

Komplex jellegek genetikai hátterének elemzése*

Rónai Zsolt dr.¹ ■ Lippai Zoltán² ■ Elek Zsuzsanna¹ ■ Somogyi Anikó dr.²

Semmelweis Egyetem, Általános Orvostudományi Kar,

¹Orvosi Vegytani, Molekuláris Biológiai és Patobiokémiai Intézet, ²II. Belgyógyászati Klinika, Budapest

Bár a Humán Genom Projekt másfél évtizede feltárta a 3 milliárd nukleotidból álló emberi genetikai információ bázissorrendjét, a betegségek – elsősorban a komplex rendellenességek – hátterének pontos megismerése még várat magára. A még azonosítatlan örökletes tényezők összességét hiányzó örökölhetőségnek nevezzük, ennek felderítése a molekuláris patomechanizmus megismerésének alapja. Ez nem csupán elméleti kérdés: ezen tudás a mindennapi gyakorlatban a diagnosztika, a megelőzés és a célzott, egyénre szabott kezelés fejlődésének lehetőségét kínálja. A még nem ismert genetikai faktorok azonosításához a mind újabb és hatékonyabb molekuláris biológiai technikák alkalmazása hozzájárul, a cél eléréséhez azonban számos klinikai és genetikai koncepció újragondolása vezethet el. Tudásunkat az eddigi genomszintű analízisek megalapozták, de további ismeretek feltárása szükséges az alábbi szempontok alapján: (1) SNP-k mellett az ismétlődési variációk (VNTR-ek és CNV-k) genotipizálása és asszociáció elemzése, (2) gén–gén és gén–környezet kölcsönhatás vizsgálata, (3) epigenetikai elemzések, (4) polimorfizmusok biológiai funkciójának meghatározása, (5) biológiailag releváns diagnosztikai kategóriák, endofenotípusok alkalmazása. Noha a genomnak csupán az 1,2%-a felelős a fehérjék kódolásáért, ugyanakkor csaknem 90%-a RNS-re átíródik, így a gén-expresszió-szintű vizsgálatok ígéretes kiindulópontot jelenthetnek, mivel rávilágíthatnak azon molekuláris szintű folyamatokra, amelyek szerepet játszanak a betegségek kialakításában.

Orv Hetil. 2018; 159(31): 1254–1261.

Kulcsszavak: mutáció, genetikai polimorfizmus, komplex jelleg, génexpresszió, heritabilitás

Investigation of the genetic background of complex diseases

Although the Human Genome Project discovered the sequence of the human genetic information 15 years ago, genetic background of the diseases – primarily that of complex disorders – is still not known. The sum of the not yet discovered inherited risk factors is termed the missing heritability; the identification of these genetic components is, however, essential, as it is the base of the understanding of the molecular pathomechanism of diseases. It is not only of theoretical importance: this knowledge can be used in the clinical practice, as it offers the possibility of improvement of diagnostics, prevention as well as targeted and individualized therapy. Application of novel and more efficient molecular biological tools contribute to the discovery of unknown genetic factors, the complete goal can only be achieved, however, by re-conceptualization of several clinical and genetic points. Our knowledge was established by genome-wide studies, however, further knowledge must be acquired according to the following points: (1) genotype and association analysis of repeat variations (VNTRs and CNVs) besides SNPs, (2) investigation of gene–gene and gene–environment interactions, (3) epigenetic studies, (4) assessing the biological function of polymorphisms, (5) application of biologically relevant diagnostic categories and endophenotypes. Although it is only 1.2% of the whole genome that codes for proteins, however, as much as 90% is transcribed to RNA, consequently it can be hypothesized that gene expression analyses might offer promising starting points for further studies, as they can shed light on the molecular processes that contribute to the development of diseases.

Keywords: mutation, genetic polymorphism, complex disorders, gene expression, heritability

Rónai Zs, Lippai Z, Elek Zs, Somogyi A. [Investigation of the genetic background of complex diseases]. Orv Hetil. 2018; 159(31): 1254–1261.

(Beérkezett: 2018. március 11.; elfogadva: 2018. április 16.)

*A Dr. Fehér János Alapítvány 2018. évi pályázatán díjazott dolgozat.

Rövidítések

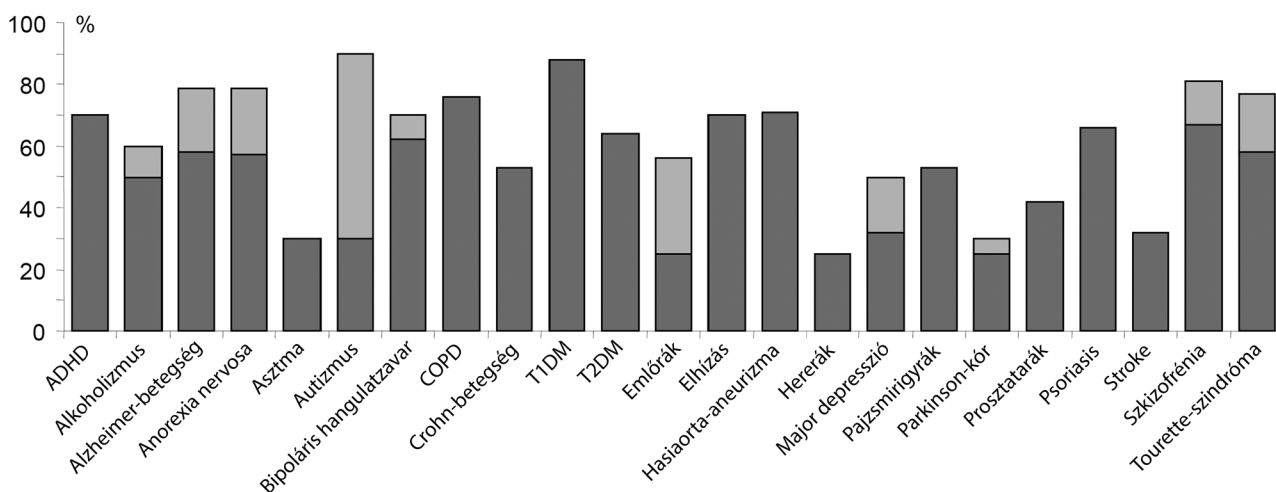
bp = bázispár; CNV = (copy number variation) kópiaszám-variáció; DNS = dezoxiribonukleinsav; DRD4 = 4-es típusú dopaminreceptor; GWAS = (genome-wide association study) genomszintű asszociációelemzés; IL = interleukin; MAF = (minor allele frequency) ritka allélfrekvencia; miRNS = mikro-RNS; NCBI = (National Center for Biotechnology Information) Nemzeti Biotechnológiai Információs Központ; PCR = (polymerase chain reaction) polimeráz-lánreakció; RNS = ribonukleinsav; siRNS = (small interfering RNS) kis interferáló RNS; SNP = (single nucleotide polymorphism) egyponos nukleotid-polimorfizmus; STR = (short tandem repeat) rövid ismétlődési variáció; VNTR = (variable number of tandem repeats) ismétlődési polimorfizmus; WES = (whole exome sequencing) teljesexom-szekvenálás

Betegségek heritabilitása, a hiányzó örökölhetőség

Az orvostudományi kutatások egyik fő irányvonala napjainkban a betegségek háttérben zajló kóros folyamatok molekuláris szintű feltárása. Ennek ismerete nemcsak elméleti, de gyakorlati klinikai szempontból is alapvető jelentőségű. Számos, rendkívüli szociális, népegészségügyi jelentőségű betegség – pszichiátriai rendellenességek, tumorok, szív-, érrendszeri megbetegedések, 1-es és 2-es típusú cukorbetegség stb. – a komplex kórképek családjába tartozik, melyek jellegzetessége, hogy létrejöttükben mind környezeti tényezők, mind pedig örökletes faktorok alapvető szerepet játszanak. A kettő arányára iránymutató az örökölhetőség (heritabilitás) értéke, amely fenotípusjegyek varianciájának genetikai komponensekkel meghatározható hányadát – azaz lényegében a háttérben álló örökletes tényezők szerepének mértékét – fejezi ki [1]. Megkülönböztethető az örökölhetőség „szűkebb” (h^2), illetve „tágabb” (H^2) értelem-

ben vett meghatározási módja („narrow-sense” és „broad-sense heritability”), melyek között a különbség az, hogy az előbbi csak a vizsgált polimorfizmusok közötti additív hatásokkal számol, míg az utóbbi figyelembe veszi a genetikai variációk közötti interakciókat (epistasz), valamint az epigenetikai hatásokat is [2]. Néhány betegség örökölhetőségértékét az 1. ábra foglalja össze.

A humán genomot alkotó körülbelül 3 milliárd nukleotidnyi szekvencia másfél évtizede ismert és mindenki számára szabadon hozzáférhető. Ez az eredmény a genetikai kutatások alapvető mérföldköve, ugyanakkor kiderült, hogy a rendelkezésre álló információ megértése, elméleti és gyakorlati alkalmazása korántsem egyszerű. Számos tanulmány célul tűzte ki és célozza meg napjainkban is a komplex kórképek háttérben álló örökletes faktorok feltárását, mégis mind a mai napig egyetlen olyan betegség (vagy jelleg) sincs, amelynek esetében a kialakulásért felelős genetikai tényezők teljes listája ismert lenne. Az intenzív kutatások ellenére a feltárt örökletes komponensek a különböző kórképek esetében a heritabilitásnak [3–6] (1. ábra) csupán a 6–50%-át magyarázzák [7, 8]: a fennmaradó, még azonosítatlan örökletes tényezők összességét hiányzó örökölhetőségnek („missing heritability”) nevezzük [9]. A komplex kórképek háttérben álló genetikai, molekuláris komponensek – azaz a még hiányzó örökletes faktorok – feltárása ugyanakkor mind elméleti, mind klinikai szempontból alapvető jelentőségű. Ez az ismeret a betegségek molekuláris patomechanizmusának pontos megismerése révén ugyanis a diagnosztika, a prognózis, a megelőzés és a hatékony, célzott, egyénre szabott kezelés alapjait jelenti. Egyre nyilvánvalóbb, hogy a hiányzó örökletes faktorok feltárása alapvetően nem „technikai” feladat, azaz nem elsősorban a dezoxiribonukleinsav (DNS)-szekvencia variációit vizsgáló módszerek további fejlesztése a kulcs, hanem koncepcionális kérdés.



1. ábra

Néhány betegség heritabilitásértéke

Egyes kórképek esetén a két különböző színű oszloprész az eltérő tanulmányokból származó legkisebb (sötét) és legnagyobb (világos) értéket mutatja.

ADHD = figyelemhiányos hiperaktivitás; COPD = krónikus obstruktív tüdőbetegség; T1DM, T2DM = 1-es, illetve 2-es típusú diabetes mellitus
 Forrás: <https://snpedia.com/index.php/Heritability>, <http://tga.nig.ac.jp/h2db/> [3–6]

A megoldást számos klinikai és genetikai szempont újragondolása és a különböző területek (klinikum, molekuláris biológia, bioinformatika) hatékony együttműködése jelentheti.

Genomszintű (GWAS) és kandidánsgén-tanulmányok

Nagy várakozás előzte meg a genomszintű asszociációanalíziseket (GWAS – genome-wide association analysis), melyek hipotézismentes vizsgálat során az összes kromoszómaszakaszra kiterjedő elemzést jelentenek. Ezekből a vizsgálatokból származik mai tudásunk számottevő része, ugyanakkor a kapott eredmények mégsem váltották be maradéktalanul a hozzájuk fűzött reményeket. A 2-es típusú cukorbetegség esetén a GWAS-ek több, mint 80 gyakori variáció szerepét vetették fel, melyek egyenként 5–40%-kal emelik a kórkép kialakulásának kockázatát [10]. Crohn-betegségben 32 – kis hatású – lókuszt az örökölhetőség körülbelül 20%-át [11], a korán kialakuló myocardialis infarctus esetén 9 genetikai variáció a heritabilitásnak csupán a 2,8%-át magyarázza [12], ugyanakkor maculadegeneráció esetén 5 nagyobb hatású genetikai komponenssel lefedhető az örökletes faktorok csaknem 50%-a [7]. Ez a néhány adat rávilágít a komplex jellegek genetikai hátterének egyik fontos aspektusára. A *gyakori betegség-gyakori variáns* elmélet szerint a komplex kórképek kialakulásában számos – egésszégesekben is megtalálható –, nagy gyakorisággal előforduló allélvariáns játszik szerepet a környezeti tényezőkkel karöltve [13], ugyanakkor látható, hogy egy-egy örökletes faktor hatása igen különböző mértékű lehet. Az alacsony (<1%) ritka allélfrekvenciájú (MAF-) variációk (mutációk) ugyanakkor szintén nem csupán a monogénis jellegek kialakításáért felelősek: a multifaktoriális etiológiájú rosszindulatú colorectalis megbetegedések hátterében több ritka variánst azonosítottak [14].

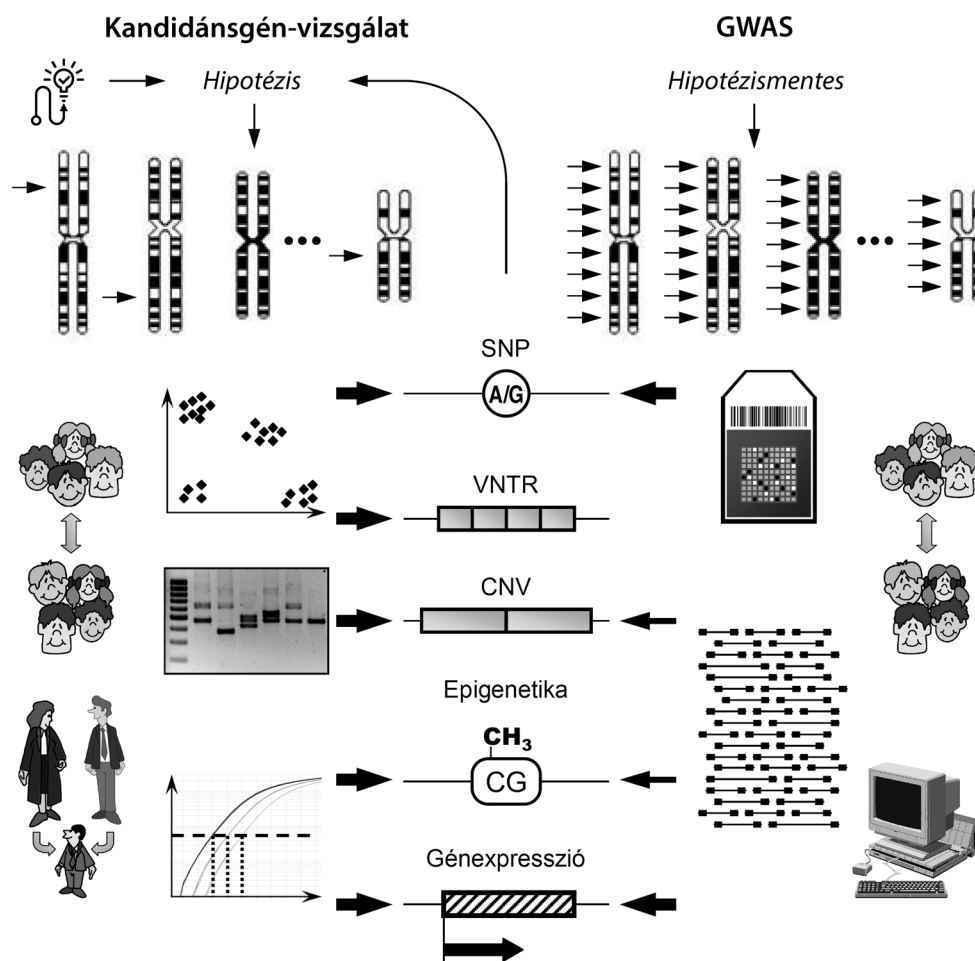
Egyebek mellett ezen a ponton is hozzájárul ismereteink bővítéséhez a komplex jellegek genetikai hátterét elemző másik fő irányvonal, a kandidánsgén-elemzés (2. ábra) [15]. Bár napjainkban a genomszintű, illetve meta- és megalízisek során rendkívül nagy létszámú minták elemzésére nyílik lehetőség, ennek ellenére a többszörös tesztes miatt szükséges statisztikai korrekció álnegatív eredményhez vezethet: a kis hatású és/vagy ritka variációk kimutatása meghiúsul.

A kandidánsgén-vizsgálat során élettani, biológiai, orvosi szempontok, illetve korábbi genomszintű vagy célzott kutatások alapján kiválasztott gének és polimorfizmusok vizsgálata történik meg. A célgének kiindulási hipotézisen alapuló szelekciója természetesen egyfajta korlátozó tényező a GWAS hipotézismentes megközelítéséhez képest. A kandidánsgén-tanulmányok a komplex képek csupán egy szűk szeletét vizsgálják, amelyből így – értelemszerűen – kimaradnak az epistasziból adódó hatások is. Ezen kutatási módot mégis több olyan szem-

pont is jellemzi, melyek miatt még a genomszintű elemzések mellett is hozzájárulnak ismereteink bővítéséhez és pontosításához.

Genetikai polimorfizmusok csoportosítása szerkezetük szerint

Két nem rokon személy genomja körülbelül 0,5%-ban (15 millió bázispár) tér el egymástól. Ezek a különbségek, azaz a genetikai variációk – szerkezeti szempontból – több csoportba sorolhatók (2. ábra). A legtöbb vizsgálat az egyponyos nukleotidpolimorfizmusokra (SNP – single nucleotide polymorphism) koncentrált, amelyek rendszerint a genom egyetlen nukleotidjának kicserélődését jelentik, de idesoroljuk az egyetlen nukleotidot érintő inszerciákat és deléciákat is. Az NCBI dbSNP-adatbázisában (<https://www.ncbi.nlm.nih.gov/snp/>) jelenleg több mint 672 millió SNP található. Az egyponyos nukleotidpolimorfizmusok mellett ugyanakkor számos ismétlődési variáció is megtalálható a genomban. Noha ezek száma az SNP-kénél alacsonyabb, nagyobb kiterjedésük és a variációs lehetőségek magasabb száma miatt mégis jelentős mértékben hozzájárulnak az emberek közötti genetikai különbözőséghez. Bár az ismétlődési polimorfizmusok vizsgálatának alapelve nem bonyolult, és – elsősorban kapilláris elektroforézis alkalmazásával – multiplex mérés is megvalósítható, ugyanakkor az SNP-k elemzéséhez hasonlítható genomszintű vizsgálat nem oldható meg. Ennek megfelelően az ismétlődési variációkat főként a kandidánsgén-tanulmányok kutatják. Ezek a polimorfizmusok több szempontból is rendkívül változatos csoportot jelentenek: mind az ismétlődési szám, mind az ismétlődő modul hossza igen széles határok között változhat, így a nevezéktan is meglehetősen összetett. A viszonylag rövid szakaszokból álló ismétlődések a VNTR-ek (variable number of tandem repeats), ezek egyik alcsoportját képezik a legrövidebb szakaszok (1–6 bp) ismétlődéséből álló mikroszatelliták (más néven STR-ek, azaz „short tandem repeat”-ek), valamint a valamivel hosszabb (10^1 – 10^2 bp nagyságrend) miniszatelliták [16]. A 4-es típusú dopaminreceptor (DRD4) 3. exonjában található egy 48 bp-os VNTR, melynek külön érdekessége, hogy ez a génszakasz a 7 transzmembránmodul alkotta receptorfehérje 3. citoplazmatikus hurokrégiójában okoz 16 aminosavnyi változást; ez a fehérjerész játszik szerepet a G-fehérjével létrejövő kapcsolatban. Az ezzel a polimorfizmussal kapott eredmények a pszichológiai, illetve pszichiátriai genetika terén úttörő jelentőségűek voltak [17, 18], a „hosszú” (7 ismétlődést tartalmazó) allél mára már a figyelemhiányos hiperaktivitás rizikófaktoraként ismert [19]. A DRD4-gén további érdekessége, hogy a szabályozó régiójában is található ismétlődési variációk, melyek feltételezhetően szerepet játszanak a génkifejeződés modulálásában [20]. Logikusan adódik a feltételezés, hogy egy gén kódoló és szabályozó régiójában elhelyezkedő polimorfizmusok egymással is kölcsönhatásban áll-



2. ábra Genomszintű (GWAS) és kandidánsgén-elemzés

Az ábra a GWAS és a célzott elemzés néhány aspektusát foglalja össze. A kandidánsgén-tanulmányok során orvosi, élettani, biokémiai alapon, illetve *in silico* adatok figyelembevételével történik meg a vizsgálatba bevont polimorfizmusok kiválasztása, melyek általában feltételezett biológiai funkcióval rendelkeznek. A genomszintű elemzések hipotézismentesek, rendszerint SNP-k elemzését célozzák, melyek – nagy felbontással – az egész genomot reprezentálják. A kandidánsgén-vizsgálatok során – célzott módszerekkel – SNP-k mellett ismétlődési variációk, a génkifejeződés módosulása és az ezzel szorosan összefüggésben álló epigenetikai mintázat is vizsgálható. A genomszintű SNP- és génexpresszió-elemzések microarray módszerrel megvalósíthatók, de ismertek nagy hatékonyságú eljárások a CNV-k, illetve a metilációs profil vizsgálatára is. A szekvenálási technikák elterjedése nagy előrelépést jelent. A kandidánsgén-tanulmányok során az eset-kontroll vizsgálat mellett a családi triók elemzése hasznos alternatíva (egyebek mellett a populációrétegződésből adódó álpozitív eredmény kizárása miatt); a GWAS-ok rendszerint eset-kontroll elemzések, kiértékelésük különleges informatikai háttérrel és ismerettel igényel

nak: az előbbieket a képződő fehérje működésére, az utóbbiak annak mennyiségére lehetnek hatással. Munkacsoportunk is felhívta a figyelmet arra, hogy esetenként az elméletileg egyszerű genotipizáló módszerek is rejtegethetnek technikai buktatókat magukban [21], ami több más tényezővel együtt hozzájárulhat ahhoz, hogy a genetikai asszociációelemzések eredményei esetenként nehezen reprodukálhatók [22].

Az ismétlődési variációk különleges csoportját képezik a CNV-k (copy number variation), melyek rendkívül hosszú (10^4 – 10^5 bp nagyságrendű) DNS-szakaszok ismétlődését jelentik. Megfigyelték, hogy ezen variációk nagy gyakorisággal megtalálhatók egészséges személyekben is [23, 24], ugyanakkor feltételezhető, hogy hozzájárulnak a komplex jellegek örökletes hátterének meghatározásához. Noha ismertek és elérhetőek módszerek a

CNV-k genomszintű elemzésére [25], mégis számos GWAS ezen polimorfizmuscsaládot figyelmen kívül hagyva csupán az SNP-k elemzésére fókuszál.

A genetikai polimorfizmusok biológiai szerepe

Az SNP-k (és rövid ismétlődések) molekuláris biológiai funkciója – a lokalizációval összefüggésben – igen eltérő lehet, és a polimorfizmusok ez alapján is különböző csoportokba sorolhatók. Kezdetben a legnagyobb figyelem a fehérjét kódoló régiókban elhelyezkedő báziscserékre irányult, mivel ezek esetében direkt összefüggés áll (illetve állhat) fenn a jelen lévő allélvariáció és a képződő fehérje primer szerkezete között. Mivel a bázisripletek

száma meghaladja a kódolt aminosavak számát (a genetikai kód „degenerált”), a kódoló régiók báziscseréinek egy része a fehérje aminosavsorrendjét nem változtatja meg (szinonim SNP-k). Ez azonban távolról sem jelenti azt, hogy ezek a polimorfizmusok biológiailag teljesen hatástalanok lennének: több esetben kimutatták, hogy ezek az SNP-k módosíthatják a ribonukleinsav (RNS) térszerkezetét és így módon a molekula stabilitását, élettartamát, vagy befolyásolhatják az intronok kivágódásának mechanizmusát is [26]. Hasonló szabályozó hatásúak lehetnek a nem kódoló szakaszokon, illetve nem fehérjét kódoló génekben elhelyezkedő variációk transzkripciófaktorok vagy a miRNS kötődésének befolyásolása vagy akár a metilációs mintázat megváltoztatása révén.

A missense polimorfizmusok a polipeptidlánc egy aminosavának megváltozását okozzák, aminek funkcionális következménye rendkívül széles skálán változhat. Egyes esetekben ráadásul a hatás közvetett, ismert olyan eset, amikor a missense polimorfizmus az RNS élettartamát is megváltoztatja, máskor a módosult primer szerkezetű fehérje el sem jut a rendeltetési helyére, hanem lebomlik [27].

A korai stopkodont létrehozó nonsense variációk, illetve a splice zavarát okozó báziscserék esetén a kódolt fehérje jelentős része hiányzik, ami rendszerint kifejezettebb funkcióvesztéssel jár. A genom szekvenciájának elemzési lehetőségei az elmúlt években – az informatika fejlődésével karöltve – jelentősen bővültek, a DNS bázissorrendjének vizsgálatára ma már az úgynevezett új generációs szekvenálási eljárások is elérhetők. Ezek lényege röviden, hogy rendkívül hatékony eszközökkel a DNS rövid és egymással jelentősen átfedő szakaszainak szekvenciáját határozzák meg. Új generációs szekvenáláson alapul a napjainkban már a diagnosztikában is alkalmazott teljesexom-szekvenálás (WES – whole exome sequencing) [28]. Ezzel a módszerrel a teljes genomnak csupán néhány százaléka: a fehérjék kódolásáért felelős kromoszómaszakaszok szekvenciája határozható meg. Ez a legtöbbször elegendő a jelentős funkcióvesztést okozó mutációk azonosítására, a komplex jellegek hátterében álló polimorfizmusok ugyanakkor gyakran a WES által lefedett régiókon kívül helyezkednek el.

Érdekes, hogy míg a teljes genom csupán alig több, mint 1%-a kódol fehérjét [16], addig RNS-re 83–90%-a átíródik [29, 30]. Ez az adat a „szemét-DNS” hipotézist megkérdőjelezi, és rávilágít ennek megfelelően arra is, hogy az ezekben a szakaszokban elhelyezkedő polimorfizmusok szintén hozzájárulhatnak egy-egy komplex jelleg vagy betegség genetikai hátterének meghatározásához.

Látható tehát, hogy a polimorfizmusok rendkívül szerteágazó biológiai funkciókkal rendelkezhetnek, ugyanakkor a GWAS-elemzések során a hatásmechanizmus feltárását nem tűzik ki célul. Az azonosított polimorfizmusok sok esetben nem rendelkeznek (az adott jelleg szempontjából releváns) biológiai funkcióval, csupán

azokkal kapcsoltan öröklődő genetikai markerek. Noha ezek a diagnosztika, illetve a rizikó becslése során eredményesen használhatók lehetnek, a molekuláris patomechanizmus megértését nem viszik előre. Említésre méltó az is, hogy még a monogénes kórképek diagnosztikája során elvégzett WES analízisek esetén is gyakran komoly kihívást jelent, jelentős bioinformatikai és elméleti szak tudást és munkát igényel a tényleges kóroki szerepű mutáció azonosítása. Nem meglepő ennek megfelelően, hogy a komplex kórképek esetén az adott rendellenességgel összefüggésben álló, általában kis hatású polimorfizmusok felkutatása nehéz feladatot jelent.

Etiológiai faktorok hálózatos kölcsönhatása, epistasis

A komplex jellegek hátterében álló örökletes faktorok hatása nem egyszerűen additív, a genetikai (és környezeti) faktorok egymással – hálózatot alkotva – bonyolult kölcsönhatásban, interakcióban állnak (epistasis) [31]. Ez a szempont ismét felveti azt a kérdést, hogy mennyiben szükséges, illetve megengedhető a hipotézis alkalmazása ezen kutatások során. Nyilvánvaló ugyanis, hogy az – SNP-k között csupán páronként vizsgált és a környezeti faktorokat figyelmen kívül hagyó – interakció megjósolása elméleti alapon csaknem lehetetlen, ugyanakkor egy GWAS-tanulmány során az összes lehetséges kombináció elemzése a többszörös tesztelést nagyságrendekkel növeli. Emiatt még nagyobb létszámú vizsgálati csoport bevonása válik szükségessé, és tovább nő az álnegatív eredmény valószínűsége, azaz a kis hatású variánsok azonosítása sikertelenné válik [32]. A genetikai és a molekuláris háttér pontos feltárása, a funkcionális alapok megismerése fokozottan nehéz, de egyben különösen jelentős azokban az esetekben, amikor valamilyen sajátos klinikai kép (például komorbiditás) hátterét elemezzük. A legfrissebb kutatások kimutatták például, hogy a megemelkedett vérlemezkeszám (thrombocytosis) hozzájárul a rosszindulatú megbetegedések során az áttétképződés kialakulásához, ami a betegség prognózisának jelentős romlását okozza [33]. Társuló cukorbetegség esetén a megváltozott szignáltranszdukciós (inzulinjelvény) folyamatok tovább bonyolítják a betegség molekuláris szintű patomechanizmusát. Állatkísérletes eredmények azt mutatták, hogy a thrombocytaszám és a plazma IL6-, illetve thrombopoetinszintje között szoros összefüggés mutatható ki [34]. Az IL6 azonban csupán egy azon citokinek közül, melyek összefüggésben állnak a megemelkedett thrombocytaszámmal, illetve általában a megacaryocytopoesissal. A thrombocytosis, illetve az áttétképzés közötti ok-okozati összefüggés pontos feltárása ugyanakkor mind elméleti, mind gyakorlati szempontból nagy jelentőségű, mivel így módon elérhetővé válhatna a thrombocytosis csökkentése révén a dagasztos megbetegedésben szenvedő betegek életkilátásainak javítása. A colorectalis carcinoma és a 2-es típusú cukorbetegség között fennálló többszintű kapcsolatáról számos

közleményben olvashatunk [35]. A colorectalis carcinoma kialakulásának kockázata 2-es típusú cukorbetegségben nagyobb [36], amihez a közös környezeti kockázati tényezők, mint például az 50 év feletti életkor, az elhízás, a helytelen táplálkozás és a mozgásszegény életmód is hozzájárulhat [37]. Mindkét betegség kialakulásában szerepe lehet a sejtkárosodásnak, a gyulladásnak, ami molekuláris szinten átfedő (például Wnt/ β -katenin) jelátviteli útvonalak révén manifesztálódik. Társuló cukorbetegség esetén a colorectalis carcinoma prognózisa rosszabb mind a kemoterápia, mind a műtéti beavatkozás vonatkozásában [38].

Molekuláris alapokon nyugvó diagnosztika és kezelés

A komplex jellegek és betegségek hátterének felkutatásával kapcsolatban egy további szempont is felvethető. A napjainkban alkalmazott klinikai diagnózisok definiálása sok esetben nem a kiváltó okok, hanem elsősorban a tünetek szerint történik. A molekuláris biológiai alapok feltárása, illetve az ezen alapuló diagnosztikai kategóriák meghatározása nem elméleti, nevezéktani kérdés csupán: a betegségek molekuláris patomechanizmusának ismerete és ennek megfelelő besorolása a célzott, oki, illetve egyénre szabott kezelési stratégiák kidolgozásának lehetőségét kínálja. A génműködés terápiai céljait történő befolyásolása napjainkban már nem mondható ritkaságnak: a leggyakrabban miRNS-ekkel vagy siRNS-ekkel történő géncsendesítés számos rosszindulatú megbetegedés esetében eredményesen alkalmazható [39]. A molekuláris háttér ismerete pedig terápiai célpontokat kínálhat, és farmakogenetikai, illetve -genomikai szempontból jelentős azon polimorfizmusok azonosítása is, amelyek egy-egy gyógyszerre adott egyéni válasz meghatározásáért felelnek [40]. Ismert például a dihidropirimidin-dehidrogenáz enzim defektusának jelentősége 5-fluorouracil-kezelés során [41], és jelentős hazai, valamint nemzetközi munkák mutatják be az asztma kezelése során alkalmazott kezelés (β_2 -agonisták, kortikoszteroidok) egyénfüggő hatékonyságának genetikai meghatározottságát [42]. Mindezek tükrében a betegségek relevánsabb besorolására a többlépcsős betegségmodell, illetve az endofenotípus alkalmazása jelenthet megoldást. Az endofenotípus fogalmát először a pszichiátriai genetika területén vezették be [43]. Olyan – rendszerint objektíven és számszerűen – mérhető paramétert jelent, mely genetikailag meghatározott, és egyértelmű összefüggést mutat az adott kórképpel, ugyanakkor azzal nem azonos: a beteg személyek egészséges családtagjainak körében nagyobb gyakorisággal figyelhető meg, mint átlagosan az egészséges populációban [44]. Ennek megfelelően egyes endofenotípusok jelenléte esetenként betegségmegelőző állapotnak is tekinthető. Az endofenotípus koncepcióját később a medicina más területeire is kiterjesztették [45], mivel a biológiai alapok feltárása során alkalmazá-

suk eredményesnek tűnik, és ily módon – a célzott, hatékony kezelés és a megelőzés területén – nagy jelentőségűek.

Az epigenetika és a génkifejeződés vizsgálatának szempontjai

A komplex jellegek hátterének kutatása során a genetikai rizikófaktorok mellett a környezeti hatások sem hagyhatók figyelmen kívül. Ez megnyilvánul egyfelől a gén-környezet interakciókban [46], és fontos szem előtt tartani, hogy az adott jelleg szempontjából releváns környezeti hatásoknak is kell, hogy legyen molekuláris szintű következménye. A teljes kép feltáráshoz így a DNS szekvenciájának vizsgálata (genetika) mellett az epigenetikai elemzések: a hisztonmódosulások, valamint a CG-dinukleotidok metilációs mintázatának [47] mérése is szükséges. Bár nem minden részletre kiterjedően ismert, de mégis logikus az is, hogy a két szint egymással is szoros összefüggésben áll: a genetikai variációk több ponton befolyásolhatják a kromatin szerkezetét, a metilációs mintázat kialakulását, illetve a génműködést befolyásoló további tényezőket: például a transzkripció faktorok vagy a mikro-RNS-ek kötődésének hatékonyságát [47]. Mindezek alapján vetődik fel az a megközelítési lehetőség, melynek során a komplex jellegek hátterében álló molekuláris tényezőket a génkifejeződés oldaláról közelítjük meg. Ez a szint ugyanis a genetikai és környezeti faktorok első eredőjének, közös megnyilvánulásának tekinthető. Itt bizonyos szempontból már közömbös, hogy a megfigyelt változások hátterében milyen faktorok állnak, de – a betegséggel vagy jelleggel összefüggésbe hozható – megváltozott működés kimutatható. Említésre méltó az is, hogy az mRNS-szint számos esetben nem korrelál a jelen lévő fehérje mennyiségével [48], és emiatt a transzkripció szintű elemzéseket alkalmanként kritika éri. Ez azonban részben amiatt van így, mert a hálózatos összefüggések RNS-szinten is megnyilvánulnak. Ugyanakkor – mint említésre került már – a genomnak csupán az 1%-a kódol fehérjét, viszont közel 90%-a átíródik RNS-re [29, 48]; ebből a szemszögből nézve tehát a transzkriptom elemzése mégis átfogóbb képet nyújt. Nem hagyható figyelmen kívül ugyanakkor az a szempont, hogy az epigenetikai és génkifejeződés terén megfigyelhető folyamatok szövetspecifikusak: a vizsgálatok – lehetőség szerint – biológiailag releváns forrásból származó minta esetén szolgáltatnak megbízható eredményt. Klinikai szempontból viszont természetesen azok a molekuláris szintű változások alkalmazhatók a mindennapokban a diagnosztika, a prognózis és a megelőzés terén úgynevezett biomarkerként hatékonyan, amelyek könnyen elérhető forrásból (például vér, szájnyálkahártyasejt) is megbízhatóan kimutathatók. Nagy jelentőségűek ennek megfelelően az olyan összehasonlító elemzések is, melyek a különböző szövetekben megfigyelhető folyamatok közötti eltérések kimutatását célozzák [49].

Következtetések

A genetikai, molekuláris biológiai kutatások és a mindennapi klinikai gyakorlat számos ponton összefonódást mutat. A betegségek hátterének génszintű megértése feltárhatja a rendellenességek molekuláris okait. Ez a biológiailag releváns diagnosztikai kategóriák kialakításának, illetve konkrétan a mindennapi klinikai munka során a diagnózis felállításának lehetőségét kínálja. A betegségek kialakulásában szerepet játszó örökletes faktorok, hajlamosító tényezők ismerete mellett új terápiás célpontok meghatározására nyújt módot, és döntő jelentőségű a primer és a szekunder prevenció szempontjából is. Mindezek alapján kétségtelen, hogy az öröklődő kórképek hátterében álló genetikai polimorfizmusok és mutációk feltárása alapvető feladat. Fontos ugyanakkor a különböző szintű genetikai vizsgálatok (célzott kandidánsgén-elmzések és GWAS-ok; alapkutatások, klinikai, diagnosztikai célú vizsgálatok) összehangolása, egyesítése, harmonizálása. Ebben fontos szerepet kapnak a több kutatómunka eredményeit közösen elemző metaanalízisek [50], és nagy jelentőségű az EuroGentest projekt hálózata is, amelynek egyik fő célkitűzése annak megvalósítása, hogy a genetikai vizsgálatok Európa-szerte hasonló alapelvek szerint történjenek a mintavételtől a genetikai tanácsadásig.

Anyagi támogatás: A munkát a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal (NKFI) K116128-as projektje, a Magyar Diabetes Társaság kutatási pályázata (Somogyi Anikó) és a STIA-KF-17. számú (Rónai Zsolt) források támogatták.

Szerzői munkamegosztás: A közleményben (illusztrációként) bemutatott saját kísérleteket és az ábrákat E. Zs. készítette, E. Zs. részt vett az átdolgozott változat fejezeteinek elkészítésében. A dolgozat klinikai vonatkozású szakaszait S. A. és L. Z., molekuláris biológiai fejezeteit L. Z. és R. Zs. írta. A közlemény szerkesztése, javítása, végső formába öntése S. A. és R. Zs. munkája. A cikk végleges változatát valamennyi szerző elolvasta és jóváhagyta.

Érdekltségek: A szerzőknek nincsenek érdekltségeik.

Irodalom

- [1] Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet.* 2013; 14: 139–149.
- [2] Mayhew AJ, Meyre D. Assessing the heritability of complex traits in humans: methodological challenges and opportunities. *Curr Genomics* 2017; 18: 332–340.
- [3] Hyttinen V, Kaprio J, Kinnunen L, et al. Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. *Diabetes* 2003; 52: 1052–1055.
- [4] Kaprio J. Twins and the mystery of missing heritability: the contribution of gene–environment interactions. *J Intern Med.* 2012; 272: 440–448.
- [5] Lundström S, Chang Z, Råstam M, et al. Autism spectrum disorders and autistic like traits: similar etiology in the extreme end and the normal variation. *Arch Gen Psychiatry* 2012; 69: 46–52.
- [6] Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet.* 2012; 13: 537–551.
- [7] Maller J, George S, Purcell S, et al. Common variation in three genes, including a noncoding variant in *CFH*, strongly influences risk of age-related macular degeneration. *Nat Genet.* 2006; 38: 1055–1059.
- [8] Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet.* 2008; 40: 638–645.
- [9] Maher B. Personal genomes: The case of the missing heritability. *Nature* 2008; 456: 18–21.
- [10] Stančáková A, Laakso M. Genetics of type 2 diabetes. *Endocr Dev.* 2016; 31: 203–220.
- [11] Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet.* 2008; 40: 955–962.
- [12] Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet.* 2009; 41: 334–341.
- [13] Hemminki K, Forsti A, Bermejo JL. The 'common disease-common variant' hypothesis and familial risks. *PLoS ONE* 2008; 3: e2504.
- [14] Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet.* 2008; 40: 695–701.
- [15] Patnala R, Clements J, Batra J. Candidate gene association studies: a comprehensive guide to useful *in silico* tools. *BMC Genet.* 2013; 14: 39.
- [16] Szalai Cs, Falus A, Oberfrank F. Medical genomics and bioinformatics. [Orvosi genomika és bioinformatika.] Semmelweis Egyetem, Budapest, 2012. [Hungarian]
- [17] Benjamin J, Li L, Patterson C, et al. Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking. *Nat Genet.* 1996; 12: 81–84.
- [18] Ebstein RP, Novick O, Umansky R, et al. Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking. *Nat Genet.* 1996; 12: 78–80.
- [19] Leung PW, Chan JK, Chen LH, et al. Family-based association study of *DRD4* gene in methylphenidate-responded Attention Deficit/Hyperactivity Disorder. *PLoS ONE* 2017; 12: e0173748.
- [20] Ronai Z, Guttman A, Keszler G, et al. Capillary electrophoresis study on DNA-protein complex formation in the polymorphic 5' upstream region of the dopamine D4 receptor (*DRD4*) gene. *Curr Med Chem.* 2004; 11: 1023–1029.
- [21] Ronai Z, Szantai E, Szmola R, et al. A novel A/G SNP in the –615th position of the dopamine D4 receptor promoter region as a source of misgenotyping of the –616 C/G SNP. *Am J Med Genet B Neuropsychiatr Genet.* 2004; 126B: 74–78.
- [22] Hirschhorn JN, Lohmueller K, Byrne E, et al. A comprehensive review of genetic association studies. *Genet Med.* 2002; 4: 45–61.
- [23] Iafraite AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. *Nat Genet.* 2004; 36: 949–951.
- [24] Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. *Science* 2004; 305: 525–528.
- [25] Haraksingh RR, Abyzov A, Urban AE. Comprehensive performance comparison of high-resolution array platforms for genome-wide Copy Number Variation (CNV) analysis in humans. *BMC Genomics* 2017; 18: 321.

- [26] Hunt R, Sauna ZE, Ambudkar SV, et al. Silent (synonymous) SNPs: should we care about them? *Methods Mol Biol.* 2009; 578: 23–39.
- [27] Zhang Y, Wang D, Johnson AD, et al. Allelic expression imbalance of human mu opioid receptor (OPRM1) caused by variant *A118G*. *J Biol Chem.* 2005; 280: 32618–32624.
- [28] Tetreault M, Bareke E, Nadaf J, et al. Whole-exome sequencing as a diagnostic tool: current challenges and future opportunities. *Expert Rev Mol Diagn.* 2015; 15: 749–760.
- [29] Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet.* 2013; 9: e1003569.
- [30] Pertea M. The human transcriptome: an unfinished story. *Genes (Basel)* 2012; 3: 344–360.
- [31] Hemani G, Knott S, Haley C. An evolutionary perspective on epistasis and the missing heritability. *PLoS Genet.* 2013; 9: e1003295.
- [32] Kooperberg C, Leblanc M. Increasing the power of identifying gene × gene interactions in genome-wide association studies. *Genet Epidemiol.* 2008; 32: 255–263.
- [33] Davis AN, Afshar-Kharghan V, Sood AK. Platelet effects on ovarian cancer. *Semin Oncol.* 2014; 41: 378–384.
- [34] Stone RL, Nick AM, McNeish IA, et al. Paraneoplastic thrombocytosis in ovarian cancer. *N Engl J Med.* 2012; 366: 610–618.
- [35] Horibe Y, Adachi S, Ohno T, et al. Alpha-glucosidase inhibitor use is associated with decreased colorectal neoplasia risk in patients with type 2 diabetes mellitus receiving colonoscopy: a retrospective study. *Oncotarget* 2017; 8: 97862–97870.
- [36] Barone BB, Yeh HC, Snyder CF, et al. Postoperative mortality in cancer patients with preexisting diabetes: systematic review and meta-analysis. *Diabetes Care* 2010; 33: 931–939.
- [37] Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer: a consensus report. *CA Cancer J Clin.* 2010; 60: 207–221.
- [38] Fransgaard T, Thygesen LC, Gögenur I. Metformin increases overall survival in patients with diabetes undergoing surgery for colorectal cancer. *Ann Surg Oncol.* 2016; 23: 1569–1575.
- [39] Osaki M, Okada F, Ochiya T. miRNA therapy targeting cancer stem cells: a new paradigm for cancer treatment and prevention of tumor recurrence. *Ther Deliv.* 2015; 6: 323–337.
- [40] Herczeg Z, Vanya M, Szili K, et al. Genetic and epigenetic factors of polycystic ovary syndrome. [Genetikai és epigenetikai faktorok polycisztás ovarium szindróma esetén.] *Orv Hetil.* 2016; 157: 1275–1281. [Hungarian]
- [41] Caudle KE, Thorn CF, Klein TE, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing. *Clin Pharmacol Ther.* 2013; 94: 640–645.
- [42] Szalai C, Tölgyesi G, Nagy A, et al. Pharmacogenomics of asthma: present and perspective. [Az asthma farmakogenomikája: jelen és perspektíva.] *Orv Hetil.* 2006; 147: 159–169. [Hungarian]
- [43] Brotman MA, Guyer AE, Lawson ES, et al. Facial emotion labeling deficits in children and adolescents at risk for bipolar disorder. *Am J Psychiatry* 2008; 165: 385–389.
- [44] Lenzenweger MF. Endophenotype, intermediate phenotype, biomarker: definitions, concept comparisons, clarifications. *Depress Anxiety* 2013; 30: 185–189.
- [45] Reitz C, Mayeux R. Endophenotypes in normal brain morphology and Alzheimer’s disease: a review. *Neuroscience* 2009; 164: 174–190.
- [46] Dick DM, Agrawal A, Keller MC, et al. Candidate gene–environment interaction research: reflections and recommendations. *Perspect Psychol Sci.* 2015; 10: 37–59.
- [47] Ziller MJ, Gu H, Müller F, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* 2013; 500: 477–481.
- [48] Edfors F, Danielsson F, Hallström BM, et al. Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol Syst Biol.* 2016; 12: 883.
- [49] Kostı I, Jain N, Aran D, et al. Cross-tissue analysis of gene and protein expression in normal and cancer tissues. *Sci Rep.* 2016; 6: 24799.
- [50] Walker E, Hernandez AV, Kattan MW. Meta-analysis: Its strengths and limitations. *Cleve Clin J Med.* 2008; 75: 431–439.

(Rónai Zsolt dr.,
Budapest, Pf. 2, 1428

e-mail: ronai.zsolt@med.semmelweis-univ.hu)

„*Non intellecti nulla est curatio morbi.*” (Cornelius Gallus)
(Mit kideríteni nem bírunk, nincs írja a bajnak.)