

A funkcionális genomikai eszköztár szerepe az onkológiai kutatásokban

Bálint Bálint L.¹, Nagy László^{1,2}

¹Debreceni Egyetem Orvos- és Egészségtudományi Centrum, Biokémiai és Molekuláris Biológiai Intézet, Klinikai Genomikai és Személyreszabott Orvoslási Központ, ²DE-MTA Lendület Immungenomikai Kutatócsoport, Debrecen

Az 1000 genom projekt publikálása döntően megváltoztatja a humán genomról alkotott képünket. A szekvenálási technológiák rohamos fejlődése számos gyakorlati kérdést vet föl, amelyeknek megválaszolásától függ, hogy miként alakulnak a hazai onkológiai kutatások. Cikkünkben áttekintjük az 1000 genom projekt néhány következtetését és elhelyezzük a funkcionális genomikai módszereket a genomikai kutatások metodikái között. A terület fejlődéséből kiindulva végül néhány gyakorlati következtetést vonunk le a következő periódus teendőiről. Magyar Onkológia 57:21-25, 2013

Kulcsszavak: genomika, exomszekvenálás, funkcionális genomika, epigenetika, onkogenomika

The 1000 genomes project changed the way how we see the human genome. The rapid development of the deep sequencing technologies is raising several practical questions, and the way how we answer these questions will affect deeply the future of the oncological research in Hungary. In our manuscript we give a short overview of the results of the 1000 genomes project and we present the place of the functional genomic investigations between other genomic tools. Based on the recent development in the field we summarize the challenges that have to be addressed in the next couple of years.

Balint BL, Nagy L. The place of functional genomics in oncological research. Hungarian Oncology 57:21-25, 2013

Keywords: genomics, exome sequencing, functional genomics, epigenetics, oncogenomics

BEVEZETÉS

A genom mélyebb megismerése átértelmezte azt a tudományismereti keretet, amelyet a kétezres évek elején a Humán Genom Projekt fölállított. A masszív, automatizált kapilláriszekvenálás technológiája bevezetésének egyik azonnali eredménye az EST (expressed sequenced tag) könyvtárak szekvenálása volt, melyek a szövetekben expresszált gének információit tartalmazták. E terület úttörője Craig Venter, aki 1992-ben 2375, agyban expresszált gén szekvenciáját tette közzé (1), és 1993-ban már a masszív EST szekvenálási adatokból azonosított célzottan receptorgéneket (2).

E könyvtárak megismerése döntően hozzájárult a humán genom annotálásához. A humán genom fragmentjeit

tartalmazó BAC könyvtárak szekvenálása és a humán EST könyvtárakból ezzel párhuzamosan extrahált információ tette lehetővé az átíródo szakaszok feltérképezését. A humán genom dekódolása lehetővé tette a gének megismerését, és egyben kitérítette a gén fogalmát. A gének száma alacsonyabbnak bizonyult az eredetileg feltételezett százezres számnál, és nyilvánvalóvá vált a génexpresszióban szerepet játszó genomik lokuszok kiemelt szerepe és elválaszthatatlan voltak a gén fehérjekódoló szakaszaitól.

A megismert gének száma viszonylag alacsonynak bizonyult, és fajoként nem tért el drasztikusan. A megismert genomik információk alapján emberi egyedek közötti különbségeknél elsődlegesen az SNP-k (single nucleotide polymorphism) szerepe került előtérbe. 1998-ban Francis Collins, a különböző eredetű BAC klónokból származó

Levelezési cím: Bálint Bálint L., Debreceni Egyetem OEC, Biokémiai és Molekuláris Biológiai Intézet, 4012 Debrecen, Nagyerdei krt. 98., Pf. 6. Tel.: 06-52-411717/50015, fax: +36 52 255 990, e-mail: lbalint@med.unideb.hu

Közlésre érkezett: 2013. január 2. • Elfogadva: 2013. február 10.

információk alapján az SNP-k számát 17 millióra jósolta, és 2001-re 100 ezer SNP azonosítását prognosztizálta (3). Az eredmények jól követték az eredeti becslést: 2008-ban a non-redundáns (rs kódjelű) humán SNP-k száma elérte a 12 milliót. A mélyszekvenálási technológiák megjelenése gyökeresen átalakította a képet: a 2012 novemberében publikált 1000 genom projekt adatai már 36,7 millió ismert SNP-ről számolnak be és átlagosan 3,6 millió SNP-t jelölnek meg személyenként (4). Általános megfigyelés, hogy minden egyed szekvenálásánál néhány száz addig ismeretlen SNP azonosítására számíthatunk (Marth Gábor, Boston College, személyes adatközlés).

Az 1000 genom projekt adatainak másik figyelemre méltó következtetése, hogy a vizsgált 1092 humán genomban összesen 1,38 millió, személyenként átlagosan 344 ezer inszerciót és deléciót (indelt) azonosítottak. Ehhez az adathoz hozzáadódik a genomszinten azonosított majd 14 ezer nagy méretű genomi variáns, melyekből személyenként átlagosan 717 nagy deléciót azonosítottak (4). Az így összesített adatok gyökeresen megváltoztatták a humán genomról alkotott képünket.

Az orvosi szemlélet a genomról leginkább a „működő gépezet” képének felel meg. Ebben a képben a genom mint egy programcsomag van jelen, amelyből szövevényként más és más funkciókat lehet előhívni. Ebben az elképzelésben a változás eleve egy funkcióvesztést jelent, mely sok esetben dokumentáltan is betegséghez vezet. Az utóbbi évek adatai a genomot egy sokkal plasztikusabb rendszernek mutatják be. A daganatokban megismert mutációk és variánsok a figyelem előterében vannak, és ezáltal nagy mennyiségű adatot generáltak, a folyamatoknak sok részlete vált ismertté. Az utóbbi években megismert genom a nagyszámú variáns miatt nagyfokú redundanciára utal a működésben. Míg némely kritikus gén vagy útvonal mutációja nagy valószínűséggel neopláziához vezet, bizonyítékok szólnak amellett, hogy betegséget okozó variánsok eltűnése gyógyuláshoz vezethet. E mutációkat a természetüknek fogva lényegesen ritkábban azonosítjuk.

A fenti jelenség neve revertáns mozaicizmus, és a szakirodalomban a „natural gene therapy” fogalmával is leírják. Természeténél fogva leginkább bőrbetegségek vizsgálata során válhat nyilvánvalóvá (5). Az egyik legegyszerűbb magyarázata az ilyen gyógyító mutációknak, amikor a betegséget egy frameshift (kereteltolás) okozza (6). Az első kereteltőlástól független második esemény az eredeti olvasási keretet visszaállíthatja. Tekintve, hogy személyenként 344 ezer indelt azonosítottak az 1000 genom projektben, ilyen változások bekövetkezése, ha ritka is, de nem kizárt.

A gének átíródo szakaszai mellett a genomszekvenálás fölvetette a nem kódoló szakaszok jelentőségét is fiziológias és patológias folyamatokban. Az onkológiai genomika területén

a vezető irány ma is a kódoló fehérjékben levő mutációk azonosítása. Ezt a megközelítést nagyrészt az onkogének korai megismerése erősítette és a sejtosztódás szabályozásában szerepet játszó fehérjék szerepének a tisztázása a daganat kialakulásában. E megközelítés igen sikeresnek tudható, és erre alapozva jelentek meg sikeres biológiai terápiák az onkológiai kezelésekben. A genomi asszociációs vizsgálatok onkológiai mintákon már 1995-ben komparatív genomi hibridizációval kimutattak prognózissal összefüggő kópiaszám-változásokat a genomban (7). Onkológiai vizsgálatok már 1998-ban kimutatták a szolid tumorokra jellemző mikroszatellita-instabilitást, mely adatok egyszerre jelezték, hogy bizonyos variánsok összefüggnek a prognózissal, illetve azt is, hogy vannak olyan variánsok, amelyek nem daganatos szövetekben is előfordulnak (8).

A genomszekvenálás kezdeti adataira nagyszámú genomasszociációs vizsgálat épült. Ezek a vizsgálatok bizonyították, hogy sok esetben nem kódoló szakaszokban levő SNP-k betegcsoportokban megnövekedett gyakorisággal mutathatók ki. Ezen SNP-k egyik típusa a transzkripció faktor kötőhelyek módosulását okozza. E módosulatok egy-egy szabályozó fehérje DNS-kapcsolódását befolyásolják. Az 1000 genom projekt az exonokat (GENCODE regions) 80 fölötti mélységben (80-szoros lefedettség) szekvenálta meg személyenként, az intergenikus régiókat ötszörös lefedettségben szekvenálták meg. Az adatok összesítésénél figyelembe vették a korábbi SNP mikroarray adatokat, minden eddiginél pontosabb haplotípus térképek előállítását téve lehetővé.

Az 1000 genom projekt adatai szerint az átíródo szakaszokhoz hasonló konzerváltságot mutatnak a genomban a transzkripció faktorok kötőhelyei. A CTCF inzulátor validált kötőhelyeit vizsgálva kiderült, hogy a motívumok variabilitása megfelel a kódoló tripletek harmadik, lötyögő nukleotidjának, míg a motívumon kívül nincs nagyfokú konzerváltságra utaló jel. Ezen adatok fölvetik annak a kérdését, hogy milyen paradigmák szerint kell dekódolni a genomban tárolt információt. A fehérjekódoló szakaszok egyértelműen a tripletek nyelvén vannak kódolva, ahol egy triplet azonosít egy aminosavat. A genom nagy része azonban nem a tripletek nyelvén íródott, hanem a kötőhelyek nyelvén. Ennek a dekódolása csak funkcionális genomikai módszerekkel lehetséges.

A KLINIKAI GENOMIKA KIHÍVÁSAI

Technológiai és pénzügyi szempontból ma egy egyedi genom szekvenálása elérhetővé vált kutatási célra. Az első genomszekvenálás több százmillió dolláros költségéhez viszonyítva, ma néhány ezer dollárból egy adott humán genom megszekvenálható. Elérhetőek olyan technológiák, amelyek segítségével a reagensár ma már a bűvös ezer dollár alatt van. A teljes genomszekvenálás nagy kérdése a di-

agnosztikai alkalmazhatóság. Ennek a megvalósulása előtt számtalan akadály van a viszonylagos magas áron kívül, melyek közül a legfontosabbak:

1. Időigény: jelenleg a teljes genomszekvenálás átfutási ideje a leggyorsabb technológiák esetében is a mintavételtől a nyers adatok előállításáig hetekben mérhető.
2. Adatértelmezési kihívások:
 - a. A jelenlegi technológiák mellett a viszonylag rövid leolvasású nyers adatokból a genom összeszerelése nem tekinthető még rutinfeladatnak.
 - b. A generált nagy mennyiségű adatból kiszűrni a releváns adatokat igen nehéz feladat, és nem tekinthető megoldottnak.
 - c. Etikai kérdések: a genomszekvenálás esetén a föltett diagnosztikai kérdés mellett nagy mennyiségű releváns, de nem az eredeti kérdést megválaszoló adat kerül azonosításra. Nem tisztázott ezen adatok sorsa, értelmezése, tárolása.
3. Szakemberhiány: jelenleg nincs kiforrott metodika az adatelemzésre. Az elemzőrendszerek igen magasan képzett biológiai, orvosi és informatikai ismeretekkel rendelkező szakembereket igényelnek. Jelenleg egy ország sem rendelkezik olyan számú szakemberrel, amely a technológia által lehetővé tett adatok értelmezését megoldhatja, ha rutin klinikai felhasználásra kerülne a mélyszekvenálási technológia. Hazai szempontból még súlyosabb a helyzet: nincs megfelelő képzési program és megfelelően megfizetett bioinformatikai életpálya, ami a szakember-utánpótlást lehetővé tenné.
4. Infrastruktúra-hiány: a legolcsóbb genomszekvenálást lehetővé tevő készülékek ára nem csökken. Bár a technológia adott, a masszív klinikai szekvenálást lehetővé tevő technológiák még nem kaphatóak kereskedelmi forgalomban. A szekvenálókészülékeken túl jelentős veszély még a megfelelő bioinformatikai kapacitások hiánya. Rutin klinikai genomikai szekvenálás esetén a keletkezett adatmennyiség tárolása, értelmezése, elemzése nagyobb kihívás, mint magának az adatnak az előállítása.
5. Pénzügyi kontextus: az egészségügyi finanszírozás szempontjából csak a hasznosuló adatnak van relevanciája. A teljes genomszekvenálásból származó adat hasznosulása igen alacsony. A mélyszekvenálás technológiája minden eddiginél olcsóbban állítja elő a szekvenanciaadatokat, a klinikai hasznosítás szempontjából azonban ma csupán ezres nagyságrendben ismertek olyan lókuszok, amelyek szekvenálása prognosztikai markernek tekinthető vagy terápiás következménnyel bír. A paradoxon feloldását jelentheti a nagy áteresztőképességű szekvenátorok segítségével előállítani a fókuszált, klinikailag releváns szekvenanciaadatokat nagyobb populációból.

A közvélekedés szerint az ember genomja élete során állandó, ezért egy személy genomját elég egyszer megszekvenálni. A publikált elemzések már igen korán kimutatták, hogy a daganatos betegségek esetében egyértelműen nagyfokú genomi átrendeződésekkel kell számolnunk, melyek időben is és helyileg is változhatnak. A daganatos sejtek előfordulása a mintában inhomogén, ezért a variánsok azonosításánál többszázaz lefedettséget szoktak alkalmazni. Külön kiemelendő a metagenomok szerepe: az emberi szervezet nyálkahártyáin és a gasztrointesztinális rendszerben igen változatos és esetenként betegséggel összefüggő kommenzális baktériumpopulációk azonosíthatóak, melyek időben is változnak. E populációk időbeli változása legegyszerűbben genomikai módszerekkel követhető nyomon.

A szekvenálással kinyerhető adatmennyiség ma meghaladja azt az adatmennyiséget, amely a felgyülemlett klinikai tudás alapján értelmezhető. Ezért a ma elérhető technológiák mellett a célzott genomszekvenálás tűnik egy olyan átmeneti technikai megoldásnak, amely lehetővé teszi célzott, terápiás következményekkel bíró kérdések megválaszolását genomikai módszerekkel.

A ma elérhető genomfrakcionálási megközelítéseket három csoportba sorolhatjuk:

1. hibridizáción alapuló exome/target capture technológiák,
2. PCR amplikonok szekvenálása,
3. funkcionális genomikai frakcionálási technológiák.

Ad 1. A hibridizáción alapuló technológiák nagyon jó eredményekkel használhatóak. Tipikus példa az exomszekvenálás. Ebben az esetben a gének átíródo szakaszaiban található genomi információt szekvenáljuk meg, miután egy hibridizációs technikával a genomból izoláljuk. A hibridizációs oligók lehetnek DNS- vagy RNS-alapúak. A kiindulási minta DNS, melynek vizsgálatával a genom átíródo szakaszaiban levő variánsok azonosíthatóak. A szekvenálás mélysége monogén betegségek vizsgálatánál a szokványos 30-szoros lefedettség. Szomatikus mutációanalízisnél ennek sokszorosára van szükség. A vizsgálat arra a hipotézisre alapul, hogy a fehérjekódoló szakaszokban levő különbségek döntően befolyásolják egy sejt működését. A mai onkológiai genomi megközelítések döntően erre a hipotézisre épülnek. A hibridizációs próbák 80–90%-ban képesek a target régiókat elválasztani a genom többi szakaszától, ezért a szekvenálás hatékonysága a föltett kérdés szempontjából lényegesen nagyobb még, mint a teljes genomszekvenálás esetén.

Tapasztalataink szerint egy-egy személy esetén az exonokban nagyságrendileg 25 ezer jó minőségű egyéni variáns azonosítható. Ennek a technikának egy módosított vál-

tozata a target capture, amikor nem exonok, hanem egyéb releváns adatokra épülő genomi szakaszok hibridizáció alapuló frakcionálása előzi meg a szekvenálást.

Ad 2. A PCR amplikonok szekvenálása tipikusan a sokszorosított multiplexelt PCR esetén költséghatékony megoldás. A szekvenálni kívánt genomi szakaszokat a polimeráz láncreakciós technika (PCR) segítségével amplifikáljuk, és hozzuk létre a szekvenálandó könyvtárat.

A két korábban említett technológia által előállított adat-típus kiindulási pontja a genom annotációs adatai. Ezen adatok alapján van tudásunk a fehérjekódoló szakaszokról vagy a bizonyos betegekben bizonyítottan releváns gének és szabályozó régiók genomi elhelyezkedéséről. A hibridizációs és PCR-alapú genomfrakcionálás hatékonysága kiemelkedő. A szekvenálás esetén az össz-szekvenált adatpontok több mint 90%-a a célterületekről kerül ki.

Ad 3. A funkcionális genomikai megközelítés alapja, hogy a genomot nem az exonok vagy ismert annotált régiók szerint frakcionáljuk, hanem a szabályozó fehérjék kötőhelyeit próbáljuk azonosítani fehérjealapú megközelítéssel. A legelterjedtebb technológia a ChIP Seq vagy kromatin-immunoprecipitációs mélyszekvenálás (chromatin immuno-precipitation with deep sequencing). A technológiát sikeresen alkalmazták onkológiai vizsgálatban (9). A vizsgálat célja ebben az esetben emlődagatok esetében azon genomi lokuszok azonosítása volt, amelyek helyén az ösztrogénreceptor azonosítható volt. A vizsgálat során egyértelműen azonosítható volt, hogy a prognózissal összefüggő módon a genomnak néhány száz ER-kötést mutató helyén más típusú kötőhelyek találhatóak. A jelenség magyarázatát az 1000 genom projekt által is megmutatott nagyfokú genomi variabilitás adhatja. A funkcionális genomikai vizsgálatokkal lehetőségünk van a genomban olyan régiókat azonosítani, amelyek immár célzott szekvenálással egy lényegesen fókuszáltabb kérdés megválaszolását tehetik lehetővé, nevezetesen, hogy azon régiókban milyen genomi polimorfizmusok találhatóak az adott betegekben.

HOGYAN ALAKÍTTJA ÁT A SEKVENÁLÁSI ÁR FOLYAMATOS CSÖKKENÉSE A KLINIKAI TEAMEKET?

Amint a genomszekvenálás költsége tovább esik, és a célzott genomfrakcionálás költsége alá esik, előreláthatóan a teljes genomszekvenálás lesz az elsőre választandó genomikai metodika. E forgatókönyv bekövetkezése esetén a teljes genomszekvenálás mint pénzügyi döntés fog előkerülni, az adatértelmezés azonban ettől még nem oldódik meg.

Előreláthatóan ahogy ma különböző metodikákkal célzottan frakcionáljuk a genomot és csak a releváns szakaszok értelmezését tekintjük célnak, hasonló metodika fog elterjedni, de nem laboratóriumi eszközökre fog épülni, hanem bioinformatikai módszerekre.

A kérdés felveti az információ árát. Ma már messze nem az a releváns kérdés, hogy mennyibe kerül a szekvenálás. A releváns kérdés, hogy milyen hasznosítható információ milyen áron extrahálható egy adott biológiai mintából. Amennyiben genomi információra gondolunk, akkor a mintafeldolgozás területén (DNS-izolálás) masszív árcsökkenésre nem számítunk. A technológiai fejlődés miatt az adatpont-előállítás költsége folyamatosan csökken. Ez a csökkenés azonban nem jár együtt a vizsgálatok árának csökkenésével. Ennek magyarázata részben az, hogy a mintatüztizálás területén egyszerűbb megoldani, hogy több adatpontot generáljunk egy vizsgálattal, mint azt, hogy kisebb reagensfelhasználással azonos adatmennyiséget nyerjünk ki. A cégek könnyebben fejlesztik termékeiket olyan irányba, hogy több adatpontot generáljanak, és nem érdekeltek egy általános költségcsökkentésben. Ez magyarázza az egyre nagyobb felbontású mikrocip- és szekvenálási technológiák erőteljes piaci jelenlétét.

Amennyiben a vizsgálatok árát elemzésnek vetjük alá, a következő adatokat találjuk (1. táblázat): monogén adat előállításának anyagára kapilláriszekvenálással nagyságrendileg 10 euró/minta, és csak nagy mintaszámoknál csökken ez alá az ár alá. Valós idejű PCR-adatsor előállítása adatpontként nem csökken (50 euró/vizsgálat). Mikroarray-adatpontok előállítási ára csökken (150 euró/vizsgálat), és mélyszekvenálás esetében masszívan csökken (néhány ezer euró/vizsgálat). Az információhasznosulás ezzel ellentétes irányú. Míg kapilláriszekvenálás esetében ez az arány 100%, a PCR esetében 90%, addig a mikroarray esetében ez az arány

1. táblázat. Klinikai genomikai vizsgálatok jövőbeni alakulásának áttekintése adatgenerálási és adatfeldolgozási szempontból

	Becsült ár (euró/minta)	Kinyerhető adatmennyiség	Bioinformatikai igény
Kapilláriszekvenálás	10 (stagnál)	100%	Minimális
Kvantitatív PCR	50 (stagnál)	90%	Minimális
Mikroarray	200 (enyhén csökken)	5–15%	Jelentős, de szabványosított
Targeted (pl. exom) szekvenálás	1000 (enyhén csökken)	0,05–0,25%	Jelentős, fejlesztés alatt
Teljes genomszekvenálás	néhány ezer (masszívan csökken)	???	Kihívás

már csak néhány százalék, és a genomszekvenálás esetén további nagyságrendekkel csökken.

A következő évekre, ha előrevetíthetjük ezeket a trendeket, azt láthatjuk, hogy a genomszekvenálás adataiból bioinformatikai módszerekkel fogja a bioinformatikus kibányászni a mindenkor értelmezhető adatokat. Az adatpont-generálás tehát nem a laboratóriumban fog eldőlni, hanem a számítógép mellett. Ebben a fázisban a fenotípus-leírások (patológiai és a klinikai adatok) a genomi adatokkal párba állítva megteremtik majd azt a tudásplatformot, amelyből rövid időn belül újabb áttörésekre számíthatunk. Ennek a feltétele a megfelelően gyors adatbányászati algoritmusok és a megfelelő bioinformatikai infrastruktúra megléte. E forгатókönyvek feltételezik, hogy ahhoz hasonlóan, ahogy a biokémikus és a biológus mára integrálódott a diagnosztikai csoportba, úgy fog a klinikai bioinformatikus is megjelenni. Jelenleg ennek a feltételei hazánkban nem adottak. Az orvostársadalomnak meg kell tennie egy bátor lépést az informatika világa felé.

A GENOMI MEDICINA MEGVALÓSÍTHATÓSÁGA HAZÁNKBAN

Mikor és hogyan érdemes Magyarországnak a kritikus infrastruktúrákat felállítania a genomi medicina megvalósulása irányába?

A cél minden esetben pragmatikus: a kinyerhető adatpontok árát kell először az egészségfinanszírozóknak meghatározniuk. Ezen adatpont előállításának árát a mellékhatásokkal és egyéb minőségi paraméterekkel együtt figyelembe venni. Az adatpontokat csak a fenotípus értelmezi. Azaz a genotípus térképezésénél lehet, hogy fontosabb ebben a fázisban a fenotípus-regiszterek felállítása és szabványosítása. Ez egy szemantikai kérdés, amelynek megoldása emberierőforrás-intenzív és megfelelő informatikai háttérrel igényel. Az informatikai háttérrel összehasonlítható fontosságú a megfelelően precíz információbevitel biztosítása. Jelenleg az orvosi dokumentációk pontossága nem elégséges ahhoz, hogy fenotípusos elemzéseket lehessen rajtuk végezni. Amennyiben a fenotípusadatok rögzítése megoldott, a biobankok rendszerezésére és vizsgálatára lesz szükség. Maga a genomikai adatpont előállítása a teljes rendszernek kisebbik része.

Kitörési lehetőségek:

1. Az egyik evidens lehetőség, hogy a nagy nemzetközi konzorciumok adatai azt mutatják, hogy nincs vagy nagyon limitált az általánosan használható genomikai információ. A vizsgálatoknak populációspecifikusnak kell lenniük, azaz a nagy nemzetközi vizsgálatokat célszerű lokálisan is megismételni, azonos metodikával. Az azonos metodika használata lehetővé teszi az eredmények összehasonlítását és a helyi specifikumok azonosítását is.

2. A másik esélyünk, hogy az eddigi évek forráshiányos egészségügyi körülményei között olyan helyi vizsgálati metodikák alakultak ki, amelyek jó megbízhatósággal a nemzetközi piacon elérhető kitek árának töredékéért képesek azonos információmennyiséget biztosítani. Hiba lenne ezeket a metodikákat veszni hagyni. Minden lokálisan jól működő vizsgálati módszert több független tesztelés után szabványosítva közös márkanevvel lehetne nemzetközi piacon értékesíteni. Ezen értékesítési csatornák léteznek.
3. Adatgenerálási lehetőség infrastruktúra-beruházás nélkül. A nyilvános adatbázisokban elérhető nagy mennyiségű és jó minőségű adatok metaanalízise és matematikai modellezése lehetővé teszi olyan hipotézisek megfogalmazását, melyek viszonylag egyszerűen tesztelhetők. E hipotézisek célzott ellenőrzése nagy nemzetközi vagy hazai centrumok infrastruktúráin töredékáron megvalósítható.

A genomika mai szemlélete szerint tehát a genom önmagában is nagyfokú variabilitást mutat. Erre a DNS-alapú variabilitásra a komplexitást tovább növelő rétegek rakódnak rá, mint az epigenom és a fehérjék poszttranszlációs módosításai. Mindezen komplexitási szintek alapján meglátásunk szerint a genomi medicina gazdagságára még csak rápillanthattunk, és a kutatási feladatok nagy része még alig ismert.

Elmondhatjuk, hogy a genom megismerésének útján csak az első lépéseket tettük meg. A továbbiakhoz elengedhetetlen az orvosi genombiológus, a bioinformatikus, a patológus, az onkológus, a klinikai genetikus és a páciens magas fokú együttműködése. Ennek törvényi, szakmai és gyakorlati keretei még kidolgozásra várnak.

IRODALOM

1. Adams MD, Dubnick M, Kerlavage AR, et al. Sequence identification of 2,375 human brain genes. *Nature* 355:632–634, 1992
2. Venter JC. Identification of new human receptor and transporter genes by high throughput cDNA (EST) sequencing. *J Pharm Pharmacol* 45(Suppl 1):355–360, 1993
3. Collins FS, Brooks LD, Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 8:1229–1231, 1998
4. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65, 2012
5. Choate KA, Lu Y, Zhou J, et al. Mitotic recombination in patients with ichthyosis causes reversion of dominant mutations in KRT10. *Science* 330:94–97, 2010
6. Lai-Cheong JE, McGrath JA, Uitto J. Revertant mosaicism in skin: natural gene therapy. *Trends Mol Med* 17:140–148, 2011
7. Isola JJ, Kallioniemi OP, Chu LW, et al. Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer. *Am J Pathol* 147:905–911, 1995
8. Arzimanoglou II, Gilbert F, Barber HR. Microsatellite instability in human solid tumors. *Cancer* 82:1808–1820, 1998
9. Ross-Innes CS, Stark R, Teschendorff AE, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481:389–393, 2012