

# DNS-chipek alkalmazása az oncohaematológiában

Uher Ferenc

Országos Haematológiai és Immunológiai Intézet, Budapest

A DNS-chipek segítségével végezhető – egyszerre akár több tízezer génre kiterjedő – összehasonlító génexpressziós vizsgálatok teljesen új távlatokat nyitottak az emberi genom felépítésének és működésének a kutatásában. Kérdés, hogy az így nyerhető – esetenként több milliós – adattömeg mennyiben hasznosítható a gyakorlatban? Alkalmasak-e ezek a génexpressziós adatok olyan kritikus gének azonosítására, amelyek hozzájárulhatnak a lympho-haematológiai eredetű tumorok eddigieknél jobb megismeréséhez, pontosabb osztályozásához és eredményesebb gyógyításához? Mivel ma még szinte lehetetlen megjósolni, hogy mennyiben viszi előre a DNS-chipek alkalmazása a daganatbiológiát, összefoglalónkban néhány konkrét példa segítségével szeretnénk bemutatni az összehasonlító génexpressziós módszer lényegét és a benne rejlő potenciális lehetőségeket. *Magyar Onkológia* 45:59-66, 2001

As the Human Genome Project hurtles towards completion, DNA microarray technology offers the potential to open wide new windows into the study of genome complexity. DNA chips can be used for many different purposes, most prominently to measure levels of gene expression (messenger RNA abundance) for tens of thousands of genes simultaneously. But how much of this data is useful and is some superfluous? Can array data be used to identify a handful of critical genes that will lead to a more detailed taxonomy of haematological malignancies and can this or similar array data be used to predict clinical outcome? It is still too early to predict what the ultimate impact of DNA chips will be on our understanding of cancer biology. There are many critically important questions about this new field that are yet unaddressed. By the publication of this article, it is hoped that the technology of DNA chips will be opened up and demystified, and that additional opportunities for creative exploration will be catalysed. *Uher F. Impact of DNA chips on haematological oncology. Hungarian Oncology* 45:59-66, 2001



Még be sem fejeződött a nemzetközi összefogással létrejött humán genom program (HGP = Human Genome Project vagy HUGO), örökítő anyagunk „ABC”-jének a megfejtése, a molekuláris biológusok jórészt már az foglalkoztatja, hogy milyen szavak (gének) rakhatók össze ennek az „ABC”-nek a betűiből és az így kapott szavak hogyan fűzhetők össze értelmes mondatokká (genetikai programokká). A cél tehát örökítő anyagunk működésének – azaz nyelvének és fő-

ként nyelvtanának – a megismerése. Ebben nyújt szinte felmérhetetlen segítséget a DNS-chip (más néven: DNS array, DNS-sorozat vagy lajstrom, gén-chip, genom-chip, stb.) technika. (A *GeneChip*® elnevezéssel vigyázzunk, ez az *Affymetrix, Inc.* tulajdona, márkavédjegy.) Összefoglalónkban azt szeretnénk bemutatni, hogy milyen változásokat hozhatnak a DNS-chipek az oncohaematológiában.

## Néhány szó a chip-technikáról

A DNS-chipek egy szilárd hordozó lapocskához sakktáblaszerű elrendezésben kötött, nagyszámú, különböző nukleotidszekvenciájú DNS-próbából állnak. A próbák lehetnek *in situ*, azaz magának a chipnek (általában nylon) a felszínén

Közlésre érkezett: 2000. október 30.  
Elfogadva: 2000. december 18.

Levelezési cím: Dr. habil. Uher Ferenc,  
Országos Haematológiai és Immunológiai Intézet,  
1519. Budapest, Pf. 424.

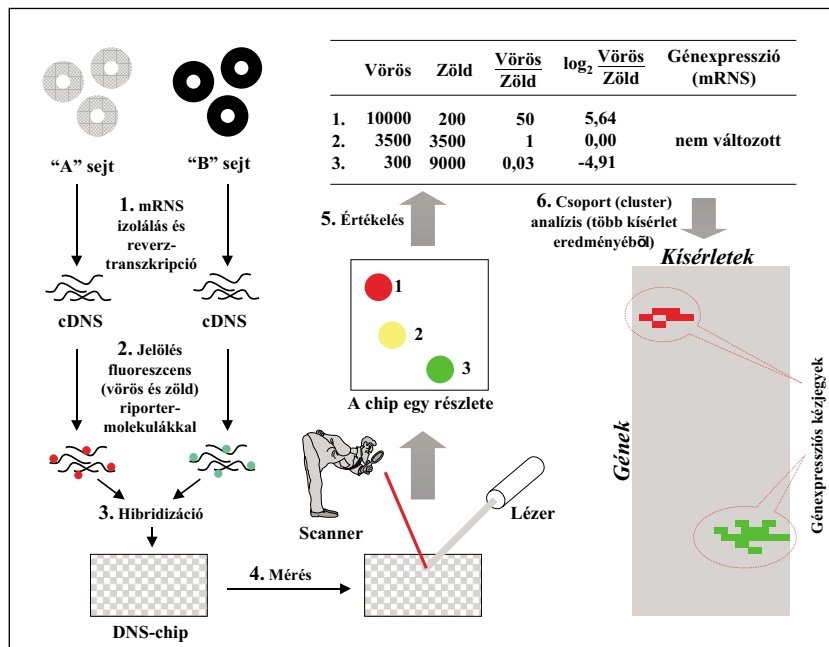
fotolitografikus úton előállított oligonukleotidok (20–25 nukleotid hosszúságúak), vagy nagyobb méretű (~500–5000 nukleotidból álló) DNS-fragmentumok. Utóbbiakat a tintasugaras nyomtatókhöz hasonló elven működő robotberendezés segítségével viszik fel a szilárd hordozó (üveg vagy szilikon) felületére. Mivel a jelenlegi technikai lehetőségek maximum 20–30 ezer, különböző próbát tartalmazó folt vagy pont (spot) kialakítását teszik lehetővé egy chip 1–2 cm<sup>2</sup>-es felületén, egy egyed teljes genetikai állományát reprezentáló – genomiális DNS-próbákból álló – chip csak prokariótákból és egyszerűbb eukarióta szervezetekből készíthető. Kereskedelmi forgalomban vannak például olyan chipek, amelyeken az élesztőgomba (*Saccharomyces*

*cerevisiae*) minden génjét (összesen ~6100 gént) két-két, fehérjére szabadon átírtató, ORF (open reading frame) szekvencia reprezentálja. Az emlős sejtek genomja azonban 60–100 ezer gént tartalmaz. Egyetlen chipen tehát csak a genom egy – a kísérlet céljának megfelelően kiválasztott – részét reprezentáló próbák helyezhetők el. Ezek a próbák általában reverz transzkriptáz enzim segítségével *in vitro* előállított komplementer DNS- (cDNS) láncok. Az ilyen chipeken végezhető el az 1. ábrán bemutatott összehasonlító génexpressziós kísérletek. (A DNS-chip technikának természetesen sok más alkalmazási területe is van: DNS-szekvenálás, pontmutációk kimutatása, gyógyszerérékenység ill. -rezisztencia meghatározása, stb., ezek ismertetése azonban messze meghaladná ennek az összefoglalónak a kereteit) (10, 13, 14, 15, 23).

1. ábra. Egy összehasonlító hibridizációs kísérlet vázlatja. Első lépésként az összehasonlítani kívánt – „A” és „B” típusú – sejtekből mRNS-t preparálnak, majd reverz transzkriptáz enzim segítségével elkészítik az mRNS molekulák kiegészítő DNS (cDNS) kópiáit. A kétféle – „A” és „B” eredetű – cDNS (target) mintákat eltérő színű (vörös vs zöld) fluoreszcens festékekkel (ún. riporter molekulákkal) jelölik, összekeverik és hibridizáltatják a DNS-chip felszínén. A chipeket mosás után lézertűnyel világítják meg és egy scannerrel leolvassák a riporter molekulák által a vörös és a zöld hullámhossztartományban kibocsátott fény intenzitását. A chip felszínén vörösön világító pontok (spots) olyan géneket jelölnek, amelyek az „A” típusú sejtekben nagyobb mértékben expresszálódnak, mint a „B” típusúakban (ld. a chip egy részlete: 1). A zölden világító pontok ennek a fordítottjára utalnak (ld. a chip egy részlete: 3). Ha egy gén mind az „A”, mind a „B” típusú sejtekben hasonló mértékben expresszálódik, akkor a DNS-chip felszínén található megfelelő próbához közel azonos mennyiségű vörös ill. zöld festékekkel jelzett target cDNS kapcsolódik és így sárga színű fénypont jön létre (vörös + zöld = sárga, ld. a chip egy részlete: 2). A génexpresszió relatív mértéke számszerűsíthető is a vörös és zöld fluoreszcencia-intenzitás hányadosának logaritmusaként (ld. az ábrán belüli táblázatot). Számos ilyen kísérlet eredményeinek hierarchikus csoport- (cluster) analizisével alakul ki az a kép (image) amelyen minden sor egy gént és minden oszlop egy kísérletet (vagy mintát) jelent. Erről a képről olvashatók le az adott sejt eredetére, funkciójára és működésére jellemző génexpressziós kézjegyek. (A végső kép kialakítása során a számítógép a sárga foltokat feketére színezi, így a közleményekben és az interneten látható ábrákon a fokozott vagy csökkent génexpresszióra utaló, vörös ill. zöld pontok fekete háttérből emelkednek ki.) (Buhler J (7) és Alizadeh AA és mtsai (1, 2, 3) nyomán.)

Az összehasonlító génexpressziós kísérletek talán legkritikusabb lépése a megfelelő szolubilis nukleinsavminták (targetek) előállítása. Hacsak nem sejtvonalakkal dolgozunk, ritkán sikerül egy mintából, például egy tumorbiopsziából annyi mRNS-t izolálni, hogy reverz transzkripció (RT) után a további munkához szükséges mennyiségű cDNS-hez jussunk. Ilyenkor a target nukleinsavat polimeráz láncreakció (PCR) segítségével kell megsokszorozni. Ezt követi a target cDNS-ek jelölése riporter molekulákkal, amik lehetnek különböző (vörös és zöld) színű fluoreszcens festékek (Cγ5 és Cγ3) vagy radioaktív izotópok (<sup>32</sup>P). A hibridizáció, a chipen inszolubilizált próbák és a komplementer szolubilis targetek közti reakció majd mosás után a chipeket mérik és a kapott eredményeket értékelik.

A mérés a riporter molekulák jellegének megfelelően történik. Fluoreszcens festékek alkalmazása esetén a chipeket lézertel világítják meg és a lapocskák különböző pontjai által kibocsátott fény színét, valamint intenzitását mérik konfokális mikroszkópban vagy valamilyen hasonló elven működő leolvasó készülékben (scannerben). Végül a több chipen végzett számos kísérlet eredményeit hierarchikus csoport- (cluster) analízis (általában a CLUSTER és a TREEVIEW számítógépprogramok) segítségével értékelik. Így meghatározhatók a mintaként feldolgozott sejtek eredetére, funkciójára, működésére, stb. jellemző génexpressziós kézjegyek. Minden kézjegy olyan gének csoportját (clusterét) jelenti, amelyek hasonló vagy inkább egymáshoz kapcsolódó funkciójú fehérjéket kódolnak. Ráadásul minden génexpressziós kézjegy tartalmaz néhány nevesített (named) gént, aminek ismerjük a funkcióját és ebből következtethetünk a csoportba tartozó többi – még ismeretlen – gén funkciójára is. Ha például olyan génexpressziós kézjeggyel találkozunk, amelyben előfordul a Fas (APO-1 vagy CD95) és több kaszpáz enzim génje, biztosak lehetünk abban, hogy a csoportba tartozó többi – fokozottan expresszáló – gén is olyan fehérjéket kódol, amelyek szerepet játszanak a Fas (CD95) – FasL (CD95L) úton indukált apoptózisban (5, 7, 20, 29).



## Egy haematopoeticus őssejt adatbázis

Az első őssejt adatbázist – „Stem Cell Database (SCDb)” – Phillips és mtsai (24, 25) hozták létre. Foetalis (14,5 napos) egérmájából izolálták az AA4.1+Sca-1<sup>+</sup>c-kit<sup>+</sup>Lin<sup>-low</sup> lympho-haematopoeticus őssejteket és az AA4.1<sup>-</sup>, a különböző vörsejt fejlődési sorok irányába differenciálódó, elkötelezett elődsejteket valamint érett vörsejteket. Mindkét – ős- és elkötelezett/érett – sejt populációból mRNS-t szeparáltak, a ribonukleinsav molekulákról DNS-másolatot készítettek, majd a két preparátumból egy szubsztrakciós (ős mínusz elkötelezett/érett sejtek) cDNS-könyvtárat hoztak létre. Ugyanezt megismételték fiatal felnőtt (2–3 hónapos) állatok csontvelőjéből izolált AA4.1+Sca-1<sup>+</sup>c-kit<sup>+</sup>Lin<sup>-low</sup>Rh<sup>low</sup> (ős) és Lin<sup>+</sup> (elkötelezett/érett) sejtekkel is. (A szubsztrakciós cDNS-könyvtárak készítésekor nem alkalmaztak polimeráz láncreakciót (PCR), így a könyvtárak a valódi génexpressziós arányokat (is) tükrözik.) A ~18 000 próbát robotberendezéssel vitték fel nylonmembránok (chipek) felületére. A hibridizációt PCR-rel felszaporított és <sup>32</sup>P izotóppal jelzett targetekkel végezték. 2119, az őssejtekben expresszálódó – eddig jórészt ismeretlen – gént találtak. Számos nukleinsav- és fehérje-adatbázissal összevetve a gének 26%-áról lehetett egyértelműen megállapítani, hogy milyen típusú fehérjéket kódolnak, nevesíteni azonban csak kisebb részüket sikerült. Eszerint 161 gén transzkripciósfaktorokat (pl. ALL-1, AML-1/CBF), 174 sejt felszíni ill. membránhoz asszociált fehérjéket (pl. Flk-2/flt-3, CD34, CD27, Notch1), 28 szekretált fehérjéket (pl. IL-12, MIP2, IL-16), és 147 jeltovábbító molekulákat (pl. Dishevelled-1, Manic Fringe, Ski, DOKL) kódol. A foetalis (máj) és a felnőttkori (csontvelő) lympho-haematopoeticus őssejtekben ugyanazok a gének expresszálódnak, mint foetalis őssejtekben, de néhány tucat gén eltérő mértékben.

### Az SCDb informatikai háttere

Az őssejtekben meghatározott nukleinsav- ill. fehérjeszekvenciákat a BLAST algoritmus segítségével hasonlították össze a SwissProt, a Genbank nr fehérje, a Genbank nr nukleotid, a dbEST, valamint az egér és a humán DOTS adatbázisokkal. A potenciális ORF szekvenciákat az ORF finder (NCBI, Bethesda, MD, USA) programmal azonosították. A fehérjék különböző szerkezeti motívumainak a meghatározását a ProDom (INRA, Franciaország), a Pfam (Washington University, St. Louis, USA), a Prosite (PBIL, Franciaország), a SMART (EMBL, Heidelberg, Németország) és az eMatrix (Stanford University, Stanford, USA) segítségével végezték. A transzmembrán hélixeket a TMPred-del, a szignálpeptideket a SignalP-vel mutatták ki. Az egyes fehérjék sejten belüli lokalizációját a PSORT II segítségével határozták meg (24, 25).

Csak szórványos adatokkal rendelkezünk az érett lymphocyták aktivációja és gátlása során bekövetkező génexpressziós változásokról. Tudjuk

például, hogy a Staphylococcus enterotoxin B – egy T-sejt szuperantigén – intravénás adása után 8 órával az érintett egér T-sejtekben sokkal több gén fokozott expressziója figyelhető meg, mint két nappal később, amikor már megindul az aktivált T-sejtek proliferációja. Ez arra utal, hogy a lymphocyták aktivációja során igen korán – néhány óra alatt – bekövetkeznek a legfontosabb, a sejtek további sorsát meghatározó génexpressziós változások (35). Ugyanakkor a legjobbnak tartott immunszuppresszív gyógyszereink is játékszernek tűnnek. Az FK506-tal gátolt emberi B-sejtekben alig tizedannyi gén expressziója csökken, mint a szervezetben fiziológiás körülmények között anergiássá váló, a szekunder nyirokcsomókban kikerült autoreaktív B-sejtekben (16, 17).

## Malignus lympho-haematopoesis

Az Egyesült Államok Nemzeti Rákkutató Intézete (NCI = National Cancer Institute) által szervezett és irányított „Cancer Genome Anatomy Project (CGAP)” keretében összehasonlító génexpressziós módszerrel vizsgált első tumor a prosztatarák volt. Ma már >13 000, a prosztatában expresszálódó gént ismerünk. Ebből közel 11 000 DNS-szekvenciát normál prosztata-sejtekből (is) azonosítottak. 2800 a praecancerosus és 4100 a malignus prosztata-sejtekben fokozottan expresszálódó gének száma (8, 32, 33). Részben a CGAP-hez kapcsolódva indult el a „Stanford NCI60 Cancer Microarray Project”, aminek a keretében elsősorban azt szerették volna tisztázni, hogy alkalmas-e az összehasonlító génexpressziós módszer az egyes daganatok eredetének a meghatározására is. Hatvan, *in vitro* kultúrában fenntartott malignus sejt vonalat vizsgálva egyértelműen meg tudták állapítani, hogy közülük melyek származnak a központi idegrendszerből, a veséből, a petefészekből, a vastagbélből, a bőrből (melanomák) és a vérképző rendszerből (leukaemiák). Igazolták tehát, hogy lehetséges olyan szerv- ill. szövetspecifikus géncsoportok (cluster) meghatározása, amelyek egyértelműen jelzik az adott tumorsejtek eredetét. Ráadásul ugyanez a kísérleti rendszer alkalmasnak bizonyult az egyes sejt vonalak gyógyszerérzékenységeinek a vizsgálatára is (26, 27, 28).

Golub és mtsai (18) voltak az elsők, akiknek sikerült akut leukaemiás betegek véréből összehasonlító génexpressziós módszerrel kimutatni, hogy a kérdéses tumor myeloid vagy lymphoid eredetű-e, azaz a beteg akut myeloid leukaemiában (AML) vagy akut lymphoid leukaemiában (ALL) szenved-e. Ehhez – mármint a myeloid és lymphoid eredetű tumorok megkülönböztetéséhez – 50-nél kevesebb gén expressziójának az összehasonlító vizsgálatára volt szükség. (A kísérletek során alkalmazott chip eredetileg több mint 6000 próbából állt.) Beigazolódott az is, hogy a HOXA9, egy homeo domén fehérjét kódoló (homeobox) gén (amiről már korábban gyanították, hogy onkogén) fokozott expressziója az AML-es betegek egy kis csoportjában nagyon rossz prognózist jelent. Ezek azok a betegek, akik általában

nem vagy csak korlátozottan reagálnak a kemoterápiára, tehát akikben szinte soha nem sikerül remissziót indukálni.

### A HOX gének

A homeotikus szelektor géneket eredetileg az ecetmuslinca vizsgálata közben fedezték fel. Olyan transzkripció faktorokat kódolnak, amelyek a gyümölcsleveg fejlődése során kulcsszerepet játszanak az egyes testtájak meghatározásában. (Maga a homeo régió az a DNS-motívum, amely a kb. 60 aminosav hosszúságú homeo domént kódolja.) A homeo domén fehérjék evolúciós konzerváltsága meglepően nagy, lényegében minden eukarióta sejtben megtalálhatók az élesztőtől az emberig. A gerincesek homeotikus szelektor génjeit HOX géneknek nevezzük. Emberben a HOX gének négy csoportba: HOXA, HOXB, HOXC, és HOXD rendeződnek. Ezek mindegyike más-más – a HOXA csoport a hetedik – szomatikus kromoszómán található. Az emberi HOX gének által kódolt homeo domén fehérjék azonban nem csak az egyedfejlődés korai szakaszában nélkülözhetetlenek, fontos szerepet játszanak a normális és a malignus haematopoiesis szabályozásában is. Egyes HOX gének fokozott expresszióját már számos különböző leukæmia/lymphoma sejtben megfigyelték. Ráadásul a HOXA9 gén közvetlenül érintett az AML-es betegek egy részében (FAB M2–M4) előforduló t(7;11)(p15;p15) kromoszóma-transzlokációban, aminek következtében egy HOXA9-nukleoprotein (NUP98) fúziós fehérje jön létre (21).

### 1. táblázat.

A „Lymphochip” segítségével meghatározott legfontosabb génextpressziós kézjegyek (1,2 és 3 nyomán)

## A Lymphochip – új lehetőség a non-Hodgkin lymphomák diagnosztikájában és osztályozásában

Az Alizadeh és mtsai (1, 2, 3) által Stanfordban és Bethesdában (USA) kifejlesztett Lymphochip 17 856 génpróbából áll. Ebből 12 069 cDNS a centrum germinatívumokból izolált B-sejtek mRNS-molekuláinak a másolata, 2338 cDNS képviseli a különböző non-Hodgkin-lymphomákat és a B-sejtes krónikus lymphoid leukaemiát (B-CLL), néhány száz próba a normális T és B lymphocyták aktivációja során indukálódó vagy gátlás alá kerülő géneket, 3186 cDNS pedig a lymphocyták biológiájában és/vagy a tumorok fejlődésében fontos szerepet játszó ismert (nevesített) gének azonosítására szolgál. A vizsgálandó mintákból származó szolubilis target cDNS-ek minden kísérletben C $\gamma$ 5-tel (vörös), a 9 lymphoid sejtvonalt kevert (poolozott) mRNS-éből kiindulva szintetizált referencia targetek pedig C $\gamma$ 3-mal (zöld) vannak jelölve. A Lymphochip segítségével – első közelítésben – hét nagyobb génextpressziós kézjegyet sikerült azonosítani a normális és malignus emberi lymphocytákban (1. táblázat). A nevesített génekkel is jellemzett kézjegyek egyértelműen jelzik, hogy a vizsgált mintákban található sejtek: (i) T és/vagy B lymphocyták, (ii) nyirokcsomóból vagy vérből származnak, (iii) nyugvó vagy aktivált állapotban vannak, netán éppen osztódnak. (Meglepő, hogy a  $\gamma$ -interferonnal aktiválható gének két, élesen elkülönülő csoportot alkotnak.)

A Lymphochipre épülő, Lymphoma/Leukæmia Molecular Profiling Project (LLMPP) távolabbi

Génextpressziós kézjegy	A kézjegyeket alkotó géncsoportok néhány fontosabb, nevesített eleme*	
Nyirokcsomó	Monocyta-differenciálódás: NK-sejt-differenciálódás: Kemokinek: Extracelluláris mátrix:	CD14, c-fms NK4 IP-10, MIG, PARC, RANTES Osteonectin, mátrix metalloproteináz-9
T-sejt	Differenciálódás: Antigénfelismerő receptor-komplex: Jeltovábbító molekulák:	CD2 TCR különböző láncai, CD3 $\epsilon$ -lánc LAT, fyn
B-sejt	Differenciálódás: Antigénfelismerő receptor-komplex:	CD19, CD20, immunglobulin J-lánc BCR különböző láncai, CD79b
Immunaktiváció	NF- $\kappa$ B/Rel fehérjék: Citokinek:  Kemokinek: Apoptózisgátló fehérjék:	NF- $\kappa$ B 1, NF- $\kappa$ B 2, RelB, c-Rel IL-1 $\alpha$ , IL-1 $\beta$ , IL-3, IL-6, IL-10, TNF- $\alpha$ , TNF- $\beta$ , SCF, G-CSF, GM-CSF IL-8, MIP-1 $\alpha$ , MIP-1 $\beta$ , MIP-2 $\beta$ , MIP-3 $\alpha$ A1 (a bcl-2 család tagja), c-IAP2 (egy IAP homológ)
Sejtosztódás	Sejtciklus-szabályozó fehérjék:  DNS-replikációs faktorok: Sejtciklust ellenőrző fehérjék:	Cdk4, cdk2; cyclin A, B1, E, F; p16 és p18 (cdk-kináz-gátlók) C, MCM2, dihidrofolát-reduktáz BUB1, CHK1
Interferon 1	Interferonnal indukálható transzkripció faktor: MHC antigének: Antigén-bemutató:	STAT1 MHC I-es gének TAP1, LMP2
Interferon 2	Interferonnal indukálható transzkripció faktorok: MHC antigének: Influenzavírus-rezisztencia:	STAT2, IRF-2, ISGF3 $\gamma$ MHC II-es gének MxB

\*Az itt szereplő rövidítések magyarázatát hely hiányában nem tudjuk megadni. Javasoljuk, hogy az olvasó az általa ismert génekre koncentráljon.

célja egy, minden vérképzőrendszeri tumor genetikai jellemzőit tartalmazó, nemzetközi adatbázis megteremtése. Ez különösen a non-Hodgkin-lymphomák esetében lenne fontos, hiszen ezeknek a tumoroknak még az osztályozása is megadatlan. A jelenleg talán legerőteljesebb, revidált európai-amerikai lymphoma osztályozás (REAL = Revised European-American Classification of Lymphoid Neoplasms) egyes kategóriái – a korábbi Kiel-i és más osztályozásokhoz hasonlóan – sokszor különböző eredetű és prognózisú betegségeket takarnak. Ilyen heterogén kategória például a diffúz nagy B-sejtes lymphoma (DLBCL = diffuse large B-cell lymphoma) is. Kemoterápiával a betegeknek csak körülbelül 40%-a vihető remisszióba és számíthat hosszabb túlélésre, nincs azonban olyan morfológiai, klinikai, immunhisztokémiai vagy genetikai marker, amivel a „jó” ill. „rossz” (pontosabban jobb vagy rosszabb) prognózisú DLBCL-es betegek egyértelműen megkülönböztethetők lennének (22, 31). *Alizadeh és mtsai* (1, 2, 3) 40 ilyen beteg tumorsejtjeinek a genetikai analizisét végezték el a *Lymphochip* segítségével, és két meghatározónak tűnő génexpressziós kézjegyet találtak. Az egyik kézjegy a centrum germinativumokból (CG) származó B-sejtekre jellemző, a másik inkább az *in vitro* aktivált perifériás B-sejtekre. Azok a DLBCL-es betegek, akiknek a tumorsejtjeiben a „CG jellegű” génexpressziós kézjegy figyelhető meg, sokkal jobban reagálnak a kezelésre és hosszabb ideig élnek, mint az „aktivált B-sejt jellegű” génexpressziós kézjegyet hordozók. Az ötéves túlélés 76 vs. 16%, tehát igen különböző a két betegcsoportban.

Mindenesetre elméleti és gyakorlati szempontból egyaránt tanulságos a két meghatározó génexpressziós kézjegy közelebbi vizsgálata. A „CG jellegű” génexpressziós kézjegyet elsősorban olyan gének alkotják, amelyek a normális CG B-sejtekben is expresszálódnak. Ilyen a CD10 és a CD38, az A-myb nukleáris faktor és a 8-oxoguanin DNS-glikoziláz (egy DNS-javító enzim) génje, sőt a BCL-6 onkogén is. Utóbbi fokozott expressziója független a DLBCL-es betegekben sokszor megfigyelhető – a BCL-6 gént is érintő – kromoszóma-transzlokációktól. Egészen más az „aktivált B-sejt jellegű” génexpressziós kézjegy összetétele. Ebben főként olyan gének dominálnak, amelyek biztosítják a sejtek folyamatos aktivációját és megvédik őket a programozott sejthaláltól, az apoptózistól. Az IRF4 (MUM1/LSIRF) például egy, számos lymphoid tumorban transzlokálódó onkogén, amelynek expressziója normális esetben csak átmenetileg, a B-sejtek BCR-en (B-cell receptor = B-sejt antigénfelismerő receptor) keresztüli aktivációjakor fokozódik. A FLIP (FLICE-like inhibitory protein/I-FLICE/Casper) a kaspáz 8 enzim működését gátolja, megvédve a sejteket a Fas (CD95) és más „halálreceptorokon” keresztül indukált apoptózistól. Fiziológiai körülmények között a FLIP is csak a lymphocyták aktivációjának korai szakaszában expresszálódik. A daganatsejtekben mindkét gén folyamatosan és nagy mennyiségben íródik át

mRNS-re. Végül az „aktivált B-sejt jellegű” kézjegy eleme a BCL-2, az egyik legfontosabb apoptózis-gátló gén is. Meglepő módon azonban – legalábbis a *Lymphochip* tanulsága szerint – a BCL-2 expressziója nem függ össze a gént érintő esetleges kromoszóma-transzlokációkkal. (Hozzá kell tennünk, hogy fokozott BCL-2-expresszió mutatható ki a „CG jellegű” génexpressziós kézjegyet hordozó DLBCL-es betegek 29%-ában is. Ez valószínűleg a DLBCL további heterogenitására utal.) (1, 2, 3).

A DLBCL-hez hasonlóan igen heterogén – részben T-, részben 0-sejtes eredetű – betegcsoport a REAL osztályozás szerinti ún. anaplasticus nagy sejtes lymphoma (ALCL = anaplastic large-cell lymphoma) is. A betegek 30%-a hordozza a t(2;5) kromoszóma-transzlokációt, amelynek során egymás mellé kerül az 5. kromoszómán elhelyezkedő nucleophosmin (NPH) és a 2. kromoszómán található anaplasticus lymphoma kináz-1 enzim (ALK vagy CD30) gén. Tapasztalatok szerint az ALCL-csoporton belül az NPM/ALK-pozitivitás vagy az ALK-expresszió inkább a T-sejtes fenotípussal és nodalis eredettel korrelál, kevésbé a 0-sejtes fenotípussal és a cutan kiindulású ALCL-lel (22, 31). *Wellman és mtsai* (38) a már kereskedelmi forgalomban lévő, 588 humán cDNS-próbából álló *Atlas* chip (Clontech Laboratories, Palo Alto, CA, USA) segítségével először harmincegy, B-, T- és myelomonocytoid eredetű sejtvonalat vizsgáltak és megállapították, hogy ez a chip alkalmas a különböző lympho-haematopoeticus eredetű (B, T vagy myeloid) sejtvonalak megkülönböztetésére. Ráadásul az ALCL-eredetű sejtvonalakban (összesen 4 ilyen volt a panelben) jelentős klaszterinexpressziót tudtak kimutatni.

#### A klaszterin

A klaszterin egy 75-80 kD molekulatömegű, két, erősen glikozilált és diszulfid-hidakkal összekapcsolt polipeptidláncból ( $\alpha$  és  $\beta$ ) álló heterodimer molekula. Gyakorlatilag minden emlősejt termelhet és szekretálhat klaszterint, ami a legkülönbözőbb stresszhatásoktól védi a sejteket. Az antiszenz klaszterin cDNS-sel kezelt sejtek rendkívül érzékenyek a hőmérsékleti ugrás, az oxidatív gyökök és a tumor-nekrózis faktor- $\alpha$  által kiváltott apoptózisra. A fokozott klaszterinexpresszió mindezen káros hatások ellen képes – legalábbis részben – megvédeni a sejteket. A klaszterin tehát egy, a hősokk-fehérjékhez hasonló, de szekretált (az extracelluláris mátrixba beépülő?) sajátos dajkafehérje (chaperone). Nem zárható azonban ki, hogy (legalábbis egyes sejtekben) intracellulárisan is aktív. A klaszterin pontos funkcionális szerepének és hatásmechanizmusának a tisztázását erősen megnehezíti, hogy nagyon sokféle ligandumot: apoA-1-et,  $\beta$ -amiloid fehérjét, glutation-S-transzferáz enzimet, komplemenseket, megalint, heparint, immunoglobulinokat, prionokat, *Staphylococcus aureus* baktériumokat, és transzformáló növekedési faktor- $\beta$  receptor molekulákat képes megkötni *in vitro* (39).

Ezután 198 primer, részben T-, részben B-sejt eredetű non-Hodgkin- valamint Hodgkin-lymphomás betegből nyert nukleinsavmintát elemeztek az *Atlas* chippel. A 36, ALCL-es betegből származó minta egyértelműen klaszterin-pozitív volt, míg a többi 162 lymphomás betegből származó minták közül csak kettőben expresszáldott ez a gén. A klaszterin fehérje jelenlétét a biopsziákban minden pozitív esetben sikerült Western blot és immunhisztokémia segítségével is igazolni. (A két „kilógó” eset közül az egyik DLBCL-ként, a másik pedig ún. T-sejt-gazdag nagy B-sejtes lymphomaként került besorolásra.) Megállapították továbbá, hogy a klaszterin gén ill. fehérje expressziója nem korrelál az ALK-éval. A klaszterin tehát egy új, értékes tumormarker lehet a non-Hodgkin-lymphomák osztályozásában (38).

### Korlátok és megfontolások

Az egyes biológiai mintákból preparálható mRNS mennyisége korlátozott és így gyakran nem elegendő egy hibridizációs kísérlet elvégzéséhez. Nem elég tehát az mRNS-eket reverz transzkriptáz enzim segítségével cDNS-re átírni és riporter molekulákkal jelölni, a target nukleinsav mennyiségét is növelni kell. Az ehhez szükséges PCR során azonban a különböző mRNS molekulák megsokszorozódása gyakran eltérő mértékű, ami később, a génextpressziós viszonyok értékelésekor félrevezető lehet. Még nagyobb probléma, hogy a biológiai minták általában nem homogének. Egy nyirokcsomóbiopszia vagy egy leukaemiás beteg vérből izolált sejttömeg a daganatsejtek mellett szép számmal tartalmazhat normális, pusztuló és különböző praecancerosus stádiumban lévő sejteket is. Így könnyen előfordulhat, hogy mRNS preparátumunknak csak egy kis része származik a tumorsejtekből. A DNS-chipen mérhető génextpressziós változások ilyenkor természetesen nem a transzformált sejtekben lezajló folyamatokat tükrözik. Ez a probléma egy új – igaz, megregrága – mintavételi módszerrel, az LCM-mel (LCM = laser capture microdissection) küszöbölhető ki, amelynek lényege, hogy a megfelelően irányított lézersugár segítségével egyetlen sejtől álló mintát lehet egy fólia (pl. nylon) felületére párologtatni és tovább vizsgálni (5, 29).

Ráadásul a nukleinsavhibridizációs módszerekkel (például DNS-chipekkel) mérhető génextpressziós változások mindig mRNS-szint-változást jelentenek. Az életjelenségek, az anyagcsere-folyamatok és szabályozásuk azonban fehérjeszinten valósulnak meg. Kérdés tehát, hogy az adott – fokozott vagy csökkent mértékben szintetizálódó – mRNS-ek egyáltalán átíródnak-e fehérjékre? Ezt a problémát – legalábbis részben – úgy próbálják meg kiküszöbölni, hogy a minták feltárása után ultracentrifugával szétválaszják a poliriboszómákhoz kötött és a szabad ill. csak egy-két riboszómához kötődő mRNS-eket. Mivel translációra csak a poliriboszómákhoz kötődő mRNS-molekulák kerülnek, a hibridizáció során ezeket használják targetként (40).

Ha az emberi genom közel 100 ezer génből áll, akkor ezek a gének a pre-mRNS-molekulák alter-

natív hasításának (splicing) és a poszttranszlációs módosításoknak (glikoziláció, foszforiláció, stb.) köszönhetően 1–2 millió különböző fehérjét kódolhatnak. Mindez elkerülhetetlenné teszi a génextpresszió fehérjeszintű vizsgálatát. A DNS-chipek után a fehérje-chipek, sőt az azonos mintákból nukleinsavat és fehérjét, valamint a különböző molekulák közti (fehérje-fehérje, fehérje-nukleinsav és nukleinsav-nukleinsav) kölcsönhatásokat is mérő chipek korszaka következik. Az első ilyen chipek már klinikai kipróbálás alatt állnak és egy-két éven belül kereskedelmi forgalomba kerülnek. Pillanatnyilag talán a prosztata- és a petefészkek-daganatok korai felismerésére kifejlesztett *ProteinChip<sup>TM</sup>* (*Ciphergen Biosystems, Inc.*, Palo Alto, CA, USA) a legígéretesebb (11, 37).

A DNS-chip technika az első olyan laboratóriumi módszer a biológiában, ahol egyetlen kísérlet vagy kísérletsorozat több millió adatot szolgáltat. Ezek az adatok azonban csak akkor érnek valamit, ha képesek vagyunk feldolgozni, értékelni és áttekinthető formában bemutatni ill. közölni őket. Ez adott esetben nem kis feladat és eddig nem is sikerült maradéktalanul megoldani. A probléma rögtön a chipek mérése után kezdődik: szinte lehetetlen meghatározni a „zaj” és a „jel” arányát. (Sajnos – elsősorban anyagi okokból – a legtöbb kísérletet csak kétszer, jó esetben háromszor ismétlik meg.) Kétséges tehát, hogy milyen mértékű fluoreszcenciaintenzitás-változás (ill. két mért pont (spot) közti különbség) tekinthető szignifikánsnak. A döntés – ma még – teljesen önkényes, függetlenül attól, hogy az adatok feldolgozása során melyik rendelkezésünkre álló számítógépes programot használjuk (4, 20). Magának a programnak a kiválasztása sem egyszerű. Az Eisen által közölt (12) CLUSTER és TREEVIEW (honlap: <http://rana.Stanford.EDU/software/>) programokkal végezhető – az LLMP programot során is alkalmazott – hierarchikus csoport- (cluster) analízis viszonylag egyszerű, nem igényel különösebb matematikai felkészültséget és eredménye többé-kevésbé áttekinthető formában ábrázolható. Hátránya viszont a hierarchikus csoportanalízis – részben kétdimenziós voltából adódó – „merevsége” és a véletlen inverziók lehetőség. Elméletileg sokkal korrektebb a clusteranalízis egy másik típusa, a SOM (self-organizing maps = önszervező térképek) módszere, amelynek lényege, hogy ha  $n$  gén expresszióját vizsgáljuk  $k$  mintában (vagy kísérletben) akkor az eredmények  $n$  pontként értelmezhetők egy  $k$ -dimenziós térben (ld. a GeneCluster programot, honlap: <http://www.genome.wi.mit.EDU/MPR>). Ez a módszer a hierarchikus csoportanalízisnél rugalmasabb és kizárja a véletlen inverziók lehetőségét (34). Matematikus segítsége nélkül azonban nem ajánlatos belevágni és az eredmények „felhasználóbarát” ábrázolása is megoldatlan. Könnyen lehet, hogy a jövő végül az irányított tanuláson alapuló programoké lesz (mint amilyen például az SVM = support vector machines). Ezek rendkívül intelligens programok, amik a „tanulás” során megismert példákat – ismert gének és géncsoportok analízisének az eredményeit – fo-

lyamatosan beépítik önmagukba és a „tapasztalokat” felhasználják a következő chip(ek) értékelésekor (6).

Kérdés persze, hogy milyen irányba fog fejlődni az összehasonlító génexpressziós módszer. A jelenlegi – préseljünk minél több próbát egy chipre és mérjük a lehető legtöbb (ismert vagy ismeretlen) gén expresszióját – stratégia folytatása néhány év múlva, amikor az SCDB-hez és az LLMPP-hez hasonló adatbázisok óriásivá növekednek és többszörösen lefedik az egész genomot, értelmét veszti. Valószínűleg egyre inkább előtérbe kerül majd a funkcionális szempontból összetartozó gének, géncsoportok (a mai genetikai kézjegyek) célzott vizsgálata. Már ma is vannak olyan munkák, amikben a szerzők az egy funkcióhoz (pl. apoptózis) kötődő gének, ill. az egy kiválasztott génhez (pl. c-myc) kapcsolódó további, különböző funkcionális szerepet betöltő (a növekedés, a sejtciklus-szabályozás, a jeltovábbítás és az adhézió irányításában érintett) gének expresszióját vizsgálják (9). Úgy tűnik, hogy ezek a célzott vizsgálatok orvosi diagnosztikai, sőt legtöbbször alap kutatási célokra is elegendőek, mivel az egyes génexpressziós kézjegyek a vártnál sokkal erőteljesebbek (markánsabbak) és az evolúció során sokkal jobban konzerválódtak, mint ezt korábban bárki gondolta volna. (Ember, pontosabban számítógép legyen a talpán, aki vagy ami egy apoptózisra jellemző génexpressziós kézjegyről eldönti, hogy pusztuló élesztősejt vagy haldokló emberi sejt „aláírását” látja-e.) Az ilyen kísérletek értékeléséhez pedig bőven elegendő a viszonylag egyszerű hierarchikus csoportanalízis. Azaz itt is érvényesül a kevesebb néha több törvénye (19).

Összefoglalásként tehát elmondhatjuk, hogy a DNS-chipekkel végzett összehasonlító génexpressziós vizsgálatok új korszakot nyitottak a normális és a malignus haematopoiesis molekuláris szabályozásának a megismerésében. Ugyanakkor azt is látnunk kell, hogy ez a módszer – jelenlegi formájában – még nem alkalmas rutin diagnosztikai feladatok megoldására. Ehhez még számos feltételnek kell teljesülnie:

- Szükség van olyan, az SCDB-hez és az LLMPP-hez hasonló, de jóval átfogóbb és minden laboratórium számára folyamatosan hozzáférhető nemzetközi adatbázisokra, amelyek referencia-ként szolgálhatnak a chipek értékelésekor.
- Meg kell határozni, hogy mikor van valóban szükség összehasonlító génexpressziós vizsgálatra és mikor célszerűbb – DNS-chip vagy más hagyományos módszer segítségével – egyes kiválasztott gének (pl. a BRCA1 vagy p53) mutációinak, esetleg a gén (pl. c-myc) transzlokációjának a kimutatása.
- Általánosan elfogadható mintavételi és adatfeldolgozási eljárás(oka)t kell kidolgozni.
- Valószínűleg olyan, a mainál jóval olcsóbb, összetett DNS-chipeket kell készíteni, amelyek legfeljebb néhány tucat gén expressziója, mutációja és/vagy transzlokációja alapján (mindezt egyetlen chipen érzékelve) képesek fontos információkat szolgáltatni a kérdéses tumor ere-

detéről, várható progressziójáról, netán gyógyszerérzékenységéről is.

Ha a DNS-chip technika is olyan ütemben fejlődik, mint azt az elektronikus chipek esetében szinte már megszoktuk, mindezek a feltételek néhány éven belül teljesülhetnek. Természetesen a fenti megszorítások csak a chip technika ill. az összehasonlító génexpressziós módszer közvetlen diagnosztikai felhasználására és nem a neoplasztikus transzformáció genetikai mechanizmusának megértését valamint új tumormarkerek azonosítását célzó kutatásokra vonatkoznak.

Köszönetnyilvánítás:

Ezt a munkát az OTKA (T030205) támogatta.

Irodalom:

Az irodalomjegyzékben és a cikkben szereplő online kapcsolatok és internet honlapok mindegyike szabadon hozzáférhető és letölthető, beleértve a Michael Eisen által közölt ScanAlyze, Cluster és TreeWiew, valamint a GeneCluster számítógép programokat is. A honlapok természetesen szinte mindenre – a kapható chipekre, műszerekre, stb. – kiterjedő keresdelmi információkat és hirdetések is tartalmazznak.

1. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503-511, 2000
2. Alizadeh AA, Staudt LM. Genomic-scale gene expression profiling of normal and malignant immune cells. *Curr Opin Immunol* 12:219-225, 2000
3. Alizadeh AA. Lymphoma/Leukemia Molecular Profiling Project (LLMPP) (<http://llmpp.nih.gov/lymphoma>)
4. Bassett DE, Eisen MB, Boguski MS. Gene expression informatics – it's all in your mine. *Nat Gen* 21 (Suppl):51-55, 1999 (<http://genetics.nature.com>)
5. Bowtell DD. Options available – from start to finish – for obtaining expression data by microarray. *Nat Gen* 21 (Suppl):25-32, 1999 (<http://genetics.nature.com>)
6. Brown MPS, Grundy WN, Lin D, et al. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc Natl Acad Sci USA* 97:262-267, 2000
7. Buhler J. Anatomy of a Comparative Gene Expression Study. (honlap) (<http://www.cs.washington.edu/homes/jbuhler/research/array>)
8. Cole KA, Krizman D, Emmert-Buck MR. The genetics of cancer – a 3D model. *Nat Gen* 21 (Suppl):38-41, 1999 (<http://genetics.nature.com>)
9. Collier HA, Grandori C, Tamayo P, et al. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc Natl Acad Sci USA* 97:3260-3265, 2000
10. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686, 1997
11. Dutt MJ, Lee KH. Proteomic analysis. *Curr Opin Biotech* 11:176-179, 2000
12. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863-14868, 1998 (<http://www.microarrays.org/software>)
13. Epstein CB, Butow RA. Microarray technology – enhanced versatility, persistent challenge. *Curr Opin Biotech* 11:36-41, 2000
14. Falus A, Váradi A, Raskó I. Az orvosi genetikai diagnosztika új eszköze, a DNS-chip. *Orv Hetil* 139:957-960, 1998
15. Fodor SP, Read JL, Pirrung MC, et al. Light-directed, spatially addressable parallel chemical synthesis. *Science* 251:767-773, 1991

16. Glynne RJ, Akkaraju S, Healy JI, et al. How self-tolerance and the immunosuppressant drug FK506 prevent B-cell mitogenesis. *Nature* 403:672-676, 2000
17. Glynne RJ, Ghandour G, Goodnow CC. Genomic-scale gene expression analysis of lymphocyte growth, tolerance and malignancy. *Curr Opin Immunol* 12:210-214, 2000
18. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531-537, 1999
19. Holter BS, Mitra M, Maritan A, et al. Fundamental patterns underlying gene expression profiles: Simplicity from complexity. *Proc Natl Acad Sci USA* 97:8409-8414, 2000
20. Lockhart DJ, Winzler EA. Genomics, gene expression and DNA arrays. *Nature* 405:827-836, 2000
21. Oostveen JW, Bijl JJ, Raaphorst FM, et al. The role of Homeobox genes in normal hematopoiesis and hematological malignancies. *Leukemia* 13:1675-1690, 1999
22. Pálóczi K, Kelényi G (szerk.). *Non-Hodgkin lymphoma*. Springer, Budapest 1998
23. Pease AC, Solas D, Sullivan E.J, et al. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 91:5022-5026, 1994
24. Phillips RL, Ernst RE, Brunk B, et al. The genetic program of hematopoietic stem cells. *Science* 288:1635-1640, 2000
25. Phillips RL. Stem Cell Database (SCDb) (<http://stemcell.princeton.edu>)
26. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Gen* 24:227-235, 2000
27. Ross DT. Stanford NCI60 Cancer Microarray Project. (<http://genome-www.stanford.edu/nci60>)
28. Scherf U, Ross DT, Waltham M, et al. A gene expression database for the molecular pharmacology of cancer. *Nat Gen* 24:236-244, 2000
29. Shi L. DNA Microarray (Genome Chip). (<http://www.gene-chips.com>)
30. Sherlock G. Analysis of large-scale gene expression data. *Curr Opin Immunol* 12:201-205, 2000
31. Stewart AK, Schuh AC. White cells 2: impact of understanding the molecular basis of haematological malignant disorders on clinical practice. *Lancet* 355:1447-1452, 2000
32. Strausberg RL, Buetow KH, Emmert-Buck MR, Klausner RD. The cancer genome anatomy project. *Trends Gen* 16:103-106, 2000
33. Strausberg RL. Cancer Genome Anatomy Project (CGAP). (<http://ncbi.nlm.nih.gov/ncicgap>)
34. Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc Natl Acad Sci USA* 96:2907-2912, 1999
35. Teague TK, Hildeman D, Kedl RM, et al. Activation changes the spectrum but not the diversity of genes expressed by T cells. *Proc Natl Acad Sci USA* 96:12691-12696, 1999
36. Vochringer DW, Hirschberg DL, Xiao J, et al. Gene microarray identification for redox and mitochondrial elements that control resistance or sensitivity to apoptosis. *Proc Natl Acad Sci USA* 97:2680-2685, 2000
37. Walter G, Büsow K, Cahill D, et al. Protein arrays for gene expression and molecular interaction screening. *Curr Opin Microbiol* 3:298-302, 2000
38. Wellman A, Thieblemont C, Pittaluga S, et al. Detection of differentially expressed genes in lymphomas using cDNA arrays: identification of clusterin as a new diagnostic marker for anaplastic large-cell lymphomas. *Blood* 96:398-404, 2000
39. Wilson M, Easterbrook-Smith SB. Clusterin is a secreted mammalian chaperone. *Trends Biol Sci* 25:95-98, 2000
40. Zong Q, Schummer M, Hood L, Morris DR. Messenger RNA translation state: The second dimension of high-throughput expression screening. *Proc Natl Acad Sci USA* 96:10632-10636, 1999