

BONYOLULTSÁG A GENOM-LÉPTÉKŰ BIOLÓGIÁBAN; ADALÉKOK A POSTGENOMIKUS AGNOSZTICIZMUSHOZ

Falus András

Semmelweis Egyetem, Genetikai, Sejt- és Immunbiológiai Intézet – faland@dgci.sote.hu

*Egyesekből áll az egész; s mégis azokból
erre és viszont ezekből azokra az ítélet
bizonytalan*

Kölcsey Ferenc

Az élettudományok korunkban minden eddigi korszakhoz képest alapvető változáson mennek át. Az eddigi biológiai „előtörténet” két szélsőség között mozgott (az angol nyelv *present perfect*-je a megfelelő igeidő), e két megoldás ma is jelen van. Az egyik az ún. *holisztikus* – a rálátó, a teljes rendszert vizsgáló megközelítés. Ez a magát szintetizálóknak értékelő tudományos módszer, amely a részletekkel nem foglalkozik, nem tud, de nem is nagyon akar kitérni a *black box* szemlélet elnagyolásai elől. A másik szélsőség a biológiai analitikai nézetrendszer, ahol az egyre kisebb elemeket (sejt, molekula, atom) tekintő tudományos közelítés, mert általában megoldani nem tudja, kitér a szintézis igénye elől. Valahol mind a két vonulat jelen van a tudományosan gondolkodó kísérletező emberben, és jelentős szerepe van jól-rosszul kezelt agnoszticizmusában. A holisztikus szemlélet elnagyolásai riasztó „határsértéseket”, az analitikus rövidlátás pedig az élővilág csodáihoz, a biológiai bonyolultsághoz méltatlan redukcionizmust eredményez.

Jelen gondolatok írója sokakkal egyetemben úgy érzi, hogy a genomika az ezred-

fordulóval együtt bekövetkező megjelenése az élettudományokban kiutat, valódi alternatívát jelent ebből a helyzetből.

*A gének „összhangzattana” a genomika,
a biológiai írásbeliség alapja*

A *genomika* genom léptékű biológiát jelent, azaz azt, hogy a vizsgálatok akár egyidejűleg is kiterjedhetnek az adott élőlény (például ember) összes génjének DNS szintű, illetve expressziós (mRNS és/vagy fehérje) analízisére. A kulcsszó a *mintázat*, tehát az egyidejűleg jelen lévő génvariánsok, illetve expresszálódnó gének összképének megjelenítése. A lényeg tehát az, hogy egy holisztikus biológiai analitika van kibontakozóban. Ma már lényegében ismert annak a hárommilliárdnál több nukleotidbázis fizikai sorrendje, amely minden sejtünkben két (testi sejtek) vagy egy (ivarsejtek) példányban fordul elő. Tudjuk azt is, ha megközelítően is, hogy ez az információ emberben mintegy 30–40 ezer gént jelent. A sejtek, embrionális programjuk és szöveti környezetük által meghatározva, bizonyos specializált funkciót látnak el, és ehhez a bennük lévő génkészlet egy részének „megszólalása” szükséges, ez a gének mRNS-re való átírását és fehérjék szintézisét jelenti.

A teljes genom szintjén történő expressziós vizsgálatokat (mRNS és fehérje szinten) funkcionális genomikának nevezzük.

A genomika elméleti jelentősége a gének funkcióinak megértésében, a modern mole-

kuláris sejtbiológiában és a biológia szinte minden ágazatában korszakos nagyságrendű. Gyakorlati szempontból a szerkezeti és funkcionális genomika jelentősége biomedicinális és mezőgazdasági területeken ugyancsak rendkívülinek látszik. Mai nézeteink, a gyakorlati felhasználás hatóköre és a kivitelezés logisztikája feltehetően gyorsan módosulni fognak, a genomikai szakirodalom nagy ütemben bővül.

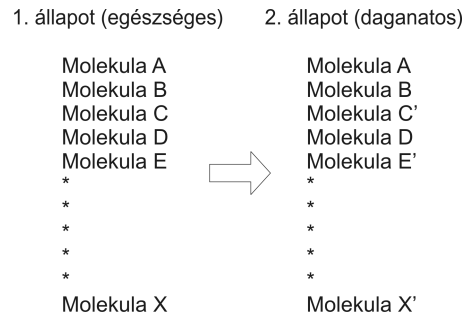
A klasszikus és a molekuláris genetika (öröklésbiológia) lényegében egyes génekekkel, azok felépítésével, variánsaival és funkcióival foglalkozik. A genomika, ezt felhasználva és ebből kiindulva, mintázatokat, génvariánsok együttes előfordulását, illetve időben változó kifejeződését képes vizsgálni. A biológiai módszerek egyik leghatékonyabb sikere a *génchip* (*microarray*, néhány cm²-en akár több 10-100 ezer ismert nukleinsavszál rendezett mátrixa) segítségével az egyidejűleg vizsgálható gének száma elérheti a vizsgált élőlény összes génjének számát. Sokan a *génchip*-ek (jó lenne magyar nevet találni rá, nevezhetnénk akár génmorzsának) bevezetését ahhoz a fejlődéshez hasonlítják, ami az elektronikában a félvezetők bevezetését követte.

Lehetővé válik annak a (*gén*)*mintázatnak*, *profilnak* a vizsgálata, ami az adott élőlény (pl. ember) szervezetében vagy egy szövetében egyszerre jelenik meg. Az időbeli egybeesés persze nem kell, hogy feltétlen oksági kapcsolatot jelentsen, de ez az egyidejűség időben változó kölcsönhatások sokaságát eredményezheti. Ha például azt feltételezzük, hogy egy sejtben egyszerre mintegy ezer gén „szólal meg” ez már sok millárd kölcsönhatást jelenthet. Ha ugyanis csak a kétkomponensű kapcsolatot számítjuk, ez már egymillió molekuláris párbeszédet jelent, márpedig a biológiai folyamatok szinte kivétel nélkül kettőnél több szereplősek, ahol az egyes események egymásutánisága, tehát időbeli rendezettsége is döntő fontosságú.

Nem véletlen, hogy viharos gyorsasággal közeledik egy vadonatúj szakma, a *bioinformatika*. Ez a matematikai, korrelációs függvényekben gondolkodó, biológiai kérdésekre válaszoló (de talán még több új kérdést feltevő) tudomány a jövő sztárágazata lesz. Legyen szabad két gondolat kísérletet bemutatni. Vegyünk például egy valódi biológiai problémát, a daganat keletkezésének és -fennmaradásának kérdését.

A genomika mai (és naponta bővülő) potenciálját tekintve kétféle logika mentén közelíthetünk:

1. Már ismert „jelölt” (*candidate*) gének (például génmutációval vagy expressziós sajátosságaival kiemelt gének) jelenlétét vizsgáljuk. Minthogy a tudományos információk halmaza óriási, óhatatlanul a tudásunkra, előítéleteinkre, érdeklődésünkre (és persze lehetőségeinkre) jellemző előválasztást végzünk abban, hogy mit vizsgáljunk. Vizsgálható objektumokként szóba jönnek a modern onkológia, immunológia, gyógyszeran által lényegesnek feltételezett molekulák (*1. ábra*). Ezek kellő alaposítással való vizsgálatával új, akár igazán lényegi felismerést is tehetünk.



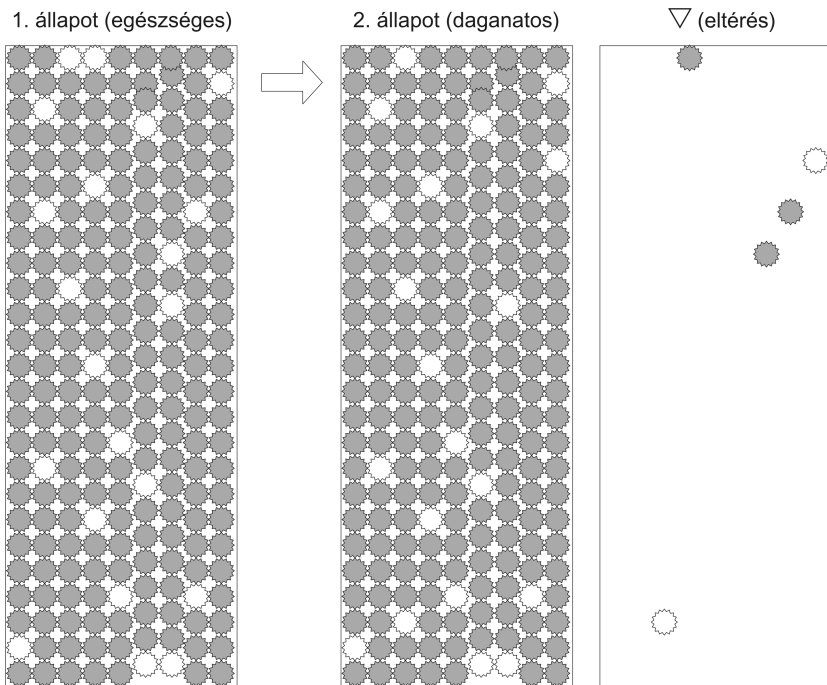
1. (és 2.) ábra • A biológiai funkcióért felelős gének kikeresésének két stratégiája – 1. Már ismert „jelölt” gének összehasonlítása két állapotban (például egészséges-beteg). Példánkban a C – C', E – E' és X – X' változások jelzik az állapotban történt módosulást.

2. A másik megközelítés a *reverz genomika*, itt a kétféle (példánkban daganatos betegség) állapot között – előszelekció nélkül – azt vizsgálom, hogy például 10-20 ezer gén egyidejű kifejeződése között mi a különbség. Kiemeljük tehát a két állapotban megfigyelhető eltéréseket, és csak azután keressük meg, hogy mi is az, ami eltér (2. ábra). Ez egyfajta „génhalászat”. Akár több száz, eddig nem ismert funkciójú gén kerülhet így napvilágra. A kezdetén vagyunk ennek a folyamatnak, és nagyon sok nehézség várja az „úttalan utat” törő génhalászt.

A kísérletet itt egy sokkal hosszabb és nehezebb bioinformatikai, korrelációs számítógépes (*in silico*) munka követi, melynek jó esetben új diagnosztikai és terápiás következményei lehetnek. Az *in silico* lehetőség sok más újszerűsége mellett egyben a tudomány nagyfokú demokratizálódását is jelenti.

Arról van szó, hogy rengeteg olyan ember is végezhet korszerű tudományos munkát a világháló adatbankjaiban dolgozva, akik eddig anyagi, képzettségi vagy más okokból nem férhettek hozzá a kísérletes munkához. Másrészt időben is kiterjed a kutatás, például akár évtizedekkel korábbi szövettani blokkokban a mai sokszorosan kiterjedtebb lehetőségekkel (például az összes emberi gént egyidejűleg használva) is végezhető kutatás.

A génexpressziós profilok esetén általában elképesztő bonyolultságot találunk: a kiértékelés, a bioinformatika időbeni (sorrend) és okozati (mi hat mire) kérdéseket oldhat meg. Géncsoportok rajzolódnak ki a *cluster-analízis* révén, géncsoportok, amelyek hasonló szabályozás alatt állnak, azonos betegségállapotokat határolnak el, vagy azonos állapotokat/szövödményeket prediktálnak.



2. ábra • (2) Nagyszámú gén mintázatainak összevetése két állapotban (például egészséges-beteg). A két mintázat „kivonásával” kapjuk meg az eltérésben részt vevő lehetséges géneket.

Pontmutációk mátrixmintázata és felhasználása

Még furcsább (hogy ne *elidegenítőt* mondjak) az a megoldás, amely a genomban átlagosan ezer bázispáronként előforduló SNP-eket (*single nucleotide polymorphism*) használja mintázatok képzésére. Arról van szó, hogy a nukleotidbázisok között *pontmutációk* (ez az SNP-k leánykori neve) fordulnak elő, a négy nukleotidbázisnak megfelelően négy variációval. Ezek genetikai, mendeli módon öröklődő jellegzetességek (allélek), melyek egymástól elvileg függetlenek. Annak az esélye, hogy valakinél egy ponton a DNS-ban hasonló változatok fordulnak elő, bizonyos valószínűségi értékkel írható le. Az hogy két ponton van hasonló eltérés, már kisebb valószínűségi értéket jelent. Annak esélye, hogy tíz (független) SNP egyforma már 1/1 000 000. Ha húsz SNP-t vizsgálunk, akkor ez egymilliárd ember közül emel ki egyet. Ehhez képest minden emberi genomban átlagosan hárommillió pontmutáció van, tehát elképesztően egyediek vagyunk. Ez akkor is így van, ha jól tudjuk, az emberi genom több mint 90 %-a nem kódol fehérjét, hanem látszólag (vagy valóban?) „értelmetlen” információkkal (például rövidebb-hosszabb repetitív szakaszok) töltik ki a teret. Dawkins „önző” génjei lennének ezek? És miért hurcoljuk őket sejtről sejtre, apáról fiúra? Mi a molekuláris „értelmetlenség” biológiai jelentése? Biztos, hogy a mai tudásunk által értelmetlen genetikai „szemétnek” kezelt, térkitöltő „vatta” nukleinsav sorrendek valóban nem tartalmaznak valamilyen információt? Tudunk-e olvasni vagy még csak a „képeket” nézzük a DNS könyvében? Mindenesetre a pontmutációk a genom ezen „kietlen” (vagy csak jelenleg annak látszó?) részére esnek.

Értjük vagy sem az SNP-k biológiai jelentését, a *pragmatikus biológia* máris használja ezeket. Ha például két, valamely tulajdonságában eltérő embercsoportot SNP minták-

zatukkal jellemezhetünk, és számítógép segítségével olyan SNP mátrixokat találhatunk ezek között, amelyek konzekvens eltérést mutatnak a két csoport között. Egy, az előbbi tulajdonságra nézve ismeretlen embernél az SNP minta prediktálhatja, hogy ő melyik csoportba fog esni. Ez így elég teoretikus, de ha például a két csoport abban tér el, hogy az egyik esetében mellékhatása van egy gyógyszernek és a másiknál nem, akkor a módszer alkalmas egy gyógyszer esetlegesen súlyos mellékhatásának előjelzésére. Ez ma már a *farmakogenomika* nevű, rohamosan fejlődő tudomány egyik fő eljárása, s óriási haszonnal jár (nemcsak a gyártónak, de a betegnek is), hiszen kisebb mellékhatások mellett gyorsabban találnak rá a megfelelő gyógyító eljárásra. Emlékeztetek arra, hogy a nagy haszon mellett az SNP-k többsége esetén fogalmunk sincs, melyik gén vagy ismétlődő „értelmetlen” géndarabban van az a pontmutáció, amelyik esetleg a több száz SNP közül valahogy jelzi a gyógyszerhatást. Kitűnő fiatal barátom Dr. V. B. erre azt mondta, hogy ez „molekuláris tenyérjólás”, és – lássuk be – igaz van.

Teljes szervezet szintű modellek, gének kiütve és beütve, „nem várt” meglepetések

A modern embriómanipulációs technika lehetővé teszi a testi sejtek kiindulási elődjei, az őssejtek genetikai módosítását, így gének ki- és bevitelét. A technika: a homológ rekombináció ma már kondicionáltan, sejtekre célzottan és időben tetszőlegesen indíthatóan végezhető el. Így egyes génekre hiányos, másokra nézve túlműködő egész szervezeteket (például egér) hozhatunk létre. Amikor az ilyen állatok tulajdonságait (fenotípusát) vizsgáljuk, a biológiai bonyolultság újabb szintén ámulhatunk el. Bizonyos gének alul- vagy túlműködése hatalmas meglepetéseket okozhat. Kidertült, hogy a gyulladássos immunológiai válaszban ismert anyagok

hiánya észrevétlenül kompenzálódik, olyan „pótkötél” mechanizmusok lépnek a kiütött gén funkciója helyébe, amelyekről eddig nem is tudtunk. Ugyanakkor megismertünk „pótolhatatlan” géneket is, melyek hiánya letális következményekkel jár.

Gének túlműköd(tet)ése esetén is sok meglepő eredményt kaptak, kiderülhet például, hogy eddig nem jó helyen keresték az adott gén termékének hatását, a funkció máshova lokalizált. Kitűnő megközelítés ez viseltes biológiai dogmáink levetésére!

Metabolom, a negyedik királyság

A „klasszikus” molekuláris biológia nagyrészt DNS-el, mRNS-sel illetve fehérjékkel foglalkozik, ezek a genomika (DNS), transzkriptomika (mRNS) és proteomika (fehérjék) *királyságai*, kisebb részt átfedő, de főként önálló metodikákkal és eltérő kérdéssel feltevési logikával. A DNS világ kérdései az egyedre specifikusak, mi öröklődik, mi a genetikai variánsok, a populációs polimorfizmus jelentősége. A mRNS kutató az expresszióval foglalkozik, fő kérdései sejt- és szövetspecifikusak, tehát például az, hogy mi szólal meg egy adott sejtben. A proteomika fehérjeszerkezettel és variánsokkal foglalkozik, és fehérjeszerkezetek prediktív vizualizációja révén új hatóanyag molekulákat, oltóanyagokat stb. tervez.

A negyedik szint a *metabolom*, a kis molekulású anyagok világa, ez az a *milieu*, ahol a DNS átíródik, s a mRNS fehérjenyelvre fordítódik. Kiderült, hogy baktériumokban ide tartoznak többek között az aminosavak, peptidek, mono- és oligonukleotidok, poliszaharidok és az aminok. Utóbbiakra példa a hisztamin, ez a 112 D méretű, a hisztidin aminosavból dekarboxilált monoamin. A hisztamin többek között szerepel az allergiás, gyulladásos és gastroenterológiai folyamatokban, átvívóanyag bizonyos neuronok között.

Kanadai, japán és magyar munkatársainmal együtt évekkkel ezelőtt sikerült egy olyan „knock-out” egeret létrehozunk, amelyben

nem termelődik a hisztamin. Ez az egér számos olyan előre nem látott, korábban nem feltételezett tulajdonságot produkál, amely arra utal, hogy e kis molekulású molekula szerepe sokkal bővebb, mint azt eddig hittük Kiderült, hogy a hisztaminnak szerepe van olyan folyamatokban is mint az étvágy-szabályozás, csontképzés és a sejtosztódás.

A jéghegy csúcsa alatt levő terület tehát láthatóvá tehető, és a *biológiai bonyolultság újabb titka* kerül elénk.

Összefoglalás (helyett)

A biológiai bonyolultság fogalma és tényanyaga további markáns árnyalatokkal bővült a genomika korszakának beköszöntével. A megközelítés lehetővé teszi az egyes genetikai és génepressziós tényezők helyett akár több tízezer gén és géntermék mintázatának együttes vizsgálatát. Ez hatalmas minőségi ugrást jelent, és az eredmények kiértékelésében nélkülözhetővé tesz egy új tudományt, a bioinformatikát. Ma már tudjuk, hogy a sejtek, szövetek fejlődése, osztódása és egyéb tulajdonságai mögött gének összetett, egymást reguláló hálózatai állnak, amelyek speciális szinkronizációjáról ma még keveset tudunk. Bár már használjuk, keveset tudunk a pontmutációk kombinatív mintázatainak biológiai jelentéséről. E sorok írója nagyon fontosnak tartja, hogy a biomedicinális tudomány ezen markáns előrelépése ne vezessen sem vak tudományimádkodáshoz, sem pedig annak méltatlan alulértékeléséhez. Tudjuk meg, mire képes a genomika ma és talán holnap, és azt is, mire nem. Ennek a józanul értékelő mértéktartásnak remélhetően pozitív hatása lesz a társadalomra a tudomány értékeinek megbecsülését illetően.

Fürküssz a Lét műhelyében

Mindig egészet a részben.

(Goethe)

Kulcsszavak: *genomika, DNS-chip, komplexitás*