

Hálózatok vizsgálata betegségekben

Vannak betegségeket, amelyek néhány vagy sok gén és a környezeti hatások révén alakulnak ki. Ezeket az ún. multifaktoriális vagy komplex betegségek, melyek általában nagyon gyakoriak, szemben a monogén betegségekkel. Ide sorolhatók olyan népbetegségek, mint az allergia, az asztma, a rák, a magas vérnyomás, a cukorbetegség, a szív-érrendszeri megbetegedések és az Alzheimer-kór. Továbbá gyakorlatilag az összes külső és belső tulajdonság is multifaktoriális jellegű, melyek vizsgálata nem különbözik a betegségektől, és mivel sokszor szoros a kapcsolat közöttük, nem választjuk el őket élesen egymástól.

A komplex betegségeknek néhány jellemzője:

- *Általában gyakoriak.* Például az asztma gyakorisága 6–10% (Magyarországon 7,5%; a tüdőgondozók által nyilvántartott asztmások száma 250 000 körüli). Szintén nagyon gyakori a magas vérnyomás (20–30%), a kóros elhízás vagy obezitás (20%), a 2-es típusú cukorbetegség (T2DM, 7%) vagy a depresszió (7–15%). A kardiovaszkuláris betegség következtében bekövetkező halál az összes halál 39%-áért felelős.
- *Családi halmozódás* figyelhető meg, de nem mutatható ki mendeli¹ öröklődés. Ez azt jelenti, hogy bizonyos betegségek egyes családokban sokkal gyakrabban fordulnak elő, mint az a populáció-gyakoriság alapján várható lenne, de a családfa alapján általában nem állapítható meg mendeli öröklődés (pl. domináns, recesszív, X-hez kötött stb.).
- *Gyakoribbak a poszt-reprodukciós korban.* Azaz, szemben a monogén betegségek többségével, a kór akkor jelentkezik, amikor már az illetőnek megszülettek a gyermekei, sőt részben fel is nevelte őket, azaz nincs kisselektálódás, a betegségre hajlamosító gének továbböröklődnek.
- *Összgazdasági jelentőségük óriási.* Például a kardiovaszkuláris betegségek az Európai Unió országaiban >170 milliárd €/év, az Egyesült Államokban 300 milliárd \$/év költséget, az asztma az Egyesült Államokban 18 milliárd \$/év összeget emészt fel. A 2-es típusú cukorbetegség kezelésére a teljes egészségügyi kiadások 12%-a megy, ami világszinten 376 milliárd \$-nak felel meg, és ez előreláthatólag 2030-ra 490 milliárd \$-ra nő.
- Egyes betegségek *gyakorisága* az elmúlt évtizedekben, főleg a fejlett országokban, jelentősen *emelkedett* (pl. elhízás, 2-es típusú cukorbetegség, magas vérnyomás, allergia, asztma stb.).
- Gyakran tapasztalható *több betegség együttes előfordulása* (komorbiditás).

Környezeti tényezők

A környezeti tényezők nagyon fontos szerepet játszanak a komplex betegségek kialakulásában, sőt, az esetek túlnyomó többségében *egy adott genetikai háttér csak bizonyos környezeti körülmények között hajlamosít valamilyen betegségre.* Környezeti tényezőnek nevezünk minden olyan faktort, ami nem genetikai. A teljesség igénye nélkül ide tartoznak például a méhen belüli hatások (az epigenetikai tényezőkön keresztül egész életre kihatnak), a táp-

lálkozás, a stressz, a dohányzás, a fertőzések, az életmód, a nevelés és az éghajlat.

Rendszerbiológiai adatok és megközelítések

Az elmúlt években a molekuláris biológiai mérés technikai módszerek, a számítógépek és a bioinformatika fejlődésével megnyílt a lehetőség arra, hogy az élőlényeket a valósághoz jobban közelítően, komplex rendszerekként modellezzük és értelmezzük. Az új mérés technikai módszerek elterjedésével soha nem tapasztalt teljességű és mennyiségű adathoz juthatunk. Ilyen módszer például a GWAS², mellyel egyszerre több millió genetikai variációt (SNP-t, vagy *single nucleotide polymorphism* ~ pontmutáció) tudunk vizsgálni. Vannak olyan ún. génexpressziós csipek, amelyekkel az összes gén kifejeződését (szintjét) mérni tudjuk egy szövetben vagy sejtcsoportban stb. Az ilyen módszereket, amelyekkel a teljes örökítő anyagot, a genomot egyszerre tudjuk vizsgálni, *genomikai módszereknek* hívjuk.

Természetesen a különböző szinteken lévő mérési eredmények, adatpontok sem függetlenek egymástól, hanem egymással kapcsolatban, kölcsönhatásban állnak. Például egy genetikai variáció, amely egy gén szabályozó régiójában helyezkedik el, nemcsak annak az adott génnek a kifejeződését változtathatja meg, hanem azokat is, amelyekkel az a gén, vagy terméke kölcsönhatásban áll. Továbbá, egy másik genetikai variáció befolyásolhatja ennek az genetikai variációnak a hatását pozitív és negatív irányban is. Egy élő szervezetben belül ezeket a kölcsönhatásokat több szinten is tapasztalhatjuk, és mára világossá vált, hogy egy szervezet működését, vagy például egy genetikai mutáció hatását csak akkor tudjuk értelmezni, ha átfogó módon, együttesen, rendszerszerűen vesszük figyelembe ezeket a kölcsönhatásokat. A biológián belül azt a tudományágat, amely ezeknek a hálózatszerűen ábrázolható kölcsönhatásoknak a feltérképezésével, és értelmezésével foglalkozik, *rendszerbiológiának* hívjuk.

Bár általános rendszerelméleti megközelítés már Bertalanffy 1960-as évekbeli munkájában megjelent, azóta is számos különböző matematikai háttérű rendszerbiológiai megközelítés jött létre. Ezek főként sejtbiológiai ihletésűek voltak, de ide sorolhatóak a teljes szerv, vagy szervrendszer modellezését célzó komplex hibrid rendszerek, mint az immunrendszer vagy az idegsejtek hálózatainak különböző modelljei. A kölcsönhatások együttesének átfogó modellezéséhez a rendszerbiológiai megközelítéseket az alábbi alapvető feltevések szerint csoportosíthatjuk:

- az entitások/faktorok páronkénti kölcsönhatásai,
- az entitások/faktorok többváltozós összefüggései-függelenségei,
- az entitások/faktorok oksági rendszerei.

¹ Gregor Mendel nevéből, aki borsókeresztezéses kísérleteivel megteremtette a tudományos öröklődést és megteremtette a modern genetika alapjait.

² Genome wide association study ~ Teljes genom asszociációs vizsgálat.

Elsőként az első ponthoz tartozó legkisebb adat- és számítás-igényű, páronkénti kölcsönhatásokra építő „hálózati” megközelítést tekintjük át, majd a harmadik ponthoz tartozó valószínűségi gráfos modelleket és végül az oksági rendszereken alapuló megközelítéseket.

Páronkénti kölcsönhatások modellezése és ábrázolása

A hálózati rendszerbiológiai megközelítésben a faktorok kölcsönhatásait páronkénti kölcsönhatásokként modellezzük, amelyeket gráfok felhasználásával ábrázolhatunk. Ekkor minden egyes faktornak egy-egy *csomópontot*, angolul „*node*”-ot feleltetünk meg, a kölcsönhatást a faktorok között a két csomópontot összekötő vonalakkal ábrázoljuk, amit a gráf *éleinek*, angolul „*edge*”-eknek nevezünk. A legtöbb csomópontnak csak kevés kapcsolata van, azonban vannak csomópontok, amelyeknek nagyon sok. Ez utóbbiakat „*hub*”-oknak, magyarul *hálózati középpontoknak* nevezzük.

Ha sejteken belüli molekuláris interakciókat modellezzünk így, akkor a csomópontok lehetnek pl. metabolitok (kismolekulájú anyagcseretermékek), illetve olyan makromolekulák, mint a fehérjék, RNS-ek, DNS-szekvenciák, míg az élek fizikai, biokémiai, vagy funkcionális interakciókat jelenthetnek.

Páronkénti kölcsönhatási hálózatok

Az elmúlt évtizedben rengeteg olyan eredmény született, amelyekkel ember-specifikus interakciós hálózatokat lehet felrajzolni. Ezek az egymásba fonódó hálózatok segítenek megérteni az emberi betegségek természetét, kialakulásának mechanizmusait. A molekuláris kölcsönhatások között megkülönböztethetünk *fehérje interakciós hálózatokat*, ahol az egyes csomópontokban fehérjék helyezkednek el, és az élek fizikai (kötő) kölcsönhatásokat jelentenek.

Az anyagcsere, vagy *metabolikus hálózatok*, amelyekben az egyes csomópontokban metabolitok vannak, és akkor kapcsolódnak egymáshoz, ha ugyanabban a biokémiai reakcióban vesznek részt. Valószínűleg a metabolikus hálózatról eddig szerzett ismeretanyag a legátfogóbb.

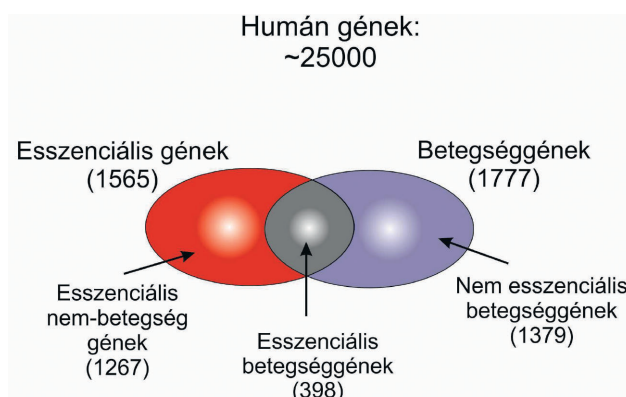
Az *RNS-hálózatok* RNS–RNS, vagy RNS–DNS interakciókat mutatnak, mint például a mikroRNS-ek génszabályozása. A mikroRNS-ek szerepét az utóbbi években ismertük meg részletesebben.

A hálózati megközelítés a betegségek gyógyszerével kapcsolatban is megjelent, amit a „hálózati gyógyszerkutatás” (*network pharmacy*) megközelítés is mutat, illetve a hatalmas méretű *hatóanyag-célpont* fehérje adatbázisok.

Betegség-gének a hálózatkutatásban

A különböző technikákkal megállapított interakciós hálózatokat a hálózatkutatás vizsgálja, amelyet orvos-biológiai hálózatoknál is sikeresen alkalmaztak. Ennek egy gyakori módja, hogy a különböző technikákkal megállapított interakciós hálózatonál, különböző módszerekkel azt vizsgálják, hogy a kapott hálózat mennyiben és miben tér el egy adott referenciának tekintett véletlenszerű hálózattípustól, és ezt hogyan lehet alkalmazni a biológiai folyamatokra.

Az előzőekben definiáltuk a hálózatokon belül a *hub*, vagyis a hálózati középpont fogalmát, amelyek olyan csomópontok, amelyekhez aránytalanul sok kapcsolat tartozik. Amikor a hub fehérjét modell-szervezetekben megvizsgálták, azt tapasztalták, hogy általában esszenciális gének kódolják őket és általában konzerváltabbak, mint a nem-hub fehérjék. A hub fehérjét kódoló gének kiütése általában nagyobb változáshoz vezet, és hiányuk szá-



1. ábra. Az emberi gének kevesebb, mint 10%-a, 1777 asszociál valamilyen betegséggel. 1565 génünk van, amelyik in utero, vagyis az embrionális fejlődésben esszenciális. A két halmaz csak részben fed át egymással

mos más fehérje funkcióját is befolyásolja. Ez ahhoz a hipotézishez vezetett, hogy a hub-oknak asszociálni kell a betegség-génekkel. Ezt támasztja alá az is, hogy amikor az OMIM adatbázisban³ szereplő betegség-géneket megvizsgálták, a kódolt fehérjék több fehérje-fehérje interakcióban vettek részt, mint a nem-betegség-gének által kódolt fehérjék.

Azoknak a géneknek az erős hatása, a funkciót jelentősen módosító mutációi azonban, melyek a korai embrionális fejlődésben létfontosságúak, általában nem tudnak továbböröklődni, így kiszekelődnek a populációból. Ezzel szemben, a humán betegség-oka-zó mutációk többségét a hordozók sokáig, sokszor felnőttkorig elnyúlóan tolerálni tudják. Ez azt jelenti, hogy emberben (illetve fejlett élőlényekben) a betegség-gének nem lehetnek mind esszenciális gének (1. ábra). A vizsgálatok alapján azt lehetett megállapítani, hogy emberben a hub proteinek kódoló gének esszenciális gének, míg a betegség-gének inkább a periférián helyezkednek el a hálózatokban (2. ábra).

Viszont, ha a gyenge hatású genetikai polimorfizmusokat is figyelembe vesszük, melyek csak kicsit módosítják a gén által kódolt fehérje működését, számos hub proteinnel találkozhatunk, amelyek kihatású genetikai variációi több komplex betegséggel is asszociációt⁴ mutattak. Ilyen például a gyulladási folyamatokban központi szerepet játszó tumor nekrozis faktor- α , mely génjének, a *TNF*-nek a variációi asztmával, obezitással, 1-es és 2-es típusú cukorbetegséggel, Alzheimer-kórral és érlelmeszesedéssel is asszociálnak. Hasonló hub fehérje még a β_2 adrenereceptor (*ADRB2*). Variációi befolyásolják az asztmát, obezitást és szerepet játszik a vérnyomás szabályozásában is.

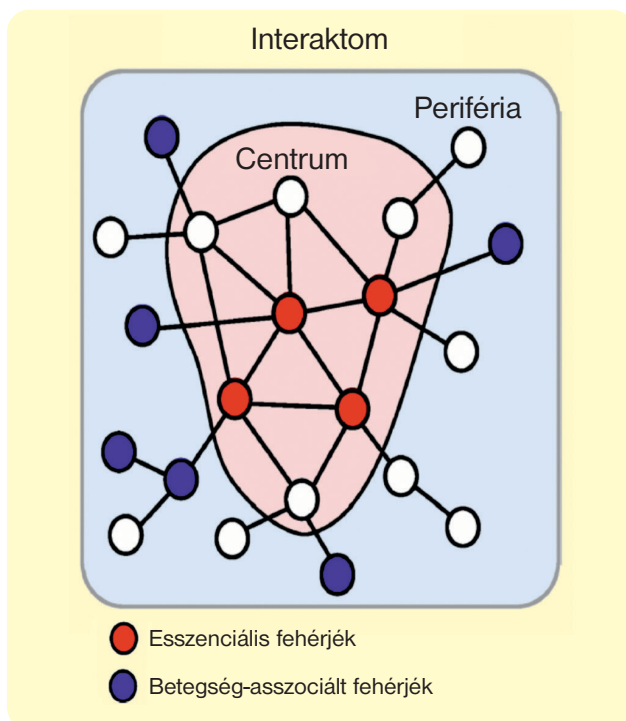
A *PPARG* gén tipikus hub fehérjét kódol, hiszen mutációja magas vérnyomást, elhízást, 2-es típusú cukorbetegséget okoz, a betegben érlelmeszesedés alakul ki.

A rendszerbiológiai módszerek alkalmazása

A betegségekkel kapcsolatban azonosított gének és anyagcsere-útvonalak mind potenciális gyógyszer és terápiás célpontok. Egy vizsgálatban, 1-es típusú cukorbetegségben (T1DM), a korábban nagy genomszűrések segítségével megismert gének által kódolt fehérjéknek megkeresték a velük interakcióban levő fehérje-partnereit, és 68 új T1DM gént sikerült azonosítani.

³ Online Mendelian Inheritance in Man; <http://www.omim.org/>

⁴ Az asszociáció itt egy statisztikai kijelentés, ami arra vonatkozik, hogy nagyobb a valószínűsége annak, hogy a genetikai variáció valamilyen kapcsolatban áll a betegséggel, mint annak, hogy nincs közöttük kapcsolat.



2. ábra. Az esszenciális és a betegség-gének közötti különbség sematikus ábrázolása. Az embrionális fejlődésben esszenciális gének által kódolt fehérjék (piros) a hálózat centrumában helyezkednek, ezek általában hub-ok. A betegség-gének inkább a periférián találhatók

Magas vérnyomásban az első nagy genomikai vizsgálatok nem vezettek eredményre. Viszont, amikor az eredményeket a rendszerbiológia eszközeivel elemezték, több olyan anyagcsere-útvonalat azonosítottak, amelyek szerepet játszanak a vérnyomás szabályozásában, de addig jelentőségüket nem ismerték fel.

GWAS és génexpressziós mérési eredmények felhasználásával, rendszerbiológiai megközelítéssel azonosítottak több új Parkinson-kórhoz kapcsolható anyagcsere-útvonalat.

Új eredményeket hozott az a rendszerbiológiai vizsgálat is, amelyben 5 multifaktoriális betegség GWAS-eredményei alapján létrehozott interakciós hálózatát tárták fel. Az 5 betegség közül két neurodegeneratív betegség (Alzheimer- és Parkinson-kór), és három autoimmun betegség (szklerózis multiplex, amely egyben neurodegeneratív is), rheumatoid arthritis és T1DM) volt. Először az öt betegségre a GWAS-eredmények alapján, adatbázisok segítségével egyenként anyagcsereútvonal-feldúsítást (*pathway enrichment*) végeztek. Azaz megnézték, hogy a GWAS-okban azonosított gének milyen anyagcsere-útvonalakhoz kapcsolódnak. Majd páronként a közös útvonalakkal rendelkező betegségeket összekötötték. Az élekhez 3 és 30 között számokat rendeltek, ahol a kisebb szám jelölte a több közös anyagcsere-útvonalat. A 3. ábrán látható hálózatból látszik, hogy mindegyik betegségnek vannak közös anyagcsere-útvonalai. Érdekes, és kissé váratlan módon, a legerősebb kapcsolat az Alzheimer-kór és a T1DM között volt. Az eredmények több, eddig nem ismert kapcsolatot, illetve betegséghez kapcsolható útvonalat tártak fel. Ezek közül csak egyet említünk. A vártak megfelelően a B-sejt és a T-sejt⁵ aktivációs útvonalak minden autoimmun betegségben szerepeltek, azonban meglepő módon Alzheimer-kórban is. Eddig az adaptív immunrendszer szerepe a betegségben nem volt ismert, bár volt né-

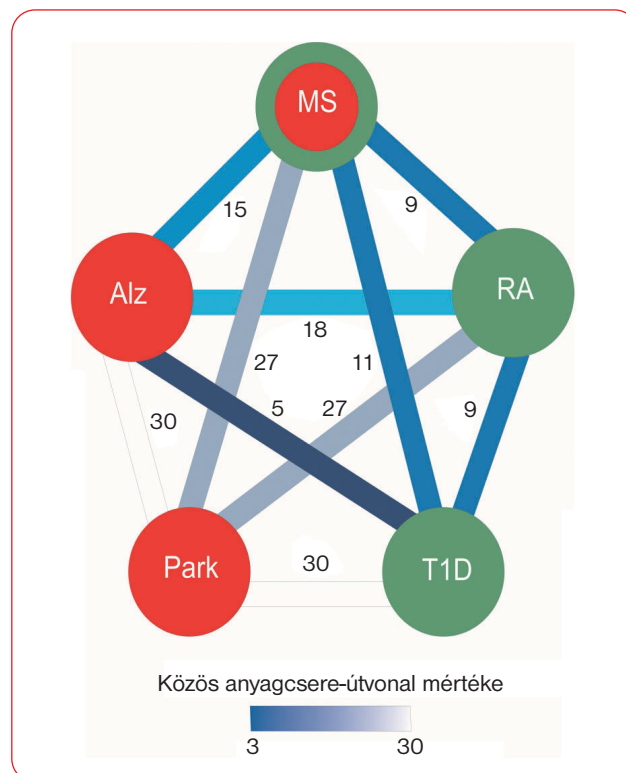
hány vizsgálat, amelyben megváltozott T-sejtválaszt tapasztaltak a betegekben. Illetve, ismert volt, hogy a gyulladáscsökkentők rendszeres használata csökkenti a betegség kialakulásának kockázatát. Az is érdekes, hogy a Parkinson-kór valamivel szorosabb kapcsolatban áll rheumatoid arthritissel és szklerózis multiplex-szel, mint Alzheimer-kórral.

A rendszerbiológia eszközeit a gyógyszerkutatásban is felhasználható. Olyan gyógyszerek, amelyek egy meghatározott célponttal rendelkeznek, sokszor javíthatják ugyan a betegség néhány hibáját, de más szomszédos, kapcsolt hálózatokat is megzavarhatnak, ami mellékhatásokhoz vezethet. A gyógyszerhatás hálózatszemléletű megközelítése alapján a legtöbb betegséget nem lehet mágikus lövedékkel (*magic bullet*) meggyógyítani, azaz olyan gyógyszerrel, amely egyetlen csomópontra hat. Erre példa a daganatok, vagy az AIDS terápiájában használt kombinált kezelések hatékonysága.

Szintén fontosak a gyógyszer-célpont-hálózatok, amelyek a forgalomban, vagy kísérleti fázisban levő gyógyszerek fehérjecélpontjait ábrázolják. Ennek elemzése alapján túlsúlyban vannak a palliatív gyógyszerek, amelyek nem közvetlenül a betegséget okozó fehérjére hatnak, hanem annak hálózati szomszédjára.

A rendszerbiológiai megközelítés alapján a racionális gyógyszerkutatásban fel kell tárnai az adott betegséghez tartozó betegség-hálózatot, és elég olyan hatóanyagokat keresni, amelyek ebben a betegségmodulban okoznak detektálható változásokat. Ez jelentősen leszűkítheti a keresés térét, és segíti a betegség diagnosztikához használható biomarkerek detektálását is, hiszen a betegségmodul komponenseinek aktivitásának változásai mutathatják a legerősebb korrelációt a betegség progressziójával.

3. ábra. 5 betegség (Alzheimer-kór (Alz), Parkinson-kór (Park), szklerózis multiplex (MS), rheumatoid arthritis (RA) és 1-es típusú cukorbetegség (T1D)) interakciós hálózata. A betegségeket összekötő élek színei, illetve a melléjük írt számok a közös anyagcsere-útvonalak mértékének rangsorát jelzik. A 3-as érték jelöli a legszorosabb, míg a 30-as a legalacsonyabb szintű rokonságot



⁵ Két fehérvérsejt-típus, melyek az ún. adaptív immunrendszer főszereplői.

Páronkénti összefüggésektől a többváltozós interakciókig

A gyakran esztétikai élményszámba menő színes interakciós gráfokat nézve, és az egyre gyarapodó élek mögötti tudásra gondolva, paradoxonnak tűnhet, hogy legalább ugyanilyen fontos az élek hiánya. Azonban az interakciós gráfokra, mint modellekre gondolva, a hálózatok ritkaságának a fontossága máris érthetővé válik, ha a gráf élsűrűségére mint a modellek egyszerűségének egy mértékére tekintünk. A faktorok többváltozós függetlenségének pontos reprezentálását és annak a tudásmodellezésben, induktív következtetésben való felhasználását a valószínűségi gráfok modellek szolgáltatják. Az irányítatlan éleket használó Markov-hálók és az irányított éleket használó Bayes-hálók közül az utóbbi vált népszerűbbé a bioinformatikában, mivel oksági értelmezést és kiterjesztési lehetőségeket is kínál.

Többváltozós interakcióktól az oksági hálózatokig

A Bayes-hálók (BN) a valószínűségi gráfok modellek egyik alosztálya, amelyben irányított, körmentes gráfokat (directed acyclic graph: DAG) használunk a sokváltozós eloszlás függetlenségeinek és kvantitatív jellemzőinek bemutatására, illetve opcionálisan az eloszlást generáló oksági mechanizmusok reprezentálására is. Az értelmezés szerint a csomópontok a véletlen változókat, az élek pedig közvetlen oki ráhatást jelentenek, így definiálva a modell struktúráját. A struktúrát definiáló gráf lokális valószínűségi modellekkel van kiegészítve, nevezetesen minden csomópontot tartalmaz egy lokális modell, amely megadja az azon csomópont által reprezentált valószínűségi változó valószínűségi függését a gráfban szülőként⁶ jelenlévő valószínűségi változóktól. Ezen lokális modellek paraméterei a modell paraméterei.

A Bayes-háló modellosztályt több tudományterületen is felhasználják. A teljesség igénye nélkül ide tartozik a döntéstámogatás, a statisztikai adatelemzés, vagy éppen a biomarker kutatások. A Bayes-hálók sokoldalúsága abból a tényből következik, hogy három autonóm kutatási szintet kapcsol egybe: az oksági modellt, a valószínűségi modell függetlenségi struktúráját és a kvantitatív eloszlást, amit a 4. ábra illusztrál. A gráfok modellek alkalmazása valószínűségi és oksági modellezésben visszavezethető az 1920-as évekig, Wright útvonal-diagramokat vizsgáló munkájáig, melyek rendszerbiológiai és genetikai alkalmazásai is nagyon széles körűek, de modern oksági kutatásokban is központi szerepet játszanak.

Hálózatok fúziója, az adat és a tudás fúziója

A nagy tömegű adatok megjelenése új korszakot nyitott a biológiában és az orvostudományban, azonban az adatok és ismeretek integrálása, az adat- és tudás-fúzió problémája továbbra is megoldatlan.

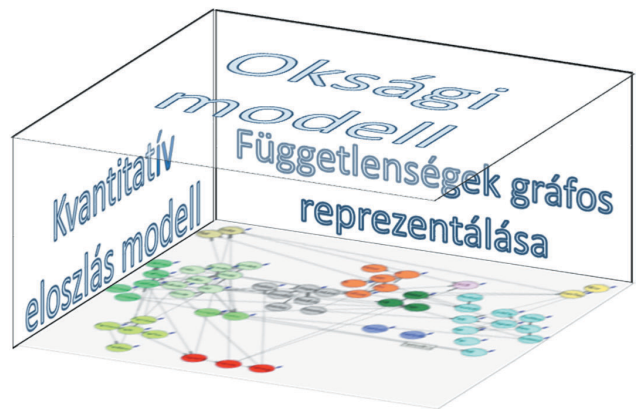
Évente kb. egymillió tudományos közlemény jelenik meg csak orvos-biológiai témakörökben, de ez egy szűkebb területen is ezres nagyságrendet jelent. Ilyen mennyiségű cikk követése meghaladja az emberi képességek határait, pedig a heterogén ismeretek integrálása, a tudományos haladás záloga, így egyre sürgetőbben merül fel az igény arra, hogy akár egyéni szempontok szerinti szá-

⁶ Itt a szülő a csomópont, ahonnan az irányított él indul.

⁷ Szilíciumban”, azaz: „a gépben” kifejezés arra a technikára utal, amikor a kísérlet nem az élő szervezetben, hanem a számítógépben, illetve számítógépes szimuláció során folyik.

⁸ Gyulladásos fehérvérsejt, melynek szintje sok asztmásban magas

⁹ Allergiás megbetegedésekben fontos immunglobulin, más néven antitest.

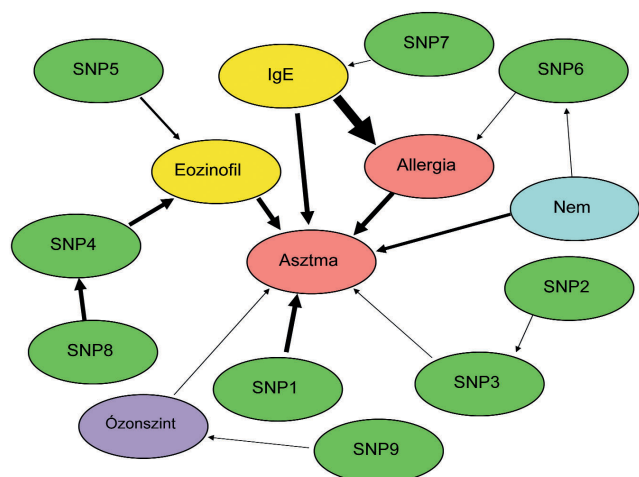


4. ábra. A Bayes-hálók nézetei: oksági, függetlenségi és parametrikus

mítási módszerek felhasználásával egyesítjük ezt a rendkívül sokrétű információt. Ez kulcsfontosságú mind az egyéni megismerés gazdagítása, mind pedig a gyógyszer-innováció és a személyre szabott gyógyászat céljainak elérése érdekében.

A komplex betegségek rendszerbiológiai vizsgálatában különösen fontosak a különböző megfigyelési szintek integrálásából származó vegyes hálózatok. Egy hálózatban szerepelhetnek például genetikai variációk (SNP-k), klinikai tünetek, környezeti tényezők, laborparaméterek stb. Ilyenkor a hálózat a különböző mérési eredményekből, bioinformatikai módszerekkel kiszámolt ok-okozati kölcsönhatásokat ábrázolhatja. Ebből megismerhetjük például, hogy melyik genetikai variáció befolyásolja közvetlenül, vagy más

5. ábra. Vegyes kölcsönhatási hálózat asztmában. A különböző jellegű csomópontokat (betegségeket, genetikai variációkat, laborparamétereket, környezeti faktorokat, egyéni jellemzőket) más-más színnel jelöltük. Az élek vastagsága arányos az összefüggés erősségével. Látszik például, hogy az SNP1 közvetlenül befolyásolja az asztmára való hajlamot, míg az SNP2 csak az SNP3-mal kölcsönhatásban. A magas eozinofil⁸-vagy IgE⁹- szint is összefüggést mutat asztmával, akárcsak az allergia. Az SNP5, vagy az SNP4 nem közvetlenül az asztmára hajlamosít, hanem befolyásolják az eozinofil-szintet és azon keresztül hatnak az asztmára. A levegő szennyezettsége (ózonszint a környezetben) szintén összefüggést mutat az asztmával. Az SNP9, amely például egy olyan útvonalon szereplő génben lehet, amely a levegőszennyezettség semlegesítéséért felel, az ózonszinten keresztül mutat összefüggést az asztmával



genetikai variációval, vagy egy környezeti tényezővel (vírussal vagy allergénnel) kölcsönhatásban a jelleg (betegség) kialakulását, vagy egy olyan laborparamétert (pl. koleszterinszintet), amely megnöveli a betegség kialakulásának esélyét (5. ábra).

Út a szingularitásba: adatgazdag szimuláció

Az adatok kutatásbeli szerepe több tudományterületen is megváltozott, sőt, ez a jelenség tudománytörténeti korszakhatárnak is tekinthető, amikor a XX. század második felére jellemző számítás-intenzív, szimulációs korszakot egy adat-intenzív, adatelemző korszak váltotta fel. A vizionált új kutatási paradigma központi eleme ezen nagy adattömegek léte és a hipotézismentes kutatás. Fontos felismerni azonban, hogy az adatgazdagság kiegészítője és nem kizárója a számítás-intenzív szimulációknak, sőt, az adatgazdagság és a szimuláció integrálása a hálózat kutatásnak is új szakaszát nyithatja meg, ugyanis az interakciós hálózatok bizonyos rendszerbiológiai modellezése lehetővé teszi a hálózatok dinamikai szimulációját is. Integrálásukkal egy öngerősítő folyamat indulhat be, amelyben a rendszerbiológiai szimulációk adatgazdagsága növekszik, illetve az egyre jobb szimulációk eredményei maguk is a felhasználható, nagy adattömegeket gyarapítják. Ennek előjele, hogy az ismeretek gyarapodása miatt az utóbbi években egyre jobb kompromisszumok mellett lehetséges a „teljes sejt” vagy a „teljes szervrendszer” dinamikai szimulációja. A betegségek megismerését az ilyen beavatkozásokat, kísérleteket is *in silico*⁷ lehetővé tevő modellek tehetik majd teljessé, amelyek gyógyszerkutatásban vagy akár egyéni terápiaválasztásban is felhasználhatóak lesznek.

Irodalom

- Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011 Jan;12(1): 56–68.
- Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. *Cell.* 2011 Mar 18;144(6): 986–98.
- Barabási AL, Albert R. Emergence of scaling in random networks. *Science.* 1999 Oct 15;286(5439):509–12. PubMed PMID: 10521342.
- Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL. The large-scale organization of metabolic networks. *Nature.* 2000 Oct 5;407(6804): 651–4.
- Albert R, Jeong H, Barabási AL. Error and attack tolerance of complex networks. *Nature.* 2000 Jul 27;406(6794): 378–82.
- Edwards YJ, et al. Identifying consensus disease pathways in Parkinson's disease using an integrative systems biology approach. *PLoS One.* 2011 Feb 22; 6(2): e16917.
- Gao S, Wang X. Predicting Type 1 Diabetes Candidate Genes using Human Protein-Protein Interaction Networks. *J Comput Sci Syst Biol.* 2009 Apr 1;2:133.
- Menon R, Farina C. Shared molecular and functional frameworks among five complex human disorders: a comparative study on interactomes linked to susceptibility genes. *PLoS One.* 2011 Apr 21;6(4): e18660.
- Binder CJ, et al. (2004) IL-5 links adaptive and natural immunity specific for epitopes of oxidized LDL and protects from atherosclerosis. *J Clin Invest* 114: 427–437.
- Taleb S, Tedgui A, Mallat Z (2010) Adaptive T cell immune responses and atherogenesis. *Curr Opin Pharmacol* 10: 197–202.
- Ludwig Von Bertalanffy: General System Theory: Foundations, Development, Applications, Penguin University Books, 1969
- S. Wright: Correlation and causation, *J. of Agricultural Research*, 20, 557–585, 1921
- N. Friedman: Inferring Cellular Networks Using Probabilistic Graphical Models, *Science*, 303, 5659, 799 – 805, 2004
- E. E. Schadt et al.: An integrative genomics approach to infer causal associations between gene expression and disease, *Nature Genetics*, 37(7): 710–717, 2005.
- R. Mourad et al: Probabilistic graphical models for genetic association studies, *Briefings in Bioinformatics*, 13, 1, 20–33, 2012.
- J. Pearl: Causality: Models, Reasoning, and Inference, Cambridge University Press, 2000.
- G. Bell et al.: Beyond the Data Deluge. *Science* 2009, 323(5919):1297–1298.
- M. Tomita: Whole-cell simulation: a grand challenge of the 21st century, *Trends in Biotechnology* Vol.19 No.6, 2001.
- J.R. Karr et al.: A whole-cell computational model predicts phenotype from genotype, *Cell.* 150(2):389–401, 2012.
- V. Narang et al: Systems immunology: a survey of modeling formalisms, applications and simulation tools, *Immunol Res.*, 53(1-3): 251–65, 2012.

Számunk szerzői