316	phosphorylation
0,00372	response to stress
0,00586	response to stimulus
0,00595	alditol biosynthetic process
0,00595	glycerol biosynthetic process
0,007	kinase activity
0,00716	carbohydrate metabolic process
0,00839	cellular polysaccharide biosynthetic process
0,00958	cell cycle phase
0,00977	polysaccharide biosynthetic process
0,0127	glycogen synthesis
0,0133	glycoside metabolic process
0,0156	single-organism carbohydrate metabolic process
0,0173	glucan metabolic process
0,0173	cellular glucan metabolic process
0,0235	alpha,alpha-trehalose-phosphate synthase complex (UDP-forming)
0,0252	cell cycle process
0,0323	cell cycle
0,0365	mitotic cell cycle

G. Élesztő fehérje-fehérje kölcsönhatás-hálózatában a stressz hatására legjobban csökkenő terjesztési központiságú pontokban dúsuló GO funkciók különféle stresszelt állapotokban

A terjesztési központiság értékeket a Turbine program segítségével számítottam, a dúsulási elemzést pedig a g:Profiler program R beépülő moduljának használatával. Szignifikánsnak minősítettem egy GO kulcsszót, ha a dúsulásra kapott p-érték szigorúan kisebb volt, mint 0,05 a Bonferroni korrekció használata után. A kulcsszavakat itt csak $p = 10^{-20}$ -ig soroltam fel, a teljes lista ugyanis igen hosszú. Az eredmények a transzlációs folyamatok egyértelmű, erős csökkenését mutatják, ami egy jól ismert stresszválasz.

A. A hősokkban legjobban csökkenő elemekben dúsuló GO funkciók	
<i>P-érték</i>	<i>GO kulcsszó</i>
$1,81 \cdot 10^{-59}$	ribosome biogenesis
$6,76 \cdot 10^{-57}$	nucleolus
$4,68 \cdot 10^{-56}$	ribonucleoprotein complex biogenesis
$7,16 \cdot 10^{-52}$	preribosome
$1,75 \cdot 10^{-49}$	cellular component biogenesis at cellular level
$3,56 \cdot 10^{-48}$	nuclear lumen
$1,91 \cdot 10^{-47}$	rRNA processing
$1,84 \cdot 10^{-46}$	rRNA metabolic process
$1,78 \cdot 10^{-39}$	ncRNA metabolic process
$7,58 \cdot 10^{-39}$	organelle lumen
$7,58 \cdot 10^{-39}$	intracellular organelle lumen
$1,14 \cdot 10^{-38}$	membrane-enclosed lumen
$2,48 \cdot 10^{-34}$	nuclear part
$6,61 \cdot 10^{-32}$	RNA processing
$1,67 \cdot 10^{-31}$	cellular component biogenesis
$1,47 \cdot 10^{-28}$	ribosomal large subunit biogenesis
$4,33 \cdot 10^{-27}$	maturation of 5,8S rRNA from tricistronic rRNA transcript
$6,41 \cdot 10^{-27}$	maturation of 5,8S rRNA
$2,17 \cdot 10^{-25}$	non-membranebounded organelle
$2,17 \cdot 10^{-25}$	intracellular nonmembrane-bounded organelle
$5,69 \cdot 10^{-25}$	ribonucleoprotein complex
$2,39 \cdot 10^{-24}$	preribosome, large subunit precursor

$5,36 \cdot 10^{-24}$	nucleus
$2,08 \cdot 10^{-22}$	ribosomal small subunit biogenesis
$3,22 \cdot 10^{-22}$	RNA metabolic process
$3,68 \cdot 10^{-22}$	90S preribosome
$6,62 \cdot 10^{-21}$	cleavage involved in rRNA processing
$9,54 \cdot 10^{-21}$	RNA phosphodiester bond hydrolysis
$4,42 \cdot 10^{-20}$	maturation of SSUrRNA from tricistronic rRNA transcript
B, Az oxidatív stresszben legjobban csökkenő elemekben dúsuló GO funkciók	
<i>P-érték</i>	<i>GO kulcsszó</i>
$6,62 \cdot 10^{-38}$	ribosome biogenesis
$5,98 \cdot 10^{-37}$	preribosome
$1,99 \cdot 10^{-35}$	ribonucleoprotein complex biogenesis
$4,58 \cdot 10^{-34}$	nucleolus
$2,38 \cdot 10^{-32}$	rRNA processing
$1,39 \cdot 10^{-31}$	rRNA metabolic process
$1,82 \cdot 10^{-30}$	cellular component biogenesis at cellular level
$9,66 \cdot 10^{-30}$	nuclear lumen
$1,34 \cdot 10^{-26}$	ncRNA processing
$3,12 \cdot 10^{-25}$	ncRNA metabolic process
$1,01 \cdot 10^{-23}$	90S preribosome
$1,2 \cdot 10^{-22}$	organelle lumen
$1,2 \cdot 10^{-22}$	intracellular organelle lumen
$1,64 \cdot 10^{-22}$	membrane-enclosed lumen
$4,57 \cdot 10^{-21}$	nuclear part
$2,28 \cdot 10^{-20}$	RNA processing
C, Az ozmotikus stresszben legjobban csökkenő elemekben dúsuló GO funkciók	
<i>P-érték</i>	<i>GO kulcsszó</i>
$1,29 \cdot 10^{-40}$	ribosome biogenesis
$2,56 \cdot 10^{-39}$	ribonucleoprotein complex biogenesis
$5,54 \cdot 10^{-37}$	preribosome
$2,1 \cdot 10^{-35}$	nucleolus
$4,88 \cdot 10^{-34}$	cellular component biogenesis at cellular level
$2,21 \cdot 10^{-32}$	rRNA processing
$1,29 \cdot 10^{-31}$	rRNA metabolic process
$1,67 \cdot 10^{-27}$	ncRNA metabolic process

$1,24 \cdot 10^{-26}$	ncRNA processing
$1,69 \cdot 10^{-25}$	nuclear lumen
$2,12 \cdot 10^{-25}$	ribosomal large subunit biogenesis
$1,41 \cdot 10^{-22}$	preribosome, large subunit precursor
$2,19 \cdot 10^{-21}$	RNA processing
$3,37 \cdot 10^{-20}$	ribonucleoprotein complex
$9,28 \cdot 10^{-20}$	cellular component biogenesis

A disszertációhoz kapcsolódó közlemények

Szalay KZ, Csermely P (2014) Attractor structures of signaling networks: consequences of different conformational barcode dynamics and their relations to network-based drug design. *Mol Inf* 33: 463–468. **IF: 2.013**

Szalay KZ, Csermely P (2013) Perturbation centrality and Turbine: a novel centrality measure obtained using a versatile network dynamics tool. *PLoS ONE* 8: e78059. **IF: 3.534**

Farkas IJ, Korcsmáros T, Kovács IA, Mihalik Á, Palotai R, Simkó GI, **Szalay KZ**, Szalay-Bekő M, Vellai T, Wang S, Csermely P (2011) Network-based tools for the identification of novel drug targets. *Science Signaling* 4: pt3.

Egyéb közlemények

Veres DV, Gyurkó MD, Thaler B, **Szalay KZ**, Fazekas D, Korcsmáros T, Csermely P (2014) ComPPI: a cellular compartment-specific database for protein-protein interaction network analysis. *Nucleic Acids Res*, doi: 10.1093/nar/gku1007 **IF: 8.808**

Csermely P, Hódsági J, Korcsmáros T, Módos D, Perez-Lopez AR, **Szalay K**, Veres DV, Lenti K, Wu LY, Zhang XS. (2014) Cancer stem cells display extremely large evolvability: alternating plastic and rigid networks as a potential mechanism. *Semin Cancer Biol* nyomtatás alatt, doi: 10.1016/j.semcancer.2013.12.004.

Hegedűs T, Gyimesi G, Gáspár ME, **Szalay KZ**, Gangal R, Csermely P (2013) Potential application of network descriptions for understanding conformational changes and protonation states of ABC transporters. *Curr Pharm Des* 19: 4155–4172. **IF: 3.288**

Köszönetnyilvánítás

Elsőként köszönetet szeretnék mondani **Prof. Csermely Péter** akadémikusnak, témavezetőmnek és immár barátomnak, hogy tanácsaival, tapasztalatával segített az elmúlt években, és hogy keményen dolgozott azon, hogy kutatót faragjon belőlem.

Köszönöm **Prof. Mandl József** akadémikusnak, és **Prof. Bánhegyi Gábornak**, hogy lehetőséget biztosítottak számomra az intézetben való munkára.

Köszönöm **Dr. Sőti Csabának** a labormunkában való megmerítkezés lehetőségét, és **Somogyvári Milánnak**, hogy megtanított a *C. elegans* vizsgálatának rejtjelmeire.

Köszönöm **Szóts Ákosnak** a Turbine CUDA modul elkészítését.

Köszönöm továbbá **Portschy Ádámnak** a Turbine összehangolását a BOINC rendszerrel és a helyi BOINC kiszolgálónk felállítását.

Köszönöm **Mihalik Ágostonnak** az élesztő stresszválaszában nyújtott tanácsadását.

Köszönöm **Dr. Amit Ghosh**-nak és **Prof. Saraswathi Vishveshwara**-nak a MetRS fehérjeszerkezeti adatokat.

Szeretném továbbá megköszönni munkacsoportomnak, a **LINK-Group** összes tagjának ötleteiket, tanácsaikat és a közös munkát.

Hálásan köszönöm édesanyámnak, **Szalay Évának** és édesapámnak, **Szalay Zsolt-nak**, hogy mindenben támogattak, segítettek, és bíztak bennem az idáig vezető úton.

Köszönettel tartozom feleségemnek, **Pintér Eszternek**, hogy mellettem állt, segített, bátorított, és támaszt nyújtott nekem mindig, amikor szükségem volt rá. Bízom benne, hogy az elkövetkező években viszonzni tudom.