

dc_718_13

MTA DOKTORI PÁLYÁZAT

Akadémiai Doktori Értekezés

**Transzkriptom szintű adatok alkalmazása a
rosszindulatú daganatos betegek várható terápiás
válaszának és túlélésének előrejelzésére**

Dr. Györffy Balázs, PhD

Budapest, 2013

Tartalomjegyzék

RÖVIDÍTÉSEK JEGYZÉKE	5
1. BEVEZETŐ	6
1.1. Rosszindulatú daganatos betegségek előfordulása és kezelése.....	6
1.2. Biomarkerek a terápiás válasz előrejelzésére.....	7
1.2.1. Tamoxifen és az ösztrogén receptor	8
1.2.2. Trastuzumab és a HER2 receptor	9
1.3. A szisztémás terápiával szembeni rezisztencia	11
1.3.1. In vitro modellrendszerek.....	11
1.3.2. Célzott terápiával szembeni rezisztencia.....	13
1.4. Transzkriptom szintű adatok.....	14
1.4.1. A DNS chip technológia	15
1.4.2. A DNS chipek alkalmazásai.....	17
1.5. Transzkriptomikus adatok rosszindulatú daganatos betegségekben.....	17
1.5.1. Multigénés vizsgálatok az emlő rosszindulatú betegségeiben	18
1.5.1. Multigénés vizsgálatok a melanoma patogenezisében.....	19
1.5.3. Multigénés vizsgálatok a petefészek rosszindulatú betegségeiben	20
1.6. Bioinformatika és biomarkerek	21
1.6.1. Onkológiai bioinformatika a világhálón	21
1.6.2. DNS chipek alkalmazásának lehetőségei	22
1.6.3. Biomarkerek azonosítása nyitott és zárt rendszerben	23
2. KUTATÁSI CÉLOK.....	25
2.1. A legjobb biomarker-jelöltek azonosítása genom-szintű adatok alkalmazásával <i>in vitro</i> kísérletekben	26
2.2. Korábban azonosított biomarkerek igazolása transzkriptom szintű adatok keresztelemzésével és klinikai mintákon való teszteléssel	27
2.3. Bioinformatikai eszközök fejlesztése, amelyek a transzkriptomikus adatok alkalmazását lehetővé teszik további biomarkerek vizsgálata során	28
3. MÓDSZEREK	30
3.1. Kemoterápiás szerekkel szembeni rezisztencia-profil meghatározása.....	30
3.2 A kemorezisztencia génjei melanoma malignumban.....	32
3.3. Túlélés előrejelzése sejtkultúra-modellben azonosított gének felhasználásával.....	34
3.4. A PSMB7 gén oki szerepének igazolása <i>in vitro</i> és <i>in silico</i>	35
3.5. Parallel evolúciós modell a párhuzamosan létrejövő rezisztencia modellezésére	38

3.6. Célzott terápiás szerekkel szembeni rezisztencia.....	40
3.7. A melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása.....	42
3.8. Az emlőrák prognózisával kapcsolatba hozott génlisták keresztelemzése	43
3.9. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének igazolása klinikai mintákban	44
3.10. Microarray adatok előfeldolgozása	45
3.11. Bioinformatikai fejlesztés génexpresszió alapú biomarkerek tesztelésére	46
3.12. Fejlesztések többgénes osztályozók és optimalizált vágópont számítására.....	49
3.13. Online diagnosztikai rendszer fejlesztése	51
4. EREDMÉNYEK.....	54
4.1. Kemoterápiás szerekkel szembeni rezisztencia-profil meghatározása.....	54
4.2 A kemorezisztencia génjei melanoma malignumban.....	57
4.3. Túlélés előrejelzése sejtkultúra modellben azonosított gének felhasználásával	59
4.4. A PSMB7 gén oki szerepének igazolása <i>in vitro</i> és <i>in silico</i>	63
4.5. Parallel evolúciós modell a párhuzamosan létrejövő rezisztencia modellezésére	67
4.6. Célzott terápiás szerekkel szembeni rezisztencia.....	72
4.7. A melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása.....	75
4.8. Az emlőrák prognózisával kapcsolatba hozott génlisták kereszt-elemzése nyers adatok újrafeldolgozása után.....	75
4.9. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének igazolása klinikai mintákban	78
4.10. Microarray adatok előfeldolgozása	81
4.11. Bioinformatikai fejlesztés génexpresszió alapú biomarkerek tesztelésére	84
4.12. Fejlesztések többgénes osztályozók és optimalizált vágópont számítására	87
4.13. Online diagnosztikai rendszer fejlesztése	89
5. MEGBESZÉLÉS	92
5.1. Tizenegy kemoterápiás szerrel szembeni rezisztencia-profil meghatározása 30 rosszindulatú daganatos sejtvonal felhasználásával.....	92
5.2 A kemorezisztencia génjei melanoma malignumban.....	93
5.3. A várható túlélés előrejelzése sejtkultúra modellben azonosított gének felhasználásával antraciklin monoterápiával kezelt betegekben	94
5.4. Genom-szintű adatokkal azonosított PSMB7 gén oki szerepének igazolása <i>in vitro</i> körülmények között.....	96
5.5. Parallel evolúciós modell alkalmazása a párhuzamosan létrejövő kemorezisztencia modellezésére	98
5.6. Öt célzott terápiás szerrel szembeni rezisztencia és érzékenység vizsgálata 45 sejtvonalban.....	99

5.7. Korábbi közleményekben a melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása	100
5.8. Korábbi közleményekben emlőrák prognózisával kapcsolatba hozott génlisták keresztelemzése nyers adatok újrafeldolgozása után	101
5.9. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének igazolása klinikai mintákban	104
5.10. Microarray adatok előfeldolgozása	105
5.11. Online rendszer fejlesztése az emlőrák prognózisát előrejelző génexpresszió alapú biomarker-jelöltek tesztelésére.....	107
5.12. A világhálón keresztül elérhető keresztelemző rendszer továbbfejlesztése többgénes osztályozók és optimalizált vágópont számítására.....	109
5.13. A világhálón keresztül elérhető diagnosztikai rendszer	110
6. A SAJÁT KUTATÁSI EREDMÉNYEK ÖSSZEFOGLALÁSA.....	112
HIVATKOZÁSOK JEGYZÉKE.....	114
AZ ÉRTEKEZÉS ALAPJÁUL SZOLGÁLÓ KÖZLEMÉNYEK.....	132
AZ ÉRTEKEZÉSTŐL FÜGGETLEN IMPAKT FAKTOROS SAJÁT KÖZELEMÉNYEK A PHD FOKOZAT MEGSZERZÉSE UTÁN.....	133
AZ ÉRTEKEZÉSTŐL FÜGGETLEN IMPAKT FAKTOROS SAJÁT KÖZELEMÉNYEK A PHD FOKOZAT MEGSZERZÉSÉIG.....	136
KÖSZÖNETNYILVÁNÍTÁS	137

RÖVIDÍTÉSEK JEGYZÉKE

MDR	multidrog-rezisztencia
ALL	akut limfoid leukémia
ASCO	Amerikai Klinikai Onkológusok Társasága
ATCC	Amerikai Sejtkultúra Gyűjtemény
AUC	görbe alatti terület
BRB	Biometric Research Branch
CIN	kromoszomális instabilitás
EGFR	=>HER
ERBB	=>HER
FDA	USA Gyógyszerfelügyelet
FDR	hibás találati arány (false discovery rate)
FISH	fluoreszcens in situ hybridizáció
GCRMA	GC Robusztus Multi-array Átlag
GEO	Gene Expression Omnibus
GI50	a növekedést 50%-ban gátló koncentráció
GIST	gastrointesztinális stróma tumor
GPLxxx	GEO-ban a gén chip platformok neve
GSExxxx	GEO-ban a sorozatok neve
HCC	májrák
HER	emberi epidermális növekedési faktor receptor
IC50	50%-os gátló koncentráció
LRD	log-arány-eltérés (log-ratio discrepancy)
MAS	microarray elemző csomag
MTT	3-[4, 5-dimethylthiazolyl-2]-2, 5-diphenyltetrazolium bromid
NCCN	National Comprehensive Cancer Network
NIH	National Institute of Health
ÖR	ösztrogén receptor
ÖR2	ösztrogén receptor 2
PAM	predikciós elemzés microarray-okra
PGR	progeszteron receptor
PLIER	Probe Logarithmic Error Intensity Estimate
PSMB7	proteasome beta-type 7 egység
RMA	Robusztus Multi-array Átlag
ROC	Receiver Operating Characteristic
RT-PCR	reverz transzkripció polimeráz láncreakció
SAM	szignifikancia elemzés microarray-okra
SEER	Surveillance, Epidemiology and End Results
siRNS	kicsi interferáló RNS
SRB	sulforhodamin B
SSL	biztonsági protokoll
STR	rövid tandem ismétlődés
TCGA	The Cancer Genome Atlas
TGF	transzformáló növekedési faktor
TMA	szöveti micro array
TOP2a	topoizomeráz 2a
VH	veszélyhányados
VSN	varianca-stabilizáló normalizálás

1. BEVEZETŐ

1.1. Rosszindulatú daganatos betegségek előfordulása és kezelése

A rosszindulatú daganatos betegségek előfordulása világszerte nő. Az EU-28-ban 1990-hez viszonyítva mára több mint másfélszeresére nőtt a daganatos eredetű halálozás. Magyarország a legrosszabb ebben a sorban, nálunk ez alatt az időszak alatt majdnem kettő és félszeresére nőtt a halálozás.

Azonban nem csak a trend szempontjából vagyunk rossz helyzetben, de egyes tumortípusok tekintetében is. Az Európai Unióban a férfiak között Magyarország az első az összes rosszindulatú daganatos betegség, az ajak- és szájüreg-, garat-, nyelőcső-, gyomor-, hasnyálmirigy-, gége-, tüdőrák és a leukémiák szempontjából, második a hererákban és harmadik a bélrendszer rosszindulatú betegségeinek előfordulását tekintve ¹. Nők esetében kicsit jobb a helyzet, itt az összesített statisztikában a második helyen állunk, elsőek vagyunk ajak és szájüreg, garat, gége és a bélrendszer rosszindulatú betegségeiben, másodikak a tüdőrák és Hodgkin-kór előfordulásában és harmadikak a leukémiák terén. Ezek a rendkívül rossz statisztikák előfordulás és mortalitás tekintetében nagyjából megegyeznek.

Az új esetek megoszlása szempontjából hazánkban a leggyakoribbak a tüdőrák (~10 ezer eset/év), a bőrrák (~11 ezer eset), a colorectalis rákok (~9 ezer eset) és az emlőrák (~7 ezer eset). Halálozást nézve a tüdőrák vezet (~7500 haláleset/év), második a colorectalis (~4500 halál/év) és harmadik az emlőrák (~2000 haláleset/év) ².

A betegek gyógyítása során három fő lehetőség áll az orvosok rendelkezésére, amely a sebészi eltávolítás és sugarazás mellett a teljes testre ható (szisztémás) gyógyszeres kezelést tartalmazza. A gyógyszeres kezelés legfontosabb gátja a rendelkezésre álló szerekkel szembeni rezisztencia. Általánosan a rezisztenciát több altípusra lehet osztani, mint szerzett és eredendően meglevő rezisztencia, vagy egy vagy több gyógyszerrel szembeni rezisztencia - ezen utóbbit a szakirodalomban multidrog-rezisztenciának (MDR) nevezzük. A kombinációban alkalmazott terápia fontossága miatt az MDR intenzív kutatások tárgya, már számos, a kialakulásához vezető mechanizmust azonosítottak. Ezek közé tartoznak a gyógyszerek sejten belüli felhalmozódását megakadályozó pumparendszerek (pld. MDR1 gén), a programozott sejthalál (apoptózis) útvonalaiiban résztvevő gének (pld. TP53), a sejtciklus ellenőrzőpontjainak változása, valamint számos további mechanizmus ³. Egyes

tumzorokban ezek akár egyszerre, egymástól függetlenül is kialakulhatnak. Klinikai szempontból a gyógyszerekkel szembeni érzékenység előrejelzése kulcsfontosságú lesz a jövőben, mert a személyre szabott gyógyszeres terápia széleskörű elterjedésének alapja lehet.

1.2. Biomarkerek a terápias válasz előrejelzésére

Az NIH szerint¹ biomarkernek nevezünk egy olyan jellemzőt, amelynek objektív mérése lehetővé teszi, hogy egy normális biológiai folyamat vagy egy patológiai folyamat vagy egy terápias célból alkalmazott farmakológiai kezelésre adott válasz indikátoraként alkalmazzuk.

A rosszindulatú daganatos betegekben alkalmazható biomarkerek minőségi vagy mennyiségi alapú mérési technológiát felhasználó, egy vagy több gént vizsgáló tesztet foglalnak magukba. Ezek a terápias válasz, terápias mellékhatások vagy túlélés előrejelzésére használhatóak fel. Egy adott biomarker lehet „megfigyeléses” vagy „kísérletes” annak függvényében, hogy milyen körülmények között azonosították. Megfigyeléses biomarkereket elsősorban klinikai vizsgálatok során, míg a kísérletes biomarkereket általában alapos preklinikai vizsgálatok során tudunk azonosítani.

A klinikailag is alkalmazott biomarkereknek további három alcsoportját különíti el az amerikai Food and Drug Administration (FDA): a széles tudományos közvélemény által elfogadott, klinikailag elterjedt tesztek jelentik az „ismert igazolt biomarkereket”. A beállított, igazolt teljesítményű biomarkerek az „igazolt biomarkerek”. A széles körben nem használt, független igazolással meg nem erősített tesztek jelentik a „valószínűleg igazolt biomarkerek” csoportját.

Technológia szempontjából a biomarkerek rendkívül sokfélék lehetnek. A mélyebb részletekben való elmélyülés nélkül kiemeljük a patológiai diagnózis során alkalmazott immunhisztokémia (ÖR), FISH (HER2) és szekvenálás (RAS) alapú teszteket. A továbbiakban a klinikailag legelterjedtebben alkalmazott két biomarkert részletesebben is bemutatom.

¹ <http://www.niehs.nih.gov/health/topics/science/biomarkers/>

1.2.1. Tamoxifen és az ösztrogén receptor

Az anti-ösztrogén tamoxifen az első célzott terápiás szer, amelyet az emlőtumor kezelésére először 1977-ben hagytak jóvá. Az ösztrogénnel versenyez a receptor kötőhelyekért, és kötődés után leállítja a sejtciklust a G0 és G1 fázisban, ezáltal megakadályozva a sejtosztódást. Az adjuvánsan alkalmazott tamoxifen kezelés az emlőrák visszatérésének esélyét csökkenti. Az USA-ban az NCCN útmutatók alapján, illetve Magyarországon is az elfogadott irányelveknek megfelelően a tamoxifen a korai és előrehaladott emlőrák kezelésére pre- és poszt-menopauzális betegekben egyaránt alkalmazható. A rosszindulatú daganatos betegek kezelésén kívül az emlőrák kialakulásának megelőzésére is használható ⁴.

A tamoxifen kezelést azért lehet célzott terápiának nevezni, mivel a várható választ előre lehet becsülni az ösztrogén receptor (ÖR) kifejeződése alapján. Csak ÖR pozitív tumorok reagálnak az endokrin terápiára, ahol a kezelés hatására az évente előforduló események aránya jelentősen csökken ($VH=0,62$; $p<1E-05$), miközben az ÖR negatív rosszindulatú daganatok egyáltalán nem reagálnak ⁵. Az ÖR negatív rosszindulatú daganatok rezisztenciáját négy klinikai vizsgálat keresztelemzésével is igazolták ⁶. Meg kell azonban jegyeznünk, hogy az ÖR pozitív betegeknek csak a fele reagál a hormonterápiára ⁷. További nehézség, hogy bár az immunhisztokémia alapú ÖR meghatározás klinikai értéke jelentős, azonban maga a meghatározás jelentős különbséget mutat az egyes laborok között ⁸. Az ÖR státusz meghatározását jelentősen lehetne javítani az objektívebb és könnyebben reprodukálható microarray alapú technológiák alkalmazásával (erre az értekezés későbbi részében részletesebben is visszatérünk ⁹).

Az ÖR-hez hasonló hatásosságot lehet elérni, ha egy olyan gént vizsgálunk, amelynek kifejeződése az ÖR által közvetlenül szabályozott – a legismertebb ilyen gén a progeszteron receptor (PGR). Az ÖR pozitív rosszindulatú daganatok 65%-a PGR-re is pozitív, miközben a PGR pozitív ÖR negatív rosszindulatú daganatok csak az összes betegek 1-2%-át adják ¹⁰. Jóllehet a PGR státusz előre jelzi a terápiás választ, azonban ezen összefüggés nem szignifikáns, ha az ÖR státuszt is bevonták a többváltozós elemzésbe ⁵. Egy tanulmányban, ahol 155.175 emlőrákos nő adatait 1990 és 2000 közötti időszakban vizsgálták, az ÖR negatív PGR pozitív rosszindulatú daganatok arányának folyamatos csökkenését figyelték meg – a szerzők szerint ezen összefüggést a diagnosztikai módszerek javulása eredményezhette ¹¹.

Ezen ellentmondások miatt az Egyesült Királyságban már nem is ajánlott a PGR státusz meghatározása (National Institute for Health and Clinical Excellence: <http://www.nice.org.uk>).

A tamoxifen egy előanyag, amelyet a citokóm P450 két izoformája, a CYP2D6 és a CYP3A4 alakítanak át endoxifenné, amely az aktív formája. Egy nagy retrospektív vizsgálat rövidebb visszaesésmentes túlélést talált azon betegeken, akik a CYP2D6 két polimorfizmusát hordozták ¹². Ezt azonban nem igazolták későbbi vizsgálatok ¹³, és ezért a jelenlegi NCCN és ASCO iránymutatások nem ajánlják a CYP2D6 tesztelését az ideális endokrin-terápiás beállításához ¹⁴.

Az ÖR önmagában nem pozitív biomarker: bár a hiányzó kifejeződés esetén terápiás válasz nem várható, azonban magas expresszió esetén is csak a betegek fele fog reagálni. Jelenleg az onkológiai betegek kezelésében lényegében csak ilyen, úgynevezett „negatív” biomarkereket alkalmazunk klinikailag is.

1.2.2. Trastuzumab és a HER2 receptor

A biomarkerek legsikeresebb alkalmazását egy másik, szintén az emlőrák kezelésében alkalmazott marker esetében dokumentálták, amely nem más, mint a HER2 gén. A HER2 magas kifejeződése először rossz prognosztikai jel volt, azonban a HER2-ellenes terápia megváltoztatta a képet: ma egy HER2 pozitív beteg várható túlélése hosszabb, mint a HER2 negatív betegek átlagos túlélése ¹⁵.

A HER2, a HER, más néven EGFR illetve ERBB család egyik tagja. Összesen négy tirozin kináz tartozik ide (HER1, HER2, HER3 és HER4), amelyekben közös, hogy homo- vagy heterodimerizáció után egy konformáció változás történik, amelynek révén másodlagos jelátviteli molekulák aktiválódnak ¹⁶. A HER jelfogók tartalmaznak egy sejten kívüli részt, amely négy további részegységből, egy sejtmembránt átvelő részből és a sejt belsejében elhelyezkedő tirozin kináz részből épül fel ¹⁷. Aktiváció előfeltétele, hogy egy ligand kapcsolódjon a dimerizációs részegységhez, amely alap (zárt) állapotban két sejten kívüli részegység között helyezkedik el. Azonban az egyes HER jelfogók között jelentősebb eltérések is vannak, a HER2 esetében például nyitott (nem zárt) állapotban van a dimerizációs részegység, a HER3 pedig nem tartalmaz tirozin kináz részegységet. A HER1 jelfogóhoz

kötődik az EGF, az amphiregulin, a TGF alfa, a HER3 jelfogóhoz kapcsolódik a betacellulin, a heparin-kötő EGF, illetve a HER3-at és a HER4-et is aktiválja a neuregulin ¹⁸.

Bár még mindig nem tudjuk, hogy a HER2 jelfogónak pontosan mi a ligandja ¹⁹, ez azonban csak azt jelenti, hogy a liganddal versengő gyógyszert egyelőre nem tudunk kifejleszteni. Egy másik lehetőség a jelfogó dimerizációjának megakadályozása, amelyet a trastuzumab a IV. sejten kívüli alegységhez kötődve gátol meg, ezáltal elnyomja a HER2 jelátvitelt és stabilizálja a jelfogót. Ennek eredményeképpen leállítja a RAS/RAF/MEK/MAPK útvonalat és a PI3K/PIP2/PIP3/AKT útvonalat, amelyeknek a sejtosztódásban, a sejtciklus szabályozásában, a programozott sejthalál elkerülésében van kifejezett szerepük, valamint gátolja az érújdonképződést és antitestfüggő sejt alapú citotoxicitást is indukál ²⁰.

A trastuzumabot elsőként az emlőrák kezelésében vezették be, amelynek alapjául egy randomizált, összesen 469 betegen végzett 3. fázisú tanulmány szolgált. Ebben a kemoterápiával kombinált trastuzumab kezelés majdnem három hónappal hosszabb kiújulásmentes és öt hónappal hosszabb teljes túlélést eredményezett, 20%-al magasabb objektív terápiás válasz mellett ²¹. A trastuzumab hatásosságát további vizsgálatok is megerősítették ²², amelyek nyomán az 1998-as első jóváhagyás indikációját 2006-tól korai emlőrákos betegekre is kiterjesztették.

A trastuzumab hatása csak HER2 pozitív betegeknél érvényesül. Hasonlóan a tamoxifen-ösztrogén jelfogó párosításhoz, a HER2 jelfogó is negatív biomarker, vagyis a kezelésre nem reagáló betegek elkülönítésére alkalmas. A HER2 pozitivitást immunhisztokémiával, FISH-el vagy gén chippel lehet meghatározni (ezen utóbbiról a későbbiekben részletesen is fogunk szólni), a HER2 pozitívitas mértéke egyenesen arányos a trastuzumab kezelésre várható válasz mértékével.

Azon betegek között, akiknek az áttétes betegsége HER2 pozitív volt, a trastuzumab kezelésre várható válasz csak 50% körül alakul ²³ és már az első évben a betegek 15%-ában további áttétek jelennek meg ²⁴. A szerrel szembeni rezisztencia legfontosabb mechanizmusai a HER2 kifejeződés elvesztése, a tömörödött HER2 jelfogó megjelenése ²⁵, a PI3K aktiváló mutáció megjelenése ²⁶ és a PTEN gén funkciójának elvesztése ²⁷.

Meg kell említenünk, hogy emlőrák mellett gyomorrák kezelése során is jelentős sikereket értek el trastuzumab kezeléssel. A HER2 pozitív, előrehaladott gyomorrákos betegek kezelése során 446 beteget osztottak két csoportra, ahol az egyik csoport csak kemoterápiát kapott, a másik csoport kemoterápia plusz trastuzumab kombinált kezelést. A trastuzumab kezelést is kapó betegek teljes túlélése több mint négy hónappal hosszabb volt (11,8 vs. 16 hónap) ²⁸. Ezen eredmények mutatják, hogy a megfelelő biomarkerek az adott rosszindulatú daganat típusától függetlenül, lényegében a tumorsejtek molekuláris jellemzése által univerzálisan bevethetőek.

1.3. A szisztémás terápiával szembeni rezisztencia

A szisztémás terápiával szembeni rezisztencia vizsgálata során elsősorban in vitro kísérletek alkalmazása jön szóba, amelynek hátterét az alábbiakban összefoglaljuk. Jelen bevezető keretein túlmutat a szisztémás kemoterápiával szembeni rezisztencia egyes génjeinek részletes tárgyalása, ezekre a 4. (eredmények) és az 5. (megbeszélés) fejezetben fogunk külön hivatkozni.

1.3.1. In vitro modellrendszerek

A rosszindulatú daganatos betegek műtéti kezelése során az elsődleges feladat és cél a tumoros szövet és sejtek eltávolítása. Azonban más betegséggel szemben ezen sejtek alapvető tulajdonsága, hogy korlátlan osztódásra képesek. Ezeknek a kombinációja teszi lehetővé, hogy elsődleges („primer”) sejtvonalat hozunk létre egy-egy eltávolított daganatból. A leghíresebb ilyen sejtvonala a HELA, és ma már több ezer hasonló, kereskedelmi forgalomban is kapható vonal érhető el.

A betegeken végzett vizsgálatok egy igen komoly problémát tartalmaztak: a vizsgálatok során összehasonlított különböző tumorszövetek különböző korú, eltérő klinikai tüneteket mutató betegekből származnak. Ezen betegek ráadásul a legtöbb esetben nem is kaptak homogén kezelést, ami az egyes azonosított gének és a klinikai paraméterek közötti kapcsolat feltárását nagymértékben megnehezíti. A sejtvonalak nem csak azt teszik lehetővé, hogy olyan kísérleteket végezzünk, amelyek egyébként élő emberben kivitelezhetetlenek

lennének, de teljesen ellenőrzött körülmények között tudunk több különböző vonalat (vagy egy vonalon belül több kezelést) egymással összehasonlítani, és ezáltal lehetővé válik az ezen vizsgált körülménnyel összefüggő gének felderítése a zavaró tényezők kizárása mellett.

A kemoterápiával szembeni rezisztencia vizsgálata során a sejtvonalak felhasználásával lehetőségünk nyílik arra, hogy olyan sejtvonalakat hozzunk létre egy szülői („parental”) vonal felhasználásával, amelyek egy adott kemoterápiás szerrel szemben már érzéketlenné váltak. Az ilyen fejlesztések általában néhány hónap - pár év alatt véghezvihetők. A munka során a sejtvonalakat olyan médiumban kell tartani, amely az adott gyógyszer folyamatosan emelkedő koncentrációjú oldatát tartalmazza. A létrejött rezisztenciát hagyományos sejtproliferációs mérések segítségével lehet dokumentálni, illetve meg lehet határozni, hogy a rezisztens vonal a szülői vonalhoz viszonyítva hányszoros rezisztenciát ért el. Ezen vizsgálatok előnye, hogy az ilyen vonalak egyéb tekintetben nagymértékű hasonlóságot mutatnak az eredeti szülői vonallal, ezért az egyes változások a vizsgált kondícióra specifikus eltéréseket takarhatnak.

A sejtkultúrák kísérletek hátránya azonban, hogy egy sejtvonal szükségszerűen csak egy szövettípust tud képviselni. Eközben tudjuk, hogy egyazon kemoterápiás szereket akár egymástól teljesen különböző tumortípusok esetén is alkalmazzák. A tumortípustól független biomarkerek azonosításához ezért célszerű több különféle sejtvonal egyidejű vizsgálata.

Az értekezés későbbi részében ismertetett és a bevezetőben is idézett vizsgálatok során is sejtkultúra alapú kísérletek adták az *in vitro* kutatás alapját. Ezek során az eltérő sejtvonalak használata lehetővé tette, hogy a rezisztencia szövettípustól független oki tényezőit lehessen azonosítani. Azonban az eredmények egy igen fontos kérdést vetnek fel: mennyire reprodukálható a rezisztencia mechanizmusa egyetlen sejtvonalon belül? Máshogy fogalmazva, ha a sejtvonalakat újra kezelnénk az adott gyógyszerrel egy független kísérletben, akkor ugyanezek a mechanizmusok és gének lennének-e szignifikánsak?

Ezen kérdés azért különösen fontos, mivel a betegekben az áttétes tumorok kezelése során nagyon hasonló kérdéssel kerülünk szembe. Az áttétek egy közös ősből indulnak el, valamint az alkalmazott terápia is azonos mindegyik esetében, hiszen a jelenleg rendelkezésre álló technológiákkal még nem lehetséges a betegek különböző testrészein elhelyezkedő elváltozások eltérő kezelése. Ebben az esetben tehát szintén azonos szelekciós nyomás

nehezedik ezekre a genetikailag nagymértékben hasonló sejtpopulációkra - vajon az egyes sejtpopulációk azonos rezisztencia-mechanizmusokat fognak-e beindítani?

Ennek a kérdésnek a vizsgálatára már az irodalomban is találunk példát. Például a MET gátlószerekkel szembeni rezisztenciát vizsgálták meg három gyomorrákos sejtvonalban ²⁹, amely vizsgálat során legalább két különböző párhuzamos mechanizmust tudtak azonosítani. A szerzők azt is megfigyelték, hogy akár egyetlen sejt is képes több, egymástól eltérő mechanizmust is beindítani a kemorezisztencia elérése céljából. Hasonló eredményekre jutottak további vizsgálatok, ahol az EGFR gátlószerekkel szembeni rezisztenciát vizsgálták ^{30, 31}. Ezekben a vizsgálatokban azonban az alkalmazott sejtvonalak száma relatíve alacsony volt, és nem vizsgálták, hogy a szóba jövő alternatív gyógyszerekkel szemben milyen mértékű keresztrezisztencia alakult ki.

Meg kell jegyezni, hogy a sejtkultúra alapú rendszerek nem tudják figyelembe venni a tumor valamennyi jellemzőjét, mint például a körülötte elhelyezkedő kötőszövetet, az érújdonképződést és az immunrendszer szerepét, amely tényezők önmagukban is jelentősen befolyásolhatják a kemorezisztencia létrejöttét ³². Ez azonban részben a sejtvonalas kísérletek előnye is, hiszen homogén kísérleti körülményeket könnyebb egy egyszerűbb modell rendszerben létrehozni, ami egy célzott kérdés megválaszolását könnyebbé teszi.

1.3.2. Célzott terápiával szembeni rezisztencia

Ma már tudjuk, hogy az onkogének aktivációjától való függés, az úgynevezett „onkogén addikció” a tumorok egyik fő jellemzője, amely a progresszió folyamatát alapvetően befolyásolja ³³. Újabb kutatások arra a meglepő jelenségre is felhívták a figyelmet, amely szerint az intenzív kezelésre adott válaszként megjelenő „gyógyszer-addikció” akár felül is múlhatja az „onkogén-addikciót” ³⁴. Ezen molekuláris mechanizmusok megértéséhez a célzott terápiás szerekkel szembeni rezisztencia vizsgálatával juthatunk közelebb.

Míg a hagyományos kemoterápiák során az összes osztódó sejt pusztítása a cél, a célzott terápiák a tumoros sejtek egy-egy jellemző génjét veszik célba. A legismertebb célzott útvonal a RAS jelátvitel, amely ma már nagyszámú gyógyszer támadáspontja. Jóllehet egyes betegek jól reagálnak a célzott terápiás kezelésekre, azonban az összes betegben nézve az

ezen szerekkel szembeni rezisztencia még felül is múlja a hagyományos kemoterápiás szerekkel szembeni rezisztenciát.

A reverzibilis EGFR gátlószer erlotinibre és gefitinibre az európai és észak-amerikai lakosság kb 10%-a reagál ³⁵, az EGFR és HER2 ellenes lapatinibre a válaszarány 24% ³⁶, a RAF, VEGFR, PDGFR, FLT3 és C-KIT receptorokon ható sorafenibre veserákos betegek 10%-a reagált ³⁷. Az áttétes veserákos betegek kezelésében első vonalban alkalmazott sunitinibra a betegek 31%-a reagál ³⁸. A célzott terápiás szerekre összességében a betegek 10-47%-a mutat objektív választ ³⁹⁻⁴³. A rezisztenciát előre lehet jelezni génmutációk vizsgálatával, mint a *KRAS* ⁴⁴, *PTEN* ⁴⁵, *BRAF* ⁴⁶, és a *PIK3CA* ⁴⁷ gének. Amennyiben a RAS jelátvitel egyik alsóbb tagja aktivált állapotban van, akkor a felsőbb gének gátlását megcélzó terápiáktól nem várható hatásosság ^{48, 49}.

Mutáción kívül azonban génexpressziós változások is összefüggenek a várható válasszal, amint a HER2 esetén azt már az 1.2. fejezetben részletesen is ismertettük. Korábbi, egy-egy gyógyszerre koncentrált vizsgálatok génexpresszió alapú rezisztencia-markereket azonosítottak a multikináz-gátló dasatinib ⁵⁰, a gefitinib ⁵¹ és az erlotinib ⁵² esetében.

Meg kell jegyezni, hogy a célzott terápiák esetében a hagyományos kemoterápiás szerekkel szemben van egy további fontos különbség: az alkalmazott molekulák általában jóval nagyobb méretűek. Ez jelentősen befolyásolhatja a szerek farmakokinetikáját is. Azonban a sejtkultúrák körülmények adta lehetőségeket kihasználva kutatásaink során a molekuláris mechanizmusok vizsgálata volt a cél. A farmakokinetikai különbségekkel összefüggő géneket illetve mutációkat állatkísérletes vizsgálatokkal lehet a jövőben azonosítani.

1.4. Transzkriptom szintű adatok

Az 1.2.1. és 1.2.2. fejezetekben bemutatott két diagnosztikai folyamat esetében egy-egy fehérje (gén) vizsgálatára van szükség. Ez klinikailag könnyen kivitelezhető, azonban alkalmazott kutatás során az összes gén külön vizsgálata a legtöbb esetben nem megvalósítható. Ez nem csak a magas költségekre és munkaigényre vezethető vissza, de a

vizsgálat tárgyát képező szövetekből sem áll a legtöbb esetben olyan sok rendelkezésre, amennyi több tízezer párhuzamos, egymástól független mérés elvégzését lehetővé tenné.

Ezzel szemben viszont egy kutató tipikusan arra kíváncsi, hogy melyik gén az, amelyik legjobban összefügg a vizsgált kórállapottal? Ezen kérdés megválaszolása genom-szintű, vagyis az összes gén egyidejű vizsgálatát lehetővé tevő technológiák felhasználásával lehetséges.

Genom-szintű módszerek ma már lényegében a sejtben jelen levő valamennyi molekuláris szinten rendelkezésre állnak, mint a DNS, az RNS, a fehérje és a funkcionális (enzim)-szintű mérések. Saját vizsgálataink szempontjából itt az összes mRNS szint mérésére alkalmas, másnéven transzkriptomikus gén chipeket emelem ki.

1.4.1. A DNS chip technológia

Az 1990-es évek végén a microarray technológia végre elérhetővé tette, hogy valamennyi gén kifejeződését egyidejűleg vizsgáljuk egy adott szövetben vagy sejtpopulációban.

A gén chipeket DNS chip-nek is szokták nevezni, azonban ez nem arra utal, hogy a DNS mérésére alkalmazzák, hanem az mRNS lesz a mérés előtt DNS-sé átírva a reverz transzkripció folyamán, és az mRNS-ről készült DNS szintjét határozzák meg a chipek. A gén chipek a hibridizáció alapelvén működnek. Ennek lényege, hogy a vizsgálandó szekvenciához kapcsolódó ellenoldali DNS szekvenciát rögzítenek egy szilárd hordozó felületen, amelyhez a vizsgált nukleinsavak ezután komplementer módon hozzákapcsolódnak. A leolvasó rendszerekben általában a hibridizáció előtt jelölik (például fluoeszcensen vagy izotóppal) a vizsgálandó szövetből származó nukleinsavakat. A jelölésre specifikus méréssel nyert jelintenzitás a mintában jelen levő nukleinsav mennyiségével arányos.

A gén chipek két fő típusa az egyféle és a kétféle jelölést tartalmazó chipeket tartalmazza. Egyféle jelölés esetén az adott mintában jelen levő összes gént lemérjük, és ezeket egymáshoz viszonyítjuk. Ilyen például az Affymetrix cég által gyártott gén chipek működése, amelyek során egy fotolitográfiának nevezett folyamat használatával egy fény általi gerjesztés révén a rögzített komplementer szálakat a hordozófelületen építik fel. A

módszer előnye, hogy rendkívül nagy sűrűséget lehet elérni, egy expressziós chip 1,3 millió oligót tartalmaz egy 1x1 cm-es felületen, egy SNP chip pedig 7 millió oligót tud mérni egy 1,3x1,3 cm-es felületen.

A kétszínű chipek esetében a mérés során két különböző mintát összekevernek, amelyek közül az egyik a vizsgálandó minta, a másik pedig referencia-(kontroll)minta. A kontrollminta lehet minden mérésben azonos, vagy pedig a mért mintához kapcsolódó, például kezeletlen szövet. Ezután a teljes keveréket ugyanarra a gén chipre hibridizálják, majd a hibridizáció befejezésekor a leolvasás után a vizsgálandó minta (szín) jelintenzitását a kontroll (másik szín) jelintenzitásához normalizálják.

A gén chip mérések a szükséges gépek magas beszerzési ára miatt jellemzően egy-egy egyetemi központi laborban történnek. A kutató az izolált RNS-t beküldi, a központi laborban elkészül a mérés és a kutató a mérési eredményeket CD-ROM-on kapja vissza. Saját vizsgálataink során a gén chip méréseket a berlini Charité Funkcionális Genomikai Központjában, a Stanford-i Egyetemen és a Semmelweis Egyetemen végeztük el.

A gén chipek elterjedt használatára jellemző, hogy az NCBI Gene Expression Omnibus lerakatában (<http://www.ncbi.nlm.nih.gov/geo/>) jelenleg 11500 különböző platformot regisztráltak, amelyek felhasználásával 930 ezer gén chip mérés nyers adatai érhetőek el. Bár léteznek további array lerakatok is, mint például az European Genome-phenome Archive (<https://www.ebi.ac.uk/ega/>), azonban a GEO az egyetlen olyan, ahol korlátozás nélkül lehetséges az adatok letöltése.

A GEO-n az eddig közölt platformok közül a tíz legelterjedtebb platformból nyolcat az Affymetrix cég készített, és kettőt az Illumina. A legelterjedtebb kettő platform az Affymetrix Human Genome U133 Plus 2.0 Array és az Affymetrix Human Genome U133A Array, ezekkel együtt eddig 110 ezer mintát mértek le. A harmadik helyezett a szintén Affymetrix által készített Mouse Genome 430 2.0 Array (33 ezer minta) és a negyedik az Illumina HumanHT-12 V3.0 expression beadchip. Vizsgálataink során még felhasználtuk az Affymetrix Human Genome U133A 2.0 Array chipet is, amelyből összesen 9 ezer minta áll rendelkezésre.

1.4.2. A DNS chipek alkalmazásai

A 2000-es évek elején jelentek meg az első olyan közlemények, amelyek keretében a transzkriptom szintű mérésekre alkalmas DNS chipeket először alkalmazták klinikai és *in vitro* kísérletek során tudományos kérdések megválaszolására is. Már jelentős számú vizsgálat igazolta, hogy gén chipek segítségével klinikai mintákat is hatásosan lehet osztályozni. A mára már klasszikusnak számító vizsgálatukban Sorlie és munkatársai klinikai altípusokat tudtak génexpressziós mintázatok alapján elkülöníteni ⁵³.

A korábbi és az aktuális klinikai útmutatók is lehetővé teszik több különféle terápia alkalmazását ugyanazon tumortípus (klinikai paraméterek) esetén. Máshogy fogalmazva ez azt jelenti, hogy nem tudjuk előre megmondani, hogy egy adott betegben levő rosszindulatú sejtek fognak-e reagálni az adott kezelési protokollra, ezért a beteget kezelő onkológus döntése lesz a gyógyszerkiválasztás alapja.

Saját kutatásaink szempontjából itt meg kell említenünk azokat a korábbi vizsgálatokat, ahol rámutattak, hogy akár sejtvonalakból ^{54, 55}, akár tumorszövetekből ^{56, 57} lehetséges több olyan, egymástól eltérő génexpressziós mintázatot azonosítani, amelyek képesek a rosszindulatú daganatos betegség kemoterápiára adott válaszát előre jelezni. Azonban ebben az időben a vizsgálatok során jellemzően nem „prediktív” módon magát a gyógyszeres választ, hanem a túlélést, mint „prognosztikai” markert vizsgálták ⁵⁸⁻⁶⁰. Ezek mellett voltak olyan vizsgálatok is, ahol egy-egy ráktípusban vizsgálták a gyógyszeres választ, pld nyelöcső ⁶¹ és colon rosszindulatú betegségeiben ⁶². Egy másik vizsgálatban emlőrákos betegek prognózisát tudták előre jelezni ⁶³. A primer tumor felhasználásával lehetséges túlélést ⁶⁴ és docetaxellel szemben várható érzékenységet ⁶⁵ előre jelezni.

Ezen vizsgálatok már felvetették annak lehetőségét is, hogy gén chipek felhasználásával diagnosztikai eszközöket fejlesszünk ki, amelyeket közvetlenül betegdiagnózis folyamán is fel lehet használni. Ez lényegében mára már realitássá vált, amit az utolsó fejezetben részletesen is ki fogunk fejteni.

1.5. Transzkriptomikus adatok rosszindulatú daganatos betegségekben

Az alábbiakban az emlő, a melanoma és a petefészek rosszindulatú betegségeihez kapcsolódó transzkriptomikus vizsgálatokat röviden áttekintjük. A többi daganatos betegség

elemzésétől eltekintünk, mivel eddigi saját vizsgálataink során csak ezen három szerv vizsgálatára szorítkoztunk.

1.5.1. Multigénos vizsgálatok az emlő rosszindulatú betegségeiben

A nagy klinikai relevancia miatt az emlő az egyik legjobban vizsgált tumortípus, és az első transzkriptomikus adatokat is az emlő rosszindulatú betegségeinek vizsgálata során generálták.

Az emlőrák vizsgálatára számos többgénos prognosztikai és prediktív tesztet fejlesztettek ki, mint az RT-PCR alapú Oncotype DX ⁶⁶, Theros Breast Cancer Index ⁶⁷, Breast bioclassifier ⁶⁸, Celera metastatic score ⁶⁹, 8-génos pontrendszer⁷⁰ és a gén chip alapú Mammaprint ⁵⁸, MapQuant Dx ⁷¹, BLN teszt ⁷², Invasive Gene Signature ⁷³ és a Wound Response Indicator ⁷⁴, valamint a FISH alapú eXagen teszt ⁷⁵ és az IHC alapú Mammostrat ⁷⁶. Ezeket a később részletesen is bemutatott keresztelemzésünk során is teszteltük ⁷⁷. Ezek közül az Oncotype DX volt a legsikeresebb, több mint 200 ezer eladott teszttel.

A nyirokcsomó-negatív, ösztrogén receptor pozitív tumorokra kifejlesztett Oncotype DX teszt 21 gén kifejeződését határozza meg, amelyek közül 5 háztartási gén. A teszthez a mintákat egy központi laboratóriumba kell eljuttatni, ahol az izolálást, a mérést és a kiértékelést végzik. Maga a teszt egy matematikai egyenletet használ fel az úgynevezett „recurrence score” kiszámítására ⁶⁶, amellyel a várható kiújulás mentes időtartam hosszát lehet meghatározni ⁷⁸. A várhatóan hamar visszatérő tumorok esetében az egyébként csak opcionális kemoterápia alkalmazásával lehet javítani a beteg kilátásain ⁷⁹. Különböző elemzések szerint a teszt a klinikai döntéshozatalt 25-44%-ban befolyásolja ^{80 81}. Azonban mind a mai napig nem készült se független igazolás, se reprodukálhatósági vizsgálat.

Meg kell azonban azt is jegyeznünk, hogy a multigénos prognosztikai tesztek kifejlesztésének alapját adó első transzkriptomikus vizsgálatokat számtalan kritika érte. Az egyik kritikus pont, hogy a tanulmányok nagy számában a tanulóhalmaz átfedett a teszthalmazzal ⁸². További gyenge pont a statisztikai elemzések eredményeinek túlzott optimista becslése az algoritmusok túloptimalizálása miatt ⁸³. Nagyszámú gén vizsgálata a betegek egy kis csoportjában a véletlen miatt szignifikánsként osztályozott mintázatok számát nagymértékben növelheti ⁸². Amikor véletlenszerűen generált génlisták hatásosságával

hasonlítottak össze génexpressziós mintázatokat, akkor hét mintázatból 5 nem teljesített jobban, mint a random gének csoportja ⁸⁴.

Eközben lassan az is evidencia lett, hogy a nagyobb mintaszámú vizsgálatok megbízhatóbb eredményt adnak, a statisztikai módszerekkel történő túloptimalizálást a módszerek a priori kiválasztásával meg kell előzni és hogy a legjobb igazolás csak független beteganyag felhasználásával lehetséges.

1.5.1. Multigénés vizsgálatok a melanoma patogenezisében

A melanoma malignum az egyik legveszélyesebb tumortípus. Bár az esetek abszolút többségében az elváltozást korán felismerik és a műtét gyógyító eredménnyel zárul, azonban azokban az esetekben, amikor áttét alakul ki, a prognózis nagyon rossz. Az átlagos túlélés mindössze 6-9 hónap ⁸⁵, az ötéves túlélés pedig 5% alatt van ⁸⁶. A betegség során alkalmazott kemoterápiás kezelések hatásossága elhanyagolható, a 2000-es évek során legelterjedtebben alkalmazott dacarbazine terápiára a betegeknek csak egy minimálisan kicsi aránya ad hosszan tartó választ ⁸⁷.

A tumorok geneziséről alkotott átfogó hipotézisek ⁸⁸ alapján a halmozódó genetikai hibák vezetnek a melanoma kialakulásához is. Amennyiben a kialakulás során sérült, aktivált vagy egyéb módon változott génekhez kapcsolódó útvonalakat fel tudjuk deríteni, akkor nemcsak a melanoma létrejöttét érthetjük meg jobban, de olyan géneket is azonosítani tudunk, amelyek később terápiás célpontok lehetnek. A 2000-es évek elején már közöltek olyan transzkriptomikus adatot alkalmazó tanulmányt, ahol a teljes genom vizsgálata alapján azonosított gének olyan csoportját írták le, amelyek a melanoma pathogenezisével kapcsolatba hozhatóak ⁸⁹.

Itt röviden meg kell említenünk az 1999 óta elérhető gén-ontológiai projektet, amelynek célja, hogy a géneket, a funkciókat és a gének által szabályozott folyamatokat egy olyan standardizált rendszerbe illesszük be, ami az egyes organizmusok és adatbázisok összehasonlíthatóságát lehetővé teszi. A projekt során valamennyi génhez hozzárendelik a kulcsszavak egy csoportját, amelyek az adott génre jellemzőek. A gén-ontológiai rendszer három szintet tartalmaz, amelyek a biológiai és molekuláris funkció, valamint a sejten belüli

elhelyezkedést tartalmazták. Vizsgálatainknál a melanoma transzkriptomikus adatainak elemzése során a gén-ontológiai projekt osztályozását is felhasználtuk.

1.5.3. Multigénus vizsgálatok a petefészek rosszindulatú betegségekben

A nőgyógyászati tumorok között a leggyakoribb halálok a petefészekrák, amelyben az ötéves teljes túlélés mindössze 10-30% között mozog. A petefészekrák klinikai kezelése során az elmúlt évtizedek alatt más tumortípusokhoz hasonlítva minimális előrelépés történt. Ennek egyik oka lehet a petefészekrákok nagyfokú heterogenitása, amely klinikailag a különböző grade és szövettani altípusok képében jelenik meg. Ezek a progresszió és a terápiára adott válasz szempontjából jelentős különbségeket mutatnak: a magas grade-ű tumorok gyorsan nőnek, de érzékenyek kemoterápiára, míg az alacsony grade-jű tumorok lassan nőnek és kemorezisztensek. Klinikailag a magas agresszivitású szerózus altípus a legjelentősebb, mivel a tumorok 80%-a ebbe a csoportba kerül. Ezek a tumorok jobban reagálnak a platina- és taxán alapú kemoterápiákra.

Eközben a petefészekrákok döntő többsége klonális sejtproliferáció jellemzőit mutatja, vagyis az egyes genetikai változások egyetlen sejtben történt mutációra vezethetőek vissza ⁹⁰. Ez a klonalitás a tumorkezdő sejt jellemzői alapján az egyes altípusok közötti molekuláris különбözőség alapját is adhatja. Ezen különбözőség a gén chip vizsgálatok után elvégzett fürtelemzés során azt eredményezi, hogy az egyes altípusok egymás mellé kerülnek ^{91, 92}.

Több korábbi közleményben felvetették, hogy a petefészek rosszindulatú betegségeinek gén chip alapú teljes transzkriptomot érintő elemzésével lehet molekuláris biomarkereket azonosítani. Ezekben a petefészek rosszindulatú daganatainak létrejöttével ⁹³⁻¹¹³, szövettani altípusaival ^{95, 102, 114, 115}, és a terápiás válasszal ¹¹⁶⁻¹²² összefüggő, valamint prognózist és progressziót előrejelző ^{104, 107, 123-127} többgénes mintázatokat írtak le. A fenti genomikus, illetve többgénes biomarkeret azonosító vizsgálatok fő gyengesége, hogy a tanulólthalmaz létrehozásához felhasznált minták száma alacsony volt, illetve a független igazolás hiányzott. Egy saját *(jelen értekezésben külön nem részletezett)* vizsgálatunkban Affymetrix gén chippek felhasználásával mi magunk is azonosítottunk prognosztikus génexpressziós mintázatot 120 petefészekrákos minta elemzése során, amelyet egy független beteganyagon is igazolni tudtunk ¹²⁸.

A széles körű vizsgálatok ellenére továbbra sem állnak rendelkezésünkre az emlőrákhoz hasonlóan alkalmazható biomarkerek, melyekkel a szisztémás kezelést optimalizálni lehetne. A progresszió figyelésére az FDA jóváhagyta a CA125 és a HE4 géneket. Számos további gént javasoltak, mint p21¹²⁹⁻¹³¹, a bax^{132, 133}, a hTERT¹³⁴, a survivin¹³⁵, a VEGFR¹³⁶, a p53^{137, 138}, a kallikrein 6¹³⁹ és 10¹⁴⁰, az IL-6¹⁴¹, a p27¹⁴²⁻¹⁴⁵, a cyclin D1^{129, 146}, D3¹⁴⁷ és E¹⁴⁸⁻¹⁵¹, a Bcl-xL¹⁵², a cIAP¹⁵³, és az ERBB1^{154, 155}. Ezekre azonban nincsenek egyértelmű javaslatok és klinikai iránymutatások, ezért szükség lesz a biomarker-jelöltek további vizsgálatára.

Összességében elmondható, hogy a petefészek rosszindulatú daganatai esetén egyrészt a klinikai paraméterek illetve a molekuláris tesztek alapján készült osztályozók hatásossága rendkívül alacsony, másrészt az alkalmazott terápiák szintén alkalmatlanok a legtöbb beteg életkilátásának tartós javítására. Ezért továbbra is különösen nagy szükség van az egyes biomarker-jelöltek független igazolására.

1.6. Bioinformatika és biomarkerek

1.6.1. Onkológiai bioinformatika a világhálón

Ma már az onkológus, patológus illetve egyéb egészségügyi szakemberek diagnosztikai és terápiás döntéseit az interneten keresztül elvégzett tesztekkel is lehet támogatni. Tipikus példa ezekre a tesztekre a patológusok esetében a digitalizált szövettani metszetek szoftveres elemzése, valamint ezen képek távoli vizsgálata egy második patológus által.

Az onkológiában alkalmazott egyik első online elérhető prognosztikai rendszer az „Adjuvant! Online” (<http://www.adjuvantonline.com>), amely a kiegészítő gyógyszeres terápia (kemoterápia, hormonterápia vagy a kettő kombinációja) várható rizikóit és előnyeit tudja megbecsülni emlőrákos betegekben¹⁵⁶. A rendszer célja, hogy megbecsülje a szisztémás kezelés elmaradása esetén a negatív kimenetel rizikóját, vagyis a rosszindulatú daganatos betegséggel összefüggő mortalitást vagy kiújulást; ezen rizikó várható csökkenését az adott terápia alkalmazása mellett és a kezelés mellékhatásainak kockázatát. A klinikailag is igazolt¹⁵⁷ Adjuvant!Online a beteg klinikai jellemzőit használja fel, mint a kor, a tumor mérete, nyirokcsomók érintettsége, szövettani grade stb. A rendszer alapja, hogy a vizsgált beteg

klinikai paramétereit összehasonlítja a SEER-ben mért nagyszámú kezelt és kezeletlen beteggel és ezen adatok alapján lényegében egy epidemiológiai összehasonlító vizsgálatot végez. A SEER („Surveillance, Epidemiology and End Results”) az USA Nemzeti Rákkutató Intézete által életre hívott program, ahol az USA lakosságának 28%-át folyamatosan követik, és valamennyi klinikai, demográfiai, kezelési, mortalitási stb. adatot egy központi adatbázisban tárolnak. Az online elérhető (<http://seer.cancer.gov/>) adatbázis lehetővé teszi, hogy az USA-ra vonatkozó részletes epidemiológiai vizsgálatokat végezzenek ¹¹.

1.6.2. DNS chippek alkalmazásának lehetőségei

A molekuláris technológiák fejlődése mára lehetővé tette, hogy ne csak klinikai paraméterek legyenek biomarkerek, hanem génexpresszió vagy szekvencia alapú tesztek is bekerülhessenek a klinikai gyakorlatba. Erre konkrét példa a bevezetőben részletesen elemzett ösztrogén receptor és a HER2 receptor. Az irodalomban eddig számos további gént azonosítottak, amelyek szintén szóba jönnek, mint onkológiai biomarkerek ¹⁵⁸. Kérdés, hogy ezeket milyen technológia segítségével lehet meghatározni?

DNS chipket, PCR-t és több további technológiát elemző összehasonlító vizsgálat igazolta, hogy a gén chippek által mért génexpresszió megbízható és reprodukálható ^{159, 160}. Technikailag a gén chippek valamennyi gén kifejeződését tartalmazzák, ami felveti, hogy alkalmazásukkal olyan teszteket is végezzünk, amelyekre eredetileg nem is tervezték az adott eszközt. Ide tartoznak azon mérések is, amelyekhez nem több, hanem csak 1-1 gén vizsgálatát tervezték, mint a jelenleg legelterjedtebb ösztrogén receptor és HER2 receptor szintjének mérése.

Ezen tesztek során kérdésként merül fel, hogy milyen vágópont alkalmazásával válasszuk szét a betegeket két (vagy több) csoportba. Bár az irodalomban a legtöbb esetben a medián (illetve alsó/felső negyed és harmad) alapján osztják a betegeket két csoportba, azonban nincsen olyan protokoll, ami egyértelmű iránymutatásként szolgálna. A medián használatának az átlaggal szemben megvan az az előnye, hogy nem függ az esetleg nagy eltérést mutató szélső értékektől. Bár Cox regresszió alkalmazásával lehetséges lenne az adott biomarker-jelöltek hatását fix vágópont nélkül is vizsgálni, azonban ebben az esetben nem tudnánk semmilyen könnyen értelmezhető grafikont rajzolni. Irodalmi adatok alapján RT-

PCR vagy immunhisztokémia felhasználásával lehetséges a vágópontok optimalizálása ¹⁶¹, azonban ezt minden génre külön el kell végezni.

A gén chipek felhasználásával értelemszerűen egy-egy vagy több gént tartalmazó mintázatok is elemezhetőek. Több gént egyidejűleg felhasználó („multigén”) mintázatok ma már egyre szélesebb körben jelentek meg. Ezekben tehát nem egy-egy gén izolált szerepét, hanem egy teljes jelátviteli utat, vagy egy transzkripció faktor által szabályozott valamennyi gént lehet vizsgálni.

Az emlőrák diagnosztikája során használt két legfontosabb gén a bevezetőben már részletesen bemutatott ösztrogén és HER2 receptorok. Mivel mind az ÖR, mind HER2, valamint a „recurrence score” számításához szükséges valamennyi gén szerepel egy teljes genomot mérő Affymetrix gén chipen is, ezért technikailag lehetséges a gén chipek felhasználása az ezen gének mérésén alapuló diagnosztikai tesztek elvégzéséhez. Fontos kiemelni, hogy az előző fejezetekben idézett további kutatások során is a mérési adatok a legtöbb bemutatott projektben gén chip alapúak voltak.

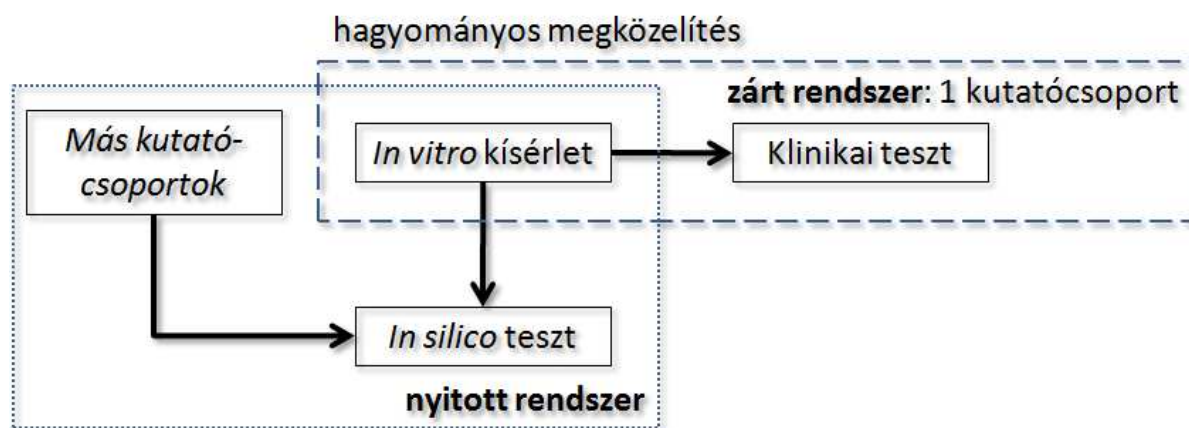
Eközben az Affymetrix chipek diagnosztikai felhasználását már az Amerikai Gyógyszerfelügyelet is jóváhagyta ². Ezek alapján logikus következő lépés a sok gén mérésére alkalmas gén chipek felhasználása betegdiagnózis készítéséhez. Ehhez azonban megfelelő informatikai (statisztikai) háttérre van szükség, amely a legtöbb kórházban vagy egyetemi klinikán nem áll rendelkezésre.

1.6.3. Biomarkerek azonosítása nyitott és zárt rendszerben

Általában a kutatócsoportok a biomarkerek in vitro azonosítása után egy további vizsgálatot végeznek, ahol egy klinikai mintagyűjtés keretein belül annyi szövetmintát gyűjtenek össze utánkövetési adatokkal, amennyi már lehetővé teszi, hogy az adott biomarkert ebben a csoportban is igazolják. Mivel itt a kezdeti kísérletet és a ráépülő tesztelési vizsgálatokat is ugyanaz a csoport csinálja, ezért ilyen esetben egy zárt rendszerrel állunk szemben.

² <http://www.businesswire.com/news/home/20110509005660/en/Affymetrix-Achieves-FDA-Clearance-Microarray-Based-Gene-Profiling>

A zárt rendszerben készült vizsgálatok eredményei megjelennek egy tudományos közleményben, ezután azonban az adatok további felhasználása csak ritka esetben fordul elő. Az általános megközelítés, hogy ezután egy másik csoport egy független vizsgálata is igazolja a megfelelő génjelöltet. A transzkriptomikus adatok mérésére alkalmas gén chipok felhasználásával lehetőség nyílik arra, hogy ne egy független beteganyagon kelljen egy újabb vizsgálatot végezni, hanem a már rendelkezésre álló egyéb kutatási projektek eredményeképpen elkészült adatokat használjuk fel erre a célra. Ez a minden szempontból hatékonyabb megközelítés nemcsak kutatási erőforrások és idő megtakarítását, de az egyes vizsgálatokban elemzett betegek összeadásával nagyobb teszhalmaz létrehozását és ezen keresztül magasabb statisztikai erő elérését is lehetővé teszi. A biomarkerek kutatói nézőpontból történő azonosítását az alábbi ábrán foglaltuk össze.



1. ábra A biomarkerek azonosításának folyamata kutatók szemszögéből. A hagyományos megközelítés során alkalmazott zárt rendszer alternatívája a több eredményt *in silico* integráló nyitott igazolások végzése, amihez az alapot a transzkriptom szintű mérések végzése adja.

2. KUTATÁSI CÉLOK

A sebészeti kezelés mellett a gyógyszeres kezelés súlya folyamatosan nő. A teljes test valamennyi sejtjét elérő („szisztémás”) kezelés előnye, hogy a rejtve maradó izolált tumorsejteket is el tudja érni. Ezen szisztémás kezelés további javítása két módon lehetséges. Az első lehetőség új, hatékonyabb gyógyszerek kifejlesztése, amelyek a jobb terápiás válaszadási arány mellett a túlélést, illetve az életminőséget jelentősen javítani tudják. Ezen opcióval foglalkozik valamennyi gyógyszergyár, illetve az akadémiai kutatás jelentős része is.

A második, általunk is vizsgált lehetőség a "személyre szabott terápia" alkalmazása, vagyis a rendelkezésre álló lehetőségek közül az adott beteg számára legmegfelelőbb gyógyszer vagy kezelési protokoll kiválasztása. Az új gyógyszer fejlesztésével szemben ezen megközelítés előnye, hogy a sokkal alacsonyabb költségigény mellett az egyén szintjén a legjelentősebb javulást (új gyógyszerek alkalmazása esetén is) csak a személyre szabott kezeléssel lehet elérni.

Vizsgálataink során onkológiai biomarkerek kutatásával foglalkoztunk. Ezen belül három fő célt tűztünk ki, úgymint:

1. **a legjobb biomarker-jelöltek azonosítása transzkriptom szintű adatok alkalmazásával sejtkultúra alapú kísérletekben**
2. **korábban azonosított biomarkerek igazolása transzkriptomikus adatok keresztelemzésével és klinikai mintákon való teszteléssel**
3. **bioinformatikai eszközök fejlesztése, amelyek a transzkriptomikus adatok alkalmazását lehetővé teszik további biomarkerek vizsgálata során is.**

Ezen három fő célkitűzés egyes kérdéseit azonban árnyalja, hogy egyrészt biomarkereket több tumortípus és több gyógyszer esetén vizsgáltuk, illetve a transzkriptomikus technológiák is különbözőek voltak az egyes vizsgálatokban. Az alábbiakban az egyes pontokhoz tartozó konkrét kérdéseket külön is részletezzük.

2.1. A legjobb biomarker-jelöltek azonosítása genom-szintű adatok alkalmazásával *in vitro* kísérletekben

2.1.1. Kemoterápiás szerekkel szembeni rezisztencia-profil meghatározása sejtvonalak vizsgálatával

Ezen első, transzkriptom szintű adatot is alkalmazó vizsgálat során a célunk az volt, hogy az egyes tumortípusoktól függetlenül tudjunk kemorezisztencia-biomarkereket azonosítani. Figyelembe véve, hogy az elképzelések alapján a tumorok általában hasonló rezisztencia-mechanizmusokat alakíthatnak ki, ezért arra törekedtünk, hogy nagyszámú sejtvonal egyidejű vizsgálatával olyan markereket azonosítsunk, amelyek a rosszindulatú daganatos betegség típusától függetlenek. **A kutatás során célunk volt, hogy összesen 30 különböző sejtvonal génexpressziós mintázatának felhasználásával 11 rákellenes gyógyszerrel szembeni rezisztencia-géneket meghatározzuk.**

2.1.2 A kemorezisztencia génjei melanoma malignumban

Rendelkezésünkre állt egy szülői melanoma sejtvonal (MeWo), valamint ennek négy különböző gyógyszerrel szemben rezisztenssé tett leányvonala. **A vizsgálatok során célunk volt a szülői- és leányvonalak génexpressziós mintázatát meghatározni, hogy ezáltal olyan géneket tudjunk azonosítani, amelyek az egyes gyógyszerekkel szembeni rezisztenciával kapcsolatba hozhatóak.**

2.1.3. Várható túlélés előrejelzése sejtkultúra modellben azonosított gének felhasználásával

Kutatásaink során célunk volt **sejtkultúrák modell rendszerben kemoterápiával szemben ellenálló és érzékeny sejtvonalak génexpressziós mintázatának összehasonlítása, hogy egy olyan mintázatot azonosítsunk, amellyel klinikai betegekben is tudunk előrejelzést adni a várható terápiás választ illetően.** Összesen négy különböző tumortípust vizsgáltunk egyidejűleg, hogy az egyes szövettípusoktól függetlenül csak a kemoterápiás válasszal kapcsolatba hozható géneket tudjuk azonosítani. Végezetül az adott gyógyszer (doxorubicin) monoterápiájával kezelt betegek elemzésével teszteltük modell rendszerünk hatásosságát.

2.1.4. Transzkriptomikus adatokkal meghatározott gének oki szerepének igazolása kísérletes körülmények között

A kutatás során elsődleges célunk volt **gén chip technológia alkalmazásával a doxorubicin-rezisztenciával kapcsolatba hozható proteaszóma-alegységek azonosítása**. A projekt második fázisában ezen alegységek rezisztenciában betöltött szerepét kísérletes módon akartuk igazolni, amelynek során **a PSMB7 gén elcsendesítésével egyidejűleg alkalmazott gyógyszeres kezelés mellett mértük az életben maradt sejtek arányát**. Végző célunk volt, hogy a fentebb már leírt kutatási folyamat egy kiterjesztett verziójával a rezisztencia egy olyan oki génjét tudjuk azonosítani, amely akár gyógyszerekkel is befolyásolható lehet, és ezáltal egy kombinált kemoterápiás kezelés alapját adhatja.

2.1.5. Parallel evolúciós modell alkalmazása a párhuzamosan létrejövő kemorezisztencia modellezésére

A fentebbi kérdések alapfeltevése az volt, hogy azok a mechanizmusok, amelyek klinikailag valóban jelentőséggel bírnak, több sejtvonalban is hasonló módon aktiválódhatnak. Felvetődik ezek után további kérdésként, hogy ha a korábbi kisebb vizsgálatok helyett **relatíve nagyszámú párhuzamos sejtvonalat fejlesztünk ki ugyanabból a genetikailag azonos ősből, akkor igazolni tudjuk-e a rezisztencia legmegbízhatóbb markereit?**

2.1.6. Célzott terápiás szerekkel szembeni rezisztencia és érzékenység vizsgálata

Az 1.-5. cél kutatása során hagyományos citotoxikus terápiákkal szembeni rezisztencia markereit kerestük. Kérdés, hogy **az egyes jelátviteli utak modulálásán keresztül ható célzott terápiás szerekkel szembeni rezisztencia markereit is** tudjuk-e azonosítani nagyszámú sejtvonal felhasználásával elvégzett sejtkultúrás mérések alapján. A legjobb biomarker-jelöltek azonosítása mellett célunk volt ezen **gének igazolása független klinikai mintákon is**.

2.2. Korábban azonosított biomarkerek igazolása transzkriptom szintű adatok keresztelemzésével és klinikai mintákon való teszteléssel

2.2.1. Korábbi közleményekben a melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása

Kutatásaink alaphipotézise alapján a melanoma pathogenezisével összefüggő gének közül a **legmegbízhatóbb géneket akkor tudjuk azonosítani, ha több vizsgálat**

eredményeit is felhasználjuk. Ezért a vizsgálat során célunk volt az **elérhető transzkriptomikus adatok alapján azonosított biomarkerek összehasonlítása és a hozzájuk rendelhető gén-ontológiai kategóriák meghatározása egy keresztelemzés során.** A kutatás további feltételezése volt, hogy nem szükséges az alapadatokat minden esetben ismernünk, elég a kutatások végeredményét elemezni.

2.2.2. Korábbi közleményekben emlőrák prognózisával kapcsolatba hozott génlisták keresztelemzése nyers adatok újrafeldolgozása után

A korábban közölt prognosztikai szereppel is bíró génlisták független klinikai anyagon való tesztelése nem megoldott. A kutatás során célunk volt **a transzkriptomikus adatok összerendezése egy adatbázisba, majd ezen adatbázis felhasználásával prognosztikus génlisták független tesztelésének elvégzése.** Végezetül meg akartuk vizsgálni, hogy az adatbázis felhasználásával lehetséges-e a korábbiaknál hatékonyabb előrejelző biomarkereket azonosítani?

2.2.3. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének azonosítása

A kutatás során a korábbi megközelítések helyett egy fordított kutatási tervet készítettünk: **először nagy in silico adatbázist építettünk, amelyen a szövettani altípusokkal és a túléléssel összefüggő géneket azonosítottuk.** Eközben klinikai mintagyűjtést folytattunk a Semmelweis Egyetemen és az Országos Onkológiai Intézetben. A **klinikai mintákon a gén chipek segítségével azonosított géneket egy független technológia (RT-PCR) alkalmazásával akartuk igazolni.**

2.3. Bioinformatikai eszközök fejlesztése, amelyek a transzkriptomikus adatok alkalmazását lehetővé teszik további biomarkerek vizsgálata során

2.3.1. Microarray adatok előfeldolgozása

Ezen módszertani vizsgálat során célunk volt az általunk két **különböző platformon (gén chip és RT-PCR) meghatározott minták felhasználásával az előfeldolgozási algoritmusok egymáshoz viszonyított teljesítményét kiszámolni.** A projekt végső célja az volt, hogy a további kísérletes és keresztelemző kutatások során ezen általunk igazolt legjobb algoritmust tudjuk a számítások során felhasználni.

2.3.2. Online rendszer fejlesztése az emlőrák prognózisát előrejelző génexpresszió alapú biomarkerek tesztelésére

A kutatás során célunk volt **a rendelkezésünkre álló gén chip adatokra építve egy olyan online elérhető rendszert kifejleszteni, ahol egy tetszőlegesen kiválasztott gén expressziójának az egyes betegekben mért szintje és a betegek várható kiújulás mentes és teljes túlélése közötti kapcsolatot lehet vizsgálni.** A rendszerbe olyan sok beteget akartunk integrálni, amennyi lehetővé teszi, hogy egyes ritkább alcsoportokban is megbízható statisztikai számítást lehessen végezni. A program létrehozása után a korábban leírt biomarker jelöltek közül a legígéretesebbek vizsgálatával demonstráltuk a software alkalmazhatóságát.

2.3.3. A világhálón keresztül elérhető keresztelemző rendszer továbbfejlesztése többgénosztályozók és optimalizált vágópont számítására

Az előző pontban említett Kaplan-Meier rajzoló programot **továbbfejlesztettük, hogy több gén egyidejű mérésére alkalmas legyen és lehessen a vágópontot optimalizálni.** A fejlesztés során **petefészekrák mintákat használtunk fel,** és a fentebb leírt rendszerrel párhuzamosan létrehoztunk egy második elemző oldalt.

2.3.4. A világhálón keresztül elérhető diagnosztikai rendszer fejlesztése

A kutatás célja egy, **a világhálón keresztül elérhető diagnosztikai rendszer** fejlesztése volt, amely **a betegből eltávolított tumorszövet feldolgozásával készült gén chipek segítségével direkt betegdiagnózis készítésére alkalmas.** A rendszerbe elsőként a három legelterjedtebb vizsgálatot (ösztrogén receptor, HER2 státusz, 21 génes „recurrence score”) terveztük beépíteni, majd nagy mintaszámon akartuk a hatásosságot ellenőrizni. Az utolsó feladat a rendelkezésre álló nagy adatbázis alapján egy, a korábbiaknál hatásosabb saját algoritmus fejlesztése volt.

3. MÓDSZEREK

3.1. Kemoterápiás szerekkel szembeni rezisztencia-profil meghatározása

A rezisztencia-index meghatározása sejtvonalakban

Vizsgálatok során 30 rosszindulatú daganatos sejtvonalat vizsgáltunk (a konkrét sejtvonalakat az eredmények fejezetben listáztam). A sejteket Leibovitz L-15-ös tápoldatban (Bio Whittaker, Walkersville, MD, USA) tartottuk, amelyet kiegészítettünk 10%-nyi borjúsérummal (Gibco BRL, Grand Island NY, USA), 1 mM L-glutaminnal, 80 IE/l inzulinnal, 2,5 mg/l transferrinnel, 1 g/l glukózzal, 1,1 g/l NaHCO₃-al és 1% minimális esszenciális vitaminnal. A sejteket 5% CO₂-t tartalmazó, nedvesített, 37 Celsius fokos inkubátorban tartottuk. A kísérletek előtt a Mycoplasma fertőzöttséget a VenorMp kit segítségével ellenőriztük a gyártó útmutatója alapján (Minerva Biolabs GmbH, Berlin, Németország).

A sejtvonalakban összesen 11 kemoterápiás szerrel szemben mértük meg a rezisztenciát. A gyógyszereket a cyclophosphamide-n kívül (amelyet aktivált formájában használtunk) a kereskedelmi forgalomban is alkalmazott kisserelésben vettük meg. Valamennyi gyógyszerrel szemben a rezisztenciát három különböző koncentrációban mértük le (minden mérést háromszor ismételtünk), amelyek közül a C2 koncentráció a gyógyszer klinikailag is alkalmazott koncentrációjával volt egyenértékű, a C1 ennek tizede, a C3 koncentráció pedig tízszerese volt.

Az egyes koncentrációknál a relatív rezisztenciát az SRB (szulforhodamin B) teszt segítségével határoztuk meg. A mérés első napján egy 96 lyukú lemez mindegyik lyukába 500 sejtet pipettáztunk ki. Kétnapos inkubáció után a mérés előtti („pre-”) kontroll sejteket megfestettük, és a vizsgált gyógyszer mindhárom koncentrációjával elindítottuk három lyukban a sejtvonal kezelését. Minden koncentrációt három párhuzamos mérésben indítottunk el. Négy nap után a kezelést megszakítottuk, a tápoldatot lecseréltük 10%-os triklórecetsavra és 1 órán át állni hagytuk. Ezután a lemezt ötször lemostuk desztillált vízzel, majd mindegyik lyukhoz hozzáadtunk 100µl SRB oldatot (Sigma, St Louis, MO, USA). Tíz perces szobahőmérsékletű inkubálás után a lemezt ötször lemostuk, hogy a nem kötött festéket eltávolítsuk. Végül a fehérjéhez kötött festéket 10mM Tris-HCL oldatban (pH=8.0) feloldottuk, és az egyes lyukakban lemértük az abszorbanciát 562 nm-en egy EL 340

Microplate Kinetics Reader ELISA-olvasóban (Bio-Tek Instruments, Winooski, VT, USA). A teljes sorozatot háromszor megismételtük, és az egyes mérési pontokat a kilenc mérési eredmény (három párhuzamos mérés és háromszoros ismételtség) átlagaként határoztuk meg.

Ezután a „rezisztencia-indexet” egy képlet alkalmazásával számoltuk, amely:

$$RI = (n_{\text{post}} / n_{\text{pre}}) \times [(n_2 - n_{\text{pre}}) / (n_{\text{post}} - n_{\text{pre}}) \times 100],$$

ahol n_{pre} a prekontrollban mért elnyelés, n_{post} a kontrollban (azonos idejű inkubáció gyógyszeres kezelés nélkül) mért elnyelés, az n_2 pedig az adott koncentrációval kezelt sejtekben mért elnyelés. Két gyógyszer (mitoxantron és topotecan) esetén a klinikailag alkalmazott koncentráció mellett nem volt elegendő rezisztens és érzékeny sejtvonal a statisztika megbízható elvégzéséhez, ezért ezekben a C3-as koncentrációt alkalmaztuk topotecan és a C1-es koncentrációt mitoxantron esetében.

Csak azokat a sejtvonalakat használtuk fel, ahol az n_{post} nagyobb volt, mint az n_{pre} és az ismételések közötti szórás 15%-nál kisebb volt. Az RI alapján sorba rendeztük a sejtvonalakat, az alsó harmadot érzékenyként, a felső harmadot rezisztensként határoztuk meg (a maradék sejtvonalat nem vettük be a statisztikai vizsgálatokba).

A logaritmusos növekedés fázisában összesen 10^7 sejtől izoláltunk RNS-t Qiagen RNeasy kit segítségével, a gyártó útmutatása alapján (Qiagen GmbH, Hilden, Németország). Az izolált RNS minőségét LabChip Bioanalyser-en (Agilent Technologies, Santa Clara, CA, USA) ellenőriztük. Az izolált mintákat -80°C fokon tartottuk hibridizációig.

Génexpresszió mérés

A génexpresszió transzkriptom szintű meghatározásához a mintákat Affymetrix HGU133 microarray chipekre (Affymetrix, Santa Clara, CA, USA) hibridizáltuk a Charité Berlin Funkcionális Genomikai Központjában. A feldolgozás az EukGE-WS2-es protokoll alapján, a leolvasás a GeneChip rendszer konfokális leolvasójával történt (Hewlett-Packard, Santa Clara, CA, USA).

A gén chipek felhasználásával kapott eredményeket TaqMan kvantitatív géneszpressziós mérés segítségével igazoltuk 46 kiválasztott gén esetében mind a 30 sejtvonalban. A méréshez az Applied Biosystems 7900HT Micro Fluidic Card rendszert, a leolvasást pedig az Abi Prism 7900HT Sequence Detection System felhasználásával végeztük

(Applied Biosystems, CA, USA). A TaqMan eredmények kiértékeléséhez az SDS 2.2 programot használtuk. A meghatározott delta Ct értékeket (amelyek a riboszomális 18S RNS kifejeződéséhez normalizált értékeket tartalmazzák) a sejtvonalakban meghatározott rezisztencia index alapján csoportosítottuk, és a rezisztens és érzékeny csoportokat t-próbával hasonlítottuk össze.

Gén chip adatok statisztikai kiértékelése

A Tumor Analysis Best Practices Working Group iránymutatási alapján ¹⁶² végeztük a microarray chipek minőségellenőrzését. Csak 25% feletti szignifikáns jel/próba arány és az RNS lebomlás mértékét mutató GAPDH 3'-5' arány három alatti értéke esetén használtuk fel az eredményeket. A fenti ajánlások alapján három normalizálási algoritmust alkalmaztunk, ezek a VSN (varianca-stabilizáló normalizálás) ¹⁶³, a MAS 5.0 (www.affymetrix.com) és az RMA ¹⁶⁴ voltak. A statisztikai kiértékelés valamennyi részlépése elkészült mindhárom normalizálási algoritmussal előkészített mintákon, hogy az eredmények reprodukálhatóságát ellenőrizni tudjuk.

A rezisztenciára vonatkozó expresszió alapú jellemzők kiválasztásához először mind a 30 expressziós mérést két csoportra osztottuk a sejtvonalak rezisztencia-profilja alapján. A rezisztens és érzékeny csoportba be nem sorolt sejtvonalakat kizártuk az elemzésből. Ezen csoportosítás alapján tehát minden gyógyszer esetében egy külön adatbázis jött létre, ezeket ezután egymástól függetlenül elemeztük tovább. A mintákat véletlenszerűen szétosztottuk tanuló- és teszhalmazra, majd a tanulóhalmazban a jellemzők kiválasztására két fő módszert alkalmaztunk, a PAM-ot („prediction analysis of microarrays”, ¹⁶⁵) és a SAM-ot („significance analysis of microarrays”, ¹⁶⁶). Megjegyzem, hogy a legjobban reprodukálható génlistát a PAM elemzéssel tudtuk azonosítani. A teljes mérési számítást 100-szor ismételtük a tanuló- és teszhalmaz folyamatos változtatásával. *(A normalizálási algoritmusokkal és az egyes gének szintjén mért génexpressziós értékek közötti korrelációval a 3.10.-es fejezetben részletesebben is fogunk foglalkozni.)*

3.2 A kemorezisztencia génjei melanoma malignumban

Sejtvonalak

Az eredetileg Fogh és munkatársai által létrehozott ¹⁶⁷ MeWo, valamint ennek négy leány-sejtvonalát használtuk, amelyek cisplatinnal, etoposiddal, vindesinnel és fotemustinnal

szemben voltak rezisztensek ¹⁶⁸. Az egyes szerekkel szembeni rezisztencia mértéke 3,8 (MeWo-CIS); 4,8 (MeWo-VIN); 32 (MeWo-ETO); és 4,8-szoros (MeWo-FOTE) volt, amelyet a 1 µg/ml cisplatin, 5 ng/ml vindesin, 1 µg/ml etoposid és 40 µg/ml fotemustin koncentráció mellett érték el. A sejtekből az RNS-t TRIzol felhasználásával izoláltuk, majd Invitrogen Fast Track 2.0 kit segítségével alakítottuk Poly(A) RNS-é. Az RNS minőségét Agilent Bioanalyser segítségével ellenőriztük.

Gén chip mérések

A génexpressziós mintázatok meghatározásához a Stanford-i egyetemen kifejlesztett, 43000 próba egyidejű meghatározására alkalmas gén chipet használtuk a korábban leírt protokollok szerint ^{169, 170}. Az elkészült chipet leolvasására GenePix 4000-es microarray leolvasót (Axon Instruments, Kalifornia, USA) használtunk, a nyers adatokat a Stanford Microarray Adatbázisban (<http://smd.stanford.edu/>) tároltuk.

Statisztikai számítások

Az előző fejezetben leírt, egyszínű gén chipet mérésére kifejlesztett kiértékelési módszerek nem használhatóak ezen chipet esetén. Itt két minta hibridizációja történik meg két különböző színnel (Cy5 és Cy3), amelyeknek az arányát adjuk meg az egyes próbákra, majd a normalizálást chipenként külön végeztük úgy, hogy a $\log_2(\text{Cy5/Cy3})$ arány az összes, a chipen elhelyezett próbára számolva 0,0 legyen. Legalább kétszeres expressziós különbséget kellett ahhoz elérnie egy génnek a gyógyszer-rezisztens és gyógyszer-érzékeny minták között, hogy az eltérést szignifikánsnak fogadjuk el. Fürtelemzést a Genesis program felhasználásával végeztünk ¹⁷¹. A számítások során Euklideszi és Pearson távolságot, illetve teljes és átlagos kapcsoltságot számoltunk.

RT-PCR

A különböző expressziót a legfontosabb gének egy kiválasztott csoportjában valósidejű RT-PCR felhasználásával igazoltuk. A méréseket Roche Light-Cycler készüléken végeztük a korábban leírt protokollok alapján ¹⁷².

3.3. Túlélés előrejelzése sejtkultúra-modellben azonosított gének felhasználásával

Kemorezisztens sejtvonalak

A sejtkultúra alapját olyan, kutatócsoportunk tagjai által korábban létrehozott leány-sejtvonalak adták, amelyek egy-egy kiválasztott gyógyszerrel szembeni derivátumait és a szülői sejtvonal kombinációját tartalmazták. Ezek között a vonalak között volt az EPG85-257P gyomorrák sejtvonal és leányvonalai ^{173, 174}, az EPP85-181P hasnyálmirigyrák vonal ¹⁷⁵, valamint a HT29-es vastagbélrák vonal ¹⁷⁶ és az MCF7 ¹⁷⁷ és MDA-MB-231¹⁷⁸ emlőrák vonalak illetve derivátumaik. A rezisztens sejtvonalakat a fejlesztés során az egyes gyógyszerek folyamatosan növekvő koncentrációjával kezeltük a rezisztencia aktuális mértékének monitorozása mellett.

A vizsgálat során három gyógyszerrel szembeni rezisztenciát tanulmányoztunk, amelyek a daunorubicin, a doxorubicin és a mitoxantron voltak. A daunorubicinnel szembeni rezisztencia mértéke 18-1857-szeres között, a mitoxantronnal szembeni rezisztencia 8-457-szeres között változott, doxorubicinnal szemben csak az MCF7 sejtvonalnak volt rezisztens leányvonala, amelyik 192-szeres rezisztenciát mutatott fel. A sejtvonalak tápoldata folyamatosan tartalmazta az adott gyógyszert, hogy a rezisztencia ne szűnjön meg. Az itt alkalmazott koncentrációk daunorubicin esetén 0,125-2,5 µg/ml között, mitoxantron esetén 0,02-0,2 µg között változtak, doxorubicin esetén pedig 0,05 µg/ml volt. A szülői sejtvonalak eközben nem részesültek gyógyszeres kezelésben.

Gén chip mérések

A sejtvonalakból TRIzol protokoll alkalmazásával izoláltuk az RNS-t, majd Invitrogen Fast Track mRNS izoláló kit segítségével készítettünk belőle Poly(A) RNS-t. Az RNS minőségét a fentebb már leírt módon, Agilent 2100-as Bioanalyser segítségével ellenőriztük. A microarray mérések során szintén a Stanford-i egyetemen készített gén chipet használtuk, amelyeken 30 ezer gén mérésére alkalmas 43009 cDNS klónt helyeztek el. Maga a hibridizáció a fentebb már idézett módon történt ^{169, 170}.

A statisztikai kiértékelés során célunk nem csak a különbözőképpen kifejezett gének azonosítása volt, de ezeket a géneket fel is akartuk használni arra, hogy független minták

várható rezisztenciáját előre jelezzük. Ehhez a Prediction Analysis of Microarrays (PAM) csomagot használtuk ¹⁶⁵. Az elmúlt 10 év alatt ez az egyik legelterjedtebb elemző algoritmussá vált, mára több mint 1600 alkalommal idézték az eredeti közleményt. A PAM előnye, hogy két független módon is használható: egyrészt létezik egy Microsoft Excel kiegészítő, valamint az R statisztikai környezeten belül is van egy működő implementációja. Ezen kutatási projekt során az Excel kiegészítőt használtuk, mivel itt az adatok mozgatása egyszerűbb volt. Statisztikai szempontból a PAM egy úgynevezett finom küszöbölés („soft thresholding”) során létrehoz egy tömörített központi vektort („shrunk centroid”), amely segítségével azon géneket tudja megjelölni, amelyeknek legnagyobb az előrejelző potenciálja.

További számítások között a korábban már említett Genesis program felhasználásával végzett fürtelemzés ¹⁷¹, illetve Winstat for Excel (R.Fitch Software, Staufen, Németország) segítségével készített Kaplan-Meier elemzés szerepeltek. A microarray adatokat a Stanford Microarray Adatbázisban nyilvánosan letölthetővé tettük.

RT-PCR mérések

A vizsgálat során egy kiválasztott gén expresszióját vizsgáltuk meg két független próbával Applied Biosystems 7900HT rendszer segítségével. Az ABCB1 próbákat (Hs00184491 és Hs00184500) a GAPDH (Hs99999905) expressziójához normalizáltuk. Emellett két további gén, a caspase 1 és a topoizomeráz 2a (TOP2a) kifejeződését LightCycler-en valósidejű RT-PCR segítségével igazoltuk - ezeket a géneket szintén a GAPDH expressziójához normalizáltuk. Az RT-PCR mérések során az EPP181 sejtvonalból vett cDNS hígítási sorát használtuk kalibrálásra mindegyik mérés esetén.

3.4. A PSMB7 gén oki szerepének igazolása *in vitro* és *in silico*

Sejtkultúra

Az előző fejezetben tárgyalt négy sejtvonalpárt használtuk fel, az EPG85-257P, az EPP85-181P, HT29 és az MCF-7 szülői vonalakat, valamint doxorubicin/daunorubicin rezisztens leányvonalait. A vonalakat fentebb leírt körülmények között tartottuk, a sejtproliferációt a gyógyszerek klinikai dózisának 0.0001x-100x-szorosa közötti tartományban határoztuk meg.

A sejtvonalak érzékenységét Roche MTT (3-[4, 5-dimethylthiazolyl-2]-2, 5-diphenyltetrazolium bromid) sejtproliferációs kittel mértük. A leolvasáshoz egy BioTek PowerWave XSA spektrofotométert használtunk.

A doxorubicin korábban leírt jelentős autofluoreszcenciája ¹⁷⁹ miatt a mérések előtt egy üres kontroll mérést végeztünk, hogy ennek a mérések eredményére gyakorolt hatását meg tudjuk becsülni. Ennek során 96 lyukú lemezen érzékeny MCF-7 sejteket helyeztünk el kezeletlen és doxorubicinnal kezelt állapotban, valamint üres és doxorubicinnal kiegészített tápoldatot. Egynapos inkubáció után MTT tesztet végeztünk és a különbségeket t-próbával hasonlítottuk össze. A kezelt és kezeletlen sejtek között nem volt különbség. A doxorubicinnal kiegészített tápoldat szignifikánsan eltért a doxorubicinnal kezelt sejtektől ($p < 0,01$), azonban nem különbözött az üres tápoldattól ($p = 0,065$).

Gén chip mérések

Valamennyi sejtvonalból izoláltunk RNS-t Qiagen RNeasy Mini kit felhasználásával. Az RNS minőségét a fent leírt módszerek segítségével ellenőriztük, majd az izolált RNS-t a standard (EukGE-WS2) Affymetrix protokollok alkalmazása mellett HGU95 és HGU133A chipekre hibridizáltuk. A leolvasást Hewlett-Packard GeneChip konfokális leolvasóval végeztük el. Valamennyi sejtvonalból három hibridizáció készült, amelyekkel az érzékeny, a rezisztens és a kezelt vonalak expressziós mintázatát határoztuk meg. A gén chip adatokat a GEO adatbázisban a GSE3926-os azonosító alatt tettük mások számára elérhetővé.

Az eredmények minőségellenőrzése során a chipék felületét *in silico* rekonstruáltuk és a karcolási, illetve cseppentési foltok hiányát ellenőriztük. Valamennyi gén chip 25% feletti sikeres gén-mérést („percentage of present calls”) mutatott, és a mintákban az RNS nem volt lebomlott. A normalizálás során az RMA algoritmust ¹⁶⁴ használtuk. Az adatok statisztikai feldolgozása az R statisztikai környezetben, Bioconductor csomagokkal történt.

A jellemző gének kiválasztásához a normalizált adatokból három csoportot képeztünk, amelyek az érzékeny szülői, a rezisztens leányvonalakat és a gyógyszerrel kezelt szülővonalakat tartalmazták. A statisztikai elemzéshez a PAM algoritmust ¹⁶⁵ használtuk, amellyel a szülői és a gyógyszerrel kezelt, illetve a szülői és a rezisztens csoportokat külön hasonlítottuk össze. Az elemzés során mindig a 100 legerősebb génnél húztuk meg a

küszöböt. A gének eltérő kifejeződésének ábrázolásához fürtelemzést végeztünk a Genesis programmal ¹⁷¹.

RNS alapú génelcsendesítés

Az RNS alapú génelcsendesítés („RNA interference”) elvégzéséhez szükséges rövid oligonuklotid szakaszokat az Ambion siRNA Target Finder programjával és a Dharmacon siDesign Center programjával terveztük. Magát a csendesítést végző oligonukleotidokat DNS oligonukleotidokból készítettük az Applied Biosystems Silencer siRNA Construction Kit-jének a felhasználásával. Az siRNS hatásosságát RT-PCR alkalmazásával igazoltuk, a PCR mérésekhez a primereket a Primer3 programmal terveztük. A kiválasztott PSMB7 génre összesen három siRNS molekulát készítettünk és ellenőriztünk.

Az siRNS kezeléshez a transzfekciót Applied Biosystems SiPort NeoFX transzfekciós reagens felhasználásával, 230 ezer sejten végeztük, 6 lyukú lemezen, 24 órás letapadási idő után. A gén hatását a kemorezisztenciára siRNS és gyógyszeres kezelés kombinációjával vizsgáltuk. Ennek alapját az adta, hogy a mérések során a PSMB7 gén magasabb kifejeződését találtuk a rezisztens vonalakban a szülői vonalakhoz viszonyítva. A kísérlet során a génelcsendesítés után 24 órával a tápoldatot kicseréltük 0,2 µg/ml doxorubicint tartalmazó oldatra, majd további 72 óra után az élő sejtek számát Casy sejtszámlálóval határoztuk meg. Valamennyi mérést háromszor ismételtük, illetve az egész kísérletet is háromszoros ismétléssel végeztük el. Végezetül az eredményeket normalizáltuk a kezeletlen sejtek számához, majd az siRNS-el és gyógyszerrel is kezelt, negatív siRNS-el és gyógyszerrel kezelt, valamint siRNS-el igen, de gyógyszerrel nem kezelt sejtekre vonatkozó mérések eredményeit t-próbával hasonlítottuk össze.

Klinikai tesztelés

A klinikai kipróbáláshoz felhasznált adatbázis létrehozását a későbbiekben részletesen is leírjuk ⁷⁷. Ebben a vizsgálatban a GSE12276, GSE16391, GSE12093, GSE11121, GSE9195, GSE7390, GSE6532, GSE5327, GSE4922, GSE3494, GSE2990, GSE2034 és a GSE1456-os közleményeket használtuk fel.

A statisztikai kiértékelés során csak a proteaszóma alegységeket vizsgáltuk (n=64), amelyekkel túlélési elemzést végeztünk a BRB Arraytools programcsomag felhasználásával,

illetve Kaplan-Meier grafikonokat készítettünk. A statisztikai számítások során a még elfogadott hibás találati arányt (FDR - „false discovery rate”) 0,05-nél határoztuk meg.

3.5. Parallel evolúciós modell a párhuzamosan létrejövő rezisztencia modellezésére

Sejtvonalak fejlesztése

Az MDA-MB-231-es és az MCF-7 sejtvonalakat az ATCC-től szereztük be és a szállító által megadott körülmények között tartottuk. A rezisztens leányvonalak fejlesztésének kezdetén a szülői vonalakból 20-20 alvonalat különítettünk el, amelyeket a fejlesztés teljes ideje alatt elkülönítve kezeltünk (tehát összesen 40 sejtvonal). A rezisztenciát két gyógyszerrel, a doxorubicinnal és a paclitaxellel szemben hoztuk létre a sejtvonalak hetente történő kezelésével. A gyógyszeres kezeléssel párhuzamosan kontrollként a szülői vonalakat is fenntartottuk a fejlesztés teljes ideje alatt.

A kezelés során hetente meghatároztuk, hogy a vonalak milyen mértékben fedik az edény felületét. Amennyiben ez nem érte el a felület felét, akkor a kezelést leállítottuk, amennyiben 50-70% között volt, akkor a kezelést folytattuk, és ha meghaladta a 70%-ot, akkor fagyasztást végeztünk a sejtvonalakból az adott koncentráció mellett. Három sikeres fagyasztás után a gyógyszerek koncentrációját növeltük. A növelés 10-szeres lépésekben történt, amíg el nem értük a klinikailag is alkalmazott dózis 10%-át, ezután a dózis 30%-ára, majd a 60%-ára emeltük a koncentrációt. Valamennyi koncentrációnövelés előtt MTT sejtproliferációs teszt felhasználásával ellenőriztük a sejtek vitalitását. Az MTT tesztet cisplatin és 5-fluorouracil kezelés mellett is meghatároztuk (illetve doxorubicinnal kezelt sejtek esetén is mértük a paclitaxel-rezisztenciát és fordítva is), hogy a rezisztencia során keletkező kereszt-rezisztenciát is dokumentálni tudjuk.

RT-PCR mérések

Irodalomkutatás alapján választottuk ki azokat a géneket, amelyek a rezisztencia létrejöttében szereppel bírhatnak, majd a gének kifejeződését Applied Biosystems Micro Fluidic Card System segítségével határoztuk meg. Az eredményeket három háztartási gén átlagához normalizáltuk, ezek a riboszómális 18S, a HPRT1 és az RPLP0 voltak. Az RT-PCR eredményeket SDS 2.0 programmal értékeltük ki, majd Spearman rang korrelációt

számoltunk normalizált génexpressziós értékek és a rezisztencia mértékének összehasonlítására. Az eredmények bemutatására fürtelemzést végeztünk a Genesis program felhasználásával.

Hagyományos citogenetika és a kromoszómális instabilitás mérése

A sejtvonalak létrehozása végén a vonalakat a 80%-os konfluencia eléréséig hagytuk nőni, majd a sejteket standard protokollok alapján megfestettük. Minden mintából 10 metafázisban levő sejtet értékeltünk ki, a kariotípusokat a Nemzetközi Humán Citogenetika Rendszer 2009-es kiadása alapján azonosítottuk. A citogenetikai elemzést az MDA-MB-231-es vonal valamennyi leányvonalán elvégeztük.

A 3-as, 17-es és 21-es kromoszómák variációinak 10 metafázisban történő számításával számoltuk a sejtekre jellemző kromoszómális instabilitást (CIN) Shannon diverzitási index¹⁸⁰ felhasználásával:

$$H = -\sum_i p_i \ln(p_i)$$

ahol p_i egy adott kromoszóma (i) frekvenciája a sejtvonalakban, és H a kromoszómális heterogenitás számszerű jellemzésére alkalmas. Amennyiben a H index értéke 2,1 felett volt, akkor CIN pozitív, ha alatta, akkor CIN negatív sejtvonalakról beszélhetünk.

Áramlásos citometriai mérések

A korábbi irodalmi adatok a kemorezisztencia létrejötte mögött a PGP transzportrendszereket többször igazolták. Ezért szükségesnek tartottuk azt is meghatározni, hogy ez a transzportrendszer milyen mértékben változott meg az egyes sejtvonalakban, illetve hogy ez összefüggést mutat-e a rezisztencia kialakulásával. Ehhez a gyógyszeres kezelés után a PGP szubsztrát rhodamine 123-at használtuk. Összesen 10^4 sejtet vizsgáltunk áramlásos citometriával, a sejtszámot a rhodamin intenzitáshoz viszonyítva ábrázoltuk.

3.6. Célzott terápiás szerekkel szembeni rezisztencia

Gén chip kiértékelések

A kutatás alapját a GSK caArray adatbázisa adta, ahol több mint 300 ismert sejtvonal transzkriptom szintű génexpresszióját határozták meg Affymetrix gén chipek felhasználásával (ftp://caftpdc.nci.nih.gov/pub/caARRAY/transcript_profiling). Valamennyi gén chip mérést háromszor ismételték. A nyers adatok letöltése után a mintákat újrnormalizáltuk, valamint a háromszoros ismétlés eredményeit próba szinten átlagoltuk. Az alacsony átlagos kifejeződést mutató próbákat kizártuk, majd az összes sejtvonalat szétszortottuk aszerint, hogy az alább ismertetendő osztályozás alapján az érzékeny vagy a rezisztens csoportba kerültek be.

Ezt mind az öt vizsgált gyógyszer esetében külön végeztük el, majd a két csoportot SAM¹⁶⁶ és „rank products”^{181, 182} algoritmusokkal hasonlítottuk össze. Bár a SAM a legelterjedtebb algoritmus gén chip adatok összehasonlítására, azonban egy korábbi összehasonlító elemzés¹⁸³ szerint a „rank products” megbízhatóbb eredményt ad alacsony esetszámnál - azért alkalmaztuk a kettős módszert, mivel az egyes gyógyszerekre alkotott sejtvonal-csoportok viszonylag alacsony esetszámot tartalmaztak. Az egyes génlisták esetében az érzékeny és rezisztens sejtvonalak elkülönítésének képességét „Prediction Analysis of Microarrays”¹⁶⁵ elemzéssel határoztuk meg. Csak azokat a géneket fogadtuk el szignifikánsként, amelyek legalább kétszeres expresszióbeli különbséget mutattak az érzékeny és a rezisztens sejtek között, és ahol az FDR 20%-nál kisebb volt.

Sejtkultúra

A fentebb elemzett gén chip adatok elkészítéséhez a sejteket az ATCC sejtvonal-bankból szereztük be, és az ATCC protokolljai szerint tartottuk fent. A sejtekből DNS-t Qiagen DNeasy Blood and Tissue kit, RNS-t pedig Qiashredder-el történő homogenizáció után Qiagen RNeasy kit felhasználásával izoláltuk. A nukleinsavak tisztaságát Nanodrop 1000 rendszerrel ellenőriztük, valamint Bioanalyzer-rel vizualizáltuk a riboszómális 18S és 28S RNS csíkok megfelelő elkülönülését. A kísérletek előtt a sejtvonalak eredetét külön is igazoltuk STR („short tandem repeat”) elemzéssel, mivel néhány vonal esetében elérte a 4 évet a beszerzés és a kísérlet befejezése közötti idő. Valamennyi sejtvonal szennyeződéstől mentes volt.

A sejtek rezisztenciáját öt kiválasztott szerrel szemben (erlotinib, gefitinib, sunitinib, sorafenib és lapatinib) MTT teszttel három koncentráció mellett határoztuk meg: a klinikailag alkalmazott, annak tizede, valamint tízszerese mellett. A rezisztencia meghatározásához a mérés kezdetén elvégzett „prekontroll” mérés értékével csökkentett értéket hasonlítottuk a gyógyszeres kezelés nélküli „postkontroll” mérés „prekontroll” mérés értékével csökkentett referenciájához. A mérések sikerességének feltétele volt, hogy a „prekontroll”-nál nagyobb legyen a „postkontroll”, illetve hogy a háromszor megismételt mérések között az eltérés 15%-nál kisebb legyen. Végezetül a sejteket szétosztottuk három csoportba, mint „érzékeny”, „rezisztens” és „közepes érzékenyséű” sejtek, ahol a közepes csoport a medián +/-10%-os tartományba eső vonalak voltak.

A statisztikai elemzéssel azonosított legerősebb gének, valamint néhány, az irodalom alapján kiválasztott gén kifejeződését TaqMan RT-PCR felhasználásával igazoltuk. Az összesen 95 gén RT-PCR mérésének eredményét a riboszómális 18S kifejeződéséhez normalizáltuk. Az RT-PCR eredmények kiértékelése az SDS 2.0-ás programmal történt, majd a csoportok közötti különbség szignifikanciáját t-próbával számoltuk.

Gyógyszer	Bevezetve	Alk. konc.* (µg/ml)	Célfehérje	Tüdő	Emlő	Vastagbél	Veserák	GIST	Hasnyálmirigy	Májrák	Fej-nyaki r.
lapatinib (Tykerb)	2007	16.66	EGFR és HER2/neu (ErbB-2)	II	J	II	I	-	L	II	III
sunitinib (Sutent)	2006	0.67	PDGF-R és VEGF-R	III	III	III	J	J	J	III	II
sorafenib (Nexavar)	2005	5.32	Raf, c Kit, PDGF, VEGF receptor 2 & 3 kinázok	III	II	II	J	II	III	J	II
erlotinib (Tarceva)	2004	2.00	EGFR tirozin kináz	J	II	III	II	II	J	II	III
gefitinib (Iressa)	2003	3.32	EGFR (Her1 or ErbB-1) tirozin kináz	J	II	III	II	-	II	II	III

* az alkalmazott koncentráció a klinikai dózishoz igazodott

Forrás: www.clinicaltrials.gov

I: Fázis I

II: Fázis II

III: Fázis III

J: jóváhagyott

L: leállított

-: nincs vizsgálat

1. táblázat. A sejtvonalakon alkalmazott öt tirozin kináz gátlószer klinikai indikációi, valamint az alkalmazott terápiás koncentráció

Klinikai mintagyűjtés

A Semmelweis Egyetem Urológiai Klinikáján 2005 és 2010 között folytattunk veserák mintagyűjtést. Mivel a veserák kezelése során a hazai gyakorlat szerint áttét jelentkezése esetén sunitinib a választandó szer, ezért ez optimális betegcsoportot alkotott. A vizsgálatba csak azok a betegek kerültek be, akik valóban kaptak sunitinibet. A betegekből a mintagyűjtés a gyógyszeres kezelés előtt történt, ami a sejtkultúrák kísérletek során végzett megközelítéssel egyezett. A tumormintákból először TMA („szöveti micro array”) készült, amely minden mintát kétszer tartalmazott. Ezután immunhisztokémiával ellenőriztük négy gén, a CD9, az LGALS8, a RAB17 és az EpCAM festődését. A festődés intenzitása és a megfestett sejtek aránya alapján osztottuk a betegeket két csoportra, majd ezeket Kaplan-Meier túlélési görbével hasonlítottuk egymással össze.

3.7. A melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása

Adatgyűjtés

A releváns tanulmányok azonosításához a PubMed-ben a „melanoma”, „gén chip”, „génexpresszió” kulcsszavak felhasználásával végeztünk keresést. Az eredmények közül csak az emberben végzett vizsgálatokat vettük figyelembe. A keresés alapján összesen 149 közleményt találtunk, ezek mindegyikét áttanulmányoztuk. Kizártuk azon vizsgálatokat, ahol csak egy gént azonosítottak, illetve ahol állatmodelleken dolgoztak.

Statisztikai számítások

A géneket unigene azonosítójuk alapján tartottuk nyilván. Az egyes génlisták egymással való összekapcsolásához az Affymetrix cég honlapján (www.affymetrix.com) elérhető gén-azonosítási táblázatokat használtuk fel. A génlisták keresztelemzése során Microsoft Access adatbáziskezelő beépített eszközeit használtuk.

A gén-ontológiai azonosítás folyamán a felülreprezentált gén-ontológiai kategóriákat a Netaffx központ (<http://www.affymetrix.com/analysis/index.affx>) felhasználásával azonosítottuk.

3.8. Az emlőrák prognózisával kapcsolatba hozott génlisták kereszt-elemzése

Gén chip adatok és génlisták

Valamennyi közleményt a Pubmed felhasználásával azonosítottunk. A gén chip adatokat az „emlőrák” és az Affymetrix HGU133A platformra specifikus „gpl96” keresőszavak kombinációjával, a génlistákat az „emlőrák” és a „microarray” szavak kombinációjával. Kihagytuk azon vizsgálatokat, ahol kevesebb mint 20 beteget elemeztek, illetve a génlisták esetében még azokat a közleményeket is, ahol kevesebb mint öt gént azonosítottak. A génlisták génjeit a gén chip adatokkal az Affymetrix által közzétett táblázatok segítségével kapcsoltuk össze Microsoft Access adatbázis kezelőben.

Statisztikai módszerek

A letöltött nyers adatokat MAS5.0 algoritmussal normalizáltuk az R statisztikai környezetben. A fentebb már leírt PAM algoritmus egy újabb verziójával ¹⁸⁴ elemeztük a génlisták hatását. A számítások során a PAM elemzés küszöbét nullára állítottuk, ezáltal a génlisták valamennyi génje bekerült az elemzésbe.

Leíró statisztikát, túlélési görbéket és a klinikai változókat Winstat for Excel felhasználásával elemeztük.

Egy új osztályozót is fejlesztettünk, amihez a BRB Arraytools 3.6.0-beta_3-as verzióját használtuk (a programot Dr Richard Simon és Amy Peng Lam fejlesztették). A program ezen verziója egy új szolgáltatást vezetett be, amelynek segítségével az egyes géneket hozzá lehet adni kiválasztott klinikai változókhoz, majd ezen kombinált osztályozóval lehet a betegek túlélését megbecsülni.

3.9. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének igazolása klinikai mintákban

Célgének azonosítása in silico

A PubMed-ben, illetve a GEO-ban keresést végeztünk a „petefészek”, „rák”, „normális”, „GPL96” és a „GPL570” kulcsszavak felhasználásával. A nyers adatokat letöltöttük, és MAS5.0 algoritmusokkal normalizáltuk. A GPL570-es platform génjei közül csak azokat tartottuk meg, amelyek a GPL96-os platformon is azonos szekvenciával szerepeltek. Az egyesített adattáblán egy második normalizálást is végeztünk, aminek során az átlagos jelintenzitást 1000-re állítottuk be.

A génextpressziós adatokat BRB ArrayTools 3.7.0-ba olvastuk be, aminek során a 10-nél kisebb jelek értékét 10-re növeltük, illetve töröltük azokat a géneket, ahol a minták 20%-a nem érte el a legalább 50%-os különbséget a legkisebb és a legnagyobb jelintenzitású minta között. Amennyiben a hiányzó vagy kiszűrt minták aránya elérte az 50%-ot, akkor a gént teljesen kihagytuk az elemzésből. Összesen 20377 próba teljesítette ezeket a feltételeket, ezen gének felhasználásával hasonlítottuk össze az egyes szövettani altípusokat, valamint a normális és rosszindulatú daganatos szöveteket.

Klinikai mintagyűjtés

A Semmelweis Egyetem I. Nőgyógyászati Klinikáján és az Országos Onkológiai Intézetben klinikai mintagyűjtést végeztünk 2005 és 2008 között az Intézeti Etikai Bizottságok jóváhagyásával. A petefészekrák mintákat -80 Celsius fokon tartottuk RNS izolálásig, amelyet Qiagen RNeasy kittel végeztünk el. Az izolált RNS minőségét Nanodrop1000-es rendszer és Agilent Bioanalyzer felhasználásával ellenőriztük. Csak azon mintákat használtuk fel, ahol a riboszómális 18S és 28S csíkok tisztán elkülöníthetőek voltak.

RT-PCR mérések

TaqMan valós idejű RT-PCR méréseket végeztünk 40 kiválasztott gén kifejeződésének meghatározására ABI PRISM 7900HT Sequence Detection System segítségével. A gének között voltak a szövettani típussal és a túléléssel kapcsolatba hozott gének is. Csak olyan géneket választottunk ki, amelyeknek legalább az egyik csoporton belüli átlaga meghaladta az összes gén átlagát. Néhány további, az irodalom alapján azonosított gént (például ABC szállítófehérjék és háztartási gének) is mértünk.

Az RT-PCR méréseket az SDS 2.0 program használatával értékeltük ki. Ennek során a mért értékeket a riboszómális 18S és az RPLP0 kifejeződéséhez normalizáltuk. Ezután az egyes mintákat klinikai paramétereik alapján csoportokba rendeztük és a csoportokat Significance Analysis of Microarrays ¹⁶⁶ segítségével hasonlítottuk össze. A számítások során a statisztikai szignifikancia küszöbszintjét 10%-os hibás találati arány (FDR) mellett határoztuk meg. Multivariancia és Kaplan-Meier elemzést Winstat 2007 alkalmazásával végeztünk.

3.10. Microarray adatok előfeldolgozása

Kísérletes adatok

Két saját korábbi vizsgálatunk adatait használtuk fel, amelyek során gén chip és RT-PCR méréseket is végeztünk a gének egy csoportjára. A 3.1. fejezetben leírt 30 sejtvonalból lemért 75 gén adta az egyik adatbázist, a másik vizsgálatunkban¹⁸⁵ 36 colon biopsziából mértünk le 84 gént. Ezen második adatbázist a Semmelweis Egyetem 2. Belgyógyászati Klinikáján gyűjtött minták adták, amelyek között adenomás, colon tumoros, gyulladásos bélbetegségben szenvedő betegek és normál szövetminták is voltak.

A gén chip mérések mindkét vizsgálatban Affymetrix HGU133 sorozatú microarray-okkal, az RT-PCR mérések pedig 96 lyukú lemezen (Applied Biosystems Taq-Man rendszer) történtek. Az egyes gének azonosítására a gyártók által rendelkezésre bocsátott táblázatokat használtuk. Meg kell jegyezni, hogy a 96 lyukú lemezen 95+1 (egy kontroll gén) mérést végeztünk, azonban nem lehetett valamennyi gént pontosan megfeleltetni a gén chipen elhelyezett próbáknak. Ennek hátterében az áll, hogy az egyes génekről átírt exonok végül különböző „transzkript”-ek formájában jutnak ki a sejtmagból, amelyeken azonban különböző helyeken kötnek be a Taq-Man és a gén chip próbák. A kiejtett esetekben a két platform eltérő „transzkript”-készletet mért az adott génre. A nyers gén chip adatok a GEO-ban (<http://www.ncbi.nlm.nih.gov/gds>) elérhetőek a GSE4183 és a GSE11812-es azonosító alatt.

Statisztikai algoritmusok

A statisztikai számítások az R környezetben (<http://www.R-project.org>) történtek, Biocondutor csomagok felhasználásával (<http://www.bioconductor.org>). Valamennyi számítást log-transzformált értékekkel végeztünk el. Az RT-PCR eredményeknél a log₂-es C_T

értékeket használtuk, amelyek a mért gén referenciagénhez viszonyított megjelenésének értékeit takarták.

A kilenc vizsgált előfeldolgozási algoritmusok között szerepelt a DFW ¹⁸⁶, a FARMS ¹⁸⁷, a GCRMA ¹⁸⁸, a MAS5 ¹⁸⁹, a MBEI ¹⁹⁰, az RMA ¹⁹¹, a VSN ¹⁶³, és a PLIER (http://www.affymetrix.com/support/technical/technotes/plier_technote.pdf és ¹⁹²). Ezek közül kettő algoritmushoz (MBEI és PLIER) az eredeti szerzők nem készítettek programot, azonban R-ben elérhető volt a tőlük függetlenül elkészített algoritmus. Valamennyi algoritmust a szerzők által javasolt alapbeállításokkal használtuk, kivéve a PLIER esetében, ahol egy alternatív kombinációt (PLIER+16) is kipróbáltunk. Az alapbeállítások alkalmazásával elkerültük, hogy az eredmények túloptimalizálása egyes algoritmusok jobb teljesítése érdekében lehetővé váljon.

3.11. Bioinformatikai fejlesztés génexpresszió alapú biomarkerek tesztelésére

Adatbázis előkészítés

Az adatbázist a fentebb már leírt módszerrel ¹⁹³ hoztuk létre. A rendszer első nyilvános kibocsátásakor 1809 emlőrákos beteg elemzése volt beépítve, ez a betegszám már nem tartalmazott ismételten leközölt gén chipeket (n=384). A betegek medián kiújulás-mentes túlélése 6,43 év volt, 1231 betegről ismertük a hormonreceptor-státuszt (968 beteg volt pozitív). Az elemzés során a teljes túlélés (594 beteg) és a távoli áttét mentes túlélés (767 beteg) adatokat is felhasználtuk.

Jóllehet a gén chip mérések nem alkalmasak közvetlenül az adott génről készült mRNS-ek számának meghatározásához, azonban a biomarkerek esetében nem is ez a célunk, elegendő az egyes mintákban mért, egymáshoz viszonyított génexpresszió mértékének meghatározása. Ehhez a gén chip adatok normalizálása során a 4.10.-es fejezetben részletesen ismertetett munkánk alapján a MAS5 normalizálást használtuk. A választás során az adatbázis későbbi bővíthetősége volt a legfontosabb szempont.

A rendszer több Affymetrix platform egyesítését tartalmazta, mint a GPL96-os (HGU133A), a GPL570-es (HGU133 Plus 2.0) és később a GPL571-es (HGU133A 2.0) gén chippek. Ezen platformok tartalmaznak 22277 közös próbát, amelyekben az egyes gének

mérésére használt próbák szekvenciája egymással teljesen megegyezik és ezért a hibridizáció során azonos érzékenységet és dinamikus tartományt várhatunk. Azonban gyári beállításokkal lefutott MAS5 algoritmus az egyes platformok esetén más és más átlagos expresszióra normalizál, ezért a platformok közös génjeinek kiszűrése után egy második normalizálást is végeztünk, amelynek során az átlagos génexpressziós értéket minden minta esetén 1000-re állítottuk be. Ez a módszer a korábbi irodalmi adatok alapján a csoportos hibát („batch effect”) is csökkenti ¹⁹⁴.

A klinikai adatokat elsősorban az eredeti közleményekből vettük. Azonban egyes paramétereket (pld ÖR státusz) a gén chip adatok alapján is meg lehet határozni. Ehhez az ÖR státusz meghatározásához az ÖR gén 205225_at próbáját használtuk a korábbi irodalmi adatok alapján javasolt 500-as vágópont felhasználásával ¹⁹⁵. Ez a próba a rendelkezésre álló 9 ÖR próba közül a legmagasabb átlagos és medián kifejeződést mutatta, valamint legjobban korrelált az immunhisztokémia alapú ÖR státusz meghatározással. A fejlesztés során az ÖR státuszt mint önálló szűrőt is alkalmaztuk.

Az génexpressziós és klinikai adatokat egy MySQL adatbázisban egyesítettük. Az adatbázis további fejlesztése során a MySQL adatbázis használata azonban nehézkesnek bizonyult, ezért később átálltunk a PostgreSQL adatbázis használatára.

Világhálón keresztül elérhető rendszer fejlesztése

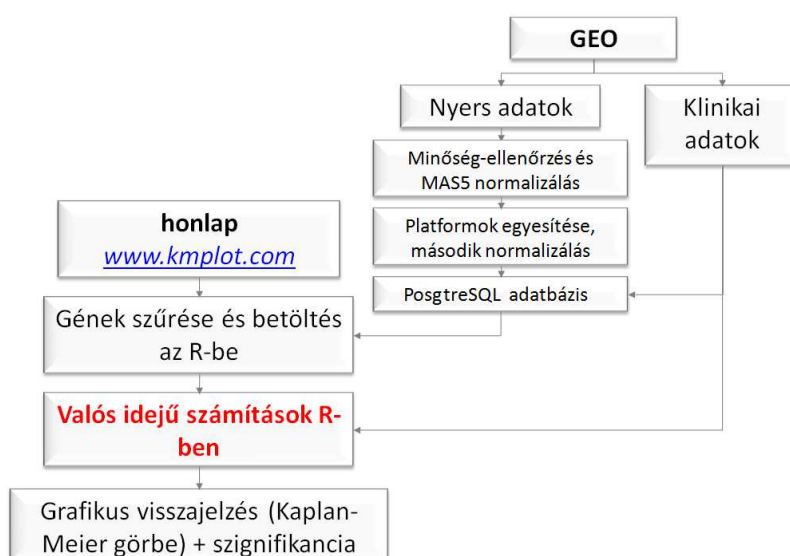
A rendszer három részből áll:

1. adatok PostgreSQL adatbázisban,
2. világhálón keresztül elérhető kezelőfelület és
3. az R statisztikai környezetet futtató virtuális kiszolgáló.

Az R környezetet futtató virtuális kiszolgálóra („R server”) azért volt szükség, mert az R-ben elérhető könyvtárak („libraries”) biztosították az eszközöket a statisztikai számítások elvégzéséhez. Azonban az R alapesetben önálló program, amelyet a felhasználó parancssorból tud kezelni. A rendszerben a fizikális kiszolgálón („server gépen”) elindított virtuális R kiszolgáló megkapja az adatokat az adatbázis kezelőből, majd a számítások elvégzése után visszaadja az eredményt. A számításhoz a „survival” csomagot használtuk fel, amellyel a Kaplan-Meier görbét elkészítettük, valamint az egyes időpontban még bent maradó betegek számát, az esélyhányadost és a szignifikancia szintjét számoltuk és a görbe mellett

feltüntettük. Bár a számítások jelentős számítási kapacitást igényelnek, a rendelkezésre álló informatikai eszközök lehetővé teszik, hogy a rendszer gyorsasága miatt szinte valós idejű számításokról beszéljünk.

Magát a kezelőfelületet PHP-ban készítettük el. A www.kmplot.com cím alatt elérhető honlapot felosztottuk több részre, és egy adott elemzés valamennyi paramétere egyetlen oldalon állítható be (a későbbiekben még lesz szó további tumortípusok beépítéséről, amelyeket az eltérő klinikai adatok miatt azonban már csak további oldalak létrehozásával tudtunk megoldani - lásd a következő fejezetben). A számítás elindítása után az eredmények egy új oldalon jelennek meg. Az alábbi ábrán összefoglaltam a rendszer felépítését.



2. ábra. A világhálón keresztül elérhető rendszer lehetővé teszi a GEO-ban elérhető transzkriptomikus adatok valós idejű elemzését az R környezetben. A blokkdiagram az egymásra épülő lépéseket mutatja, amelyek eredménye a biomarker-jelölt hatásosságának elemzése.

Irodalomban azonosított biomarkerek keresztelemzése

Mivel jelen kutatás során nem volt célunk az irodalomban leírt összes szóba jövő biomarker-jelölt elemzése, ezért úgy döntöttünk, hogy az Amerikai Klinikai Onkológusok Társasága által már vizsgált sejtproliferáció mérésre alkalmazható biomarkereket¹⁵⁸ fogjuk a rendszer felhasználásával megvizsgálni. Az egyes biomarkerek estén a legjobban működő

próbákat az eredeti körülmények - mint szűrők - figyelembevételével osztályoztuk az online rendszer felhasználásával.

3.12. Fejlesztések többgénés osztályozók és optimalizált vágópont számítására

Gén chip adatok

A génexpressziós adatokat a fent leírt módon, a GEO és a TCGA központokból gyűjtöttük össze. A keresés során a „petefészek”, „rák”, „túlélés”, „GPL96”, „GPL570” és „GPL571” kulcsszavakat használtuk. Az adatok feldolgozása szintén a korábban már leírt módon, az R statisztikai környezetben történt.

A klinikai adatbázis összesen 1287 beteget tartalmaz, a medián teljes túlélés 31 hónap, esemény összesen 692 betegben következett be. A tumorok közül 927 szerózus, 27 pedig endometrioid, 33 betegben a tumor grade 1, 142-ben grade 2 és 779-ben grade 3. A sebészet sikeressége 1121 betegnél ismert, ezek közül 674-ben volt optimális a műtét (semennyi vagy kevesebb mint 1 cm visszamaradó tumorszövet). A betegek közül 985 esetében tudjuk a kemoterápiás kezelési protokollt is, közülük 914 kapott platina alapú terápiát és 616 kapott taxol kezelést is.

Kiszolgáló felépítése és statisztikai algoritmusok

Jelen rendszerben az adatokat már eleve a PostgreSQL adatbázisban tároltuk. A kiszolgáló operációs rendszere Debian Linux (www.debian.org), amely egy Apache webkiszolgálót (www.apache.org) futtat. Az R és a PostgreSQL adatbázis között a kapcsolatot az RODB Csomag biztosítja (<http://cran.r-project.org/package=RODBC>). A rendszer létrehozása során kizárólag ingyenesen elérhető, nyílt forráskódú programokat használtunk.

Több gén számítása során először valamennyi minta esetében a program kiválasztja a legjobb próbákat, majd kiszámítja ezek átlagos expresszióját. Ezután a Kaplan-Meier görbe készítésekor és a Cox regresszió számítása során ezzel az egy értékkel számít.

Mivel egy gént több próba is mér az Affymetrix chipeken, ezért fontos a legjobban teljesítő próbák kiválasztása. A korábbi rendszerben már alkalmaztunk erre egy módszert,

amelynek során a legmagasabb átlagos kifejeződést mutató próbát használtuk fel. Ezt azonban célszerű a felhasználó számára vizualizálni, aminek segítségével nemcsak a próba expresszió főbb leíró számait (pld medián, alsó és felső érték), de a próbák egyes mintákon belül mutatott kifejeződését is egy skálán lehet egyszerre bemutatni. Erre az R környezetben elérhető „beeswarm” csomagot alkalmaztuk, amely egydimenziós térben képes a próbák expresszióját ábrázolni (<http://www.cbs.dtu.dk/~eklund/beeswarm/>).

Figyelembe véve, hogy a CA125 már klinikailag is alkalmazott biomarker, ezért szükséges lehet egy olyan elemzés, ahol a CA125 által már diagnosztizált betegek alcsoportjaiban vagyunk kíváncsiak egy új biomarker-jelölt teljesítményére. Ezért beépítettünk a rendszerbe egy további szűrőt, amellyel csak azokat a betegeket vizsgáljuk, akiknél CA125 gén chip alapú meghatározása alapján az expresszió az alsó negyed alatt volt.

A rendszert szintén kiegészítettük egy további lehetőséggel, amely a legjobb vágópont meghatározását tartalmazza. Ennek során az adott gén kifejeződése alapján (vagy többgénos osztályozó esetén a gének átlagos expressziója alapján) az alsó és a felső negyed között valamennyi percentilis esetén elvégzi a program a Cox regressziós számítást. Ezután a legalacsonyabb p érték melletti vágópontot, illetve azonos p értékek esetén ezek között a legmagasabb HR-el társuló vágópontot fogadjuk el, mint legjobb értéket. A felhasználó által visszkapott Kaplan-Meier görbe ezen vágópont felhasználásával készül el.

Petefészekrák biomarkerek tesztelése

Korábban közölt biomarker jelölteket a PubMed-ben elvégzett irodalomkutatás során azonosítottunk. Ennek során a „petefészekrák”, „túlélés”, „biomarker”, és „génexpresszió” kulcsszavakat, valamint ezek kombinációit alkalmaztuk. Ezután valamennyi gént a HUGO gén nevezéktan szerint azonosítottunk (<http://www.genenames.org>).

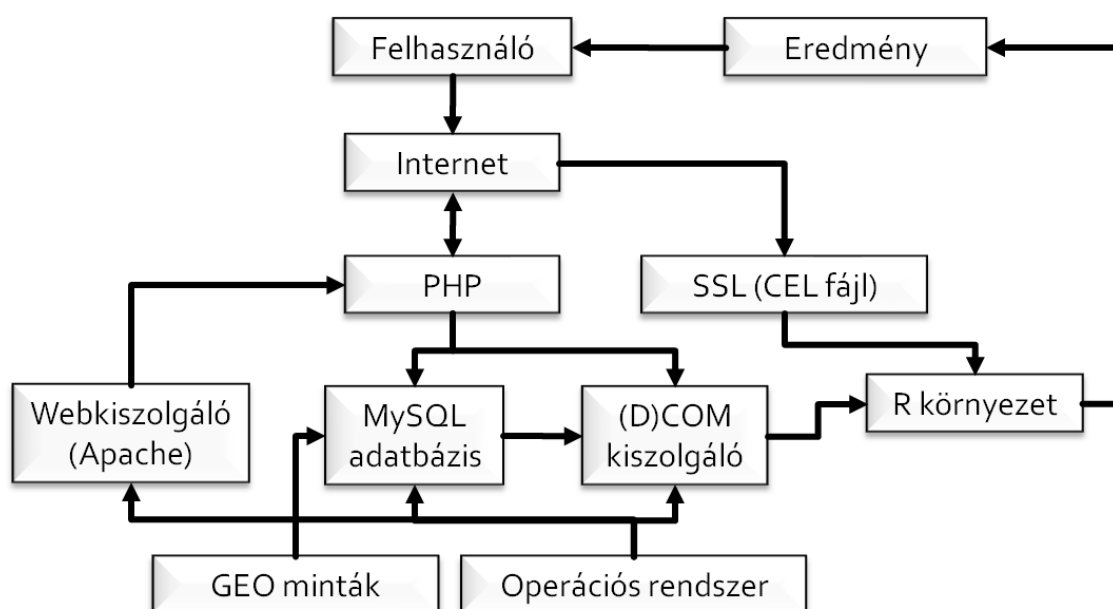
Az egyes gének hatásosságának vizsgálata során az elemzést minden esetben leszűkítettük azokra a betegekre, akikben az adott biomarkert az eredeti közleményben azonosították. Egy példán bemutatva az ERCC1 gént a taxol és platina terápiában részesült 3. stádiumú betegeken írták le. A validációs vizsgálat során mi is csak ezekben a betegeken vizsgáltuk a gén expressziójának összefüggését a túléléssel.

3.13. Online diagnosztikai rendszer fejlesztése

A kiszolgálórendszer fejlesztésének folyamata

A rendszer teszteléséhez használt adatbázis készítését és microarray adatok minőségellenőrzését és normalizálását az előző fejezetekben már részletesen bemutattuk.

A diagnosztikai szolgáltatást biztosító, szünetmentes tápegységgel ellátott kiszolgáló számítógép Debian Linux operációs rendszert futtat. Magát a honlapot Drupal (www.drupal.org) tartalomkezelő környezetben készítettük el. A honlapon belül a diagnosztikai rendszert egy különálló modul biztosítja, amely a (D)COM szerveren keresztül kapcsolódik az R kiszolgálóhoz, amely magát a statisztikai számításokat végzi. A számítások elvégzése előtt a gén chip mérés eredményét tartalmazó CEL fájlt (Affymetrix által meghatározott kiterjesztés, amely minden gén chip fájl esetében megegyezik) fel kell tölteni a kiszolgálóra. Az elvégzett számítások után a felhasználó a honlapon kapja vissza a számítások összefoglalását, illetve egy további részletesebb jelentést is letölthető PDF fájl formájában. A rendszer létrehozásához felhasznált valamennyi program ingyenesen elérhető, illetve felhasználható.



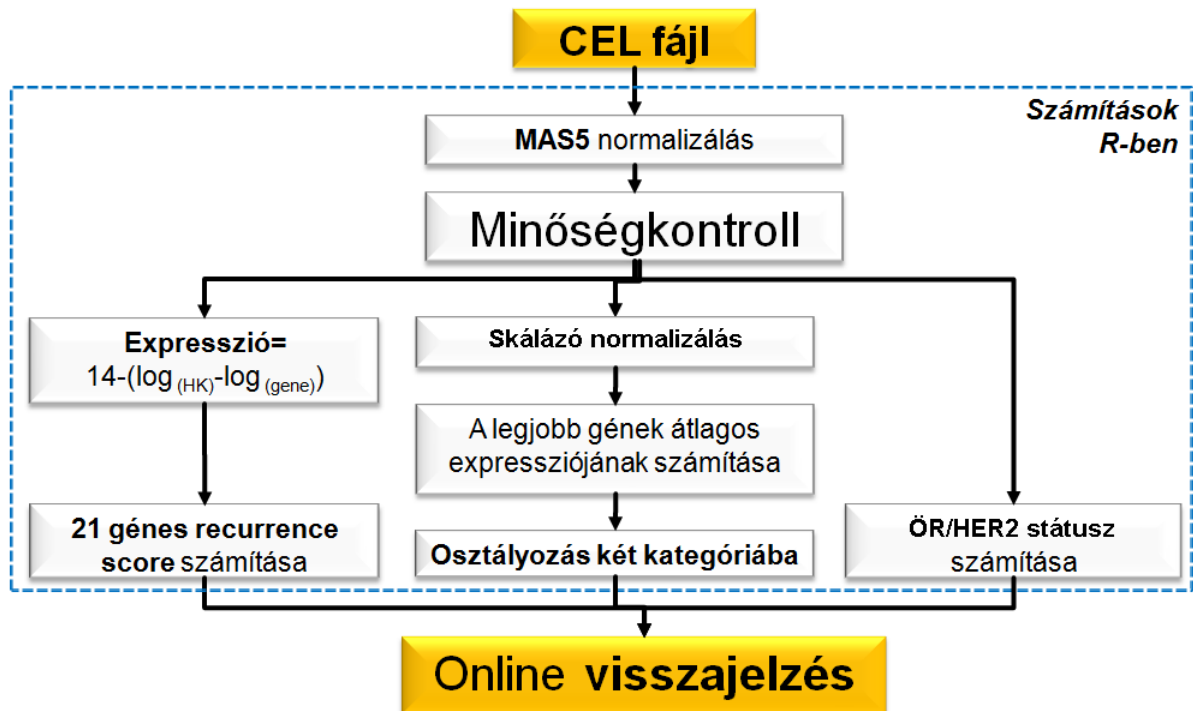
3. ábra. A kiszolgáló számítógépen működő rendszer sematikus ábrája. Az R környezetben történnek a számítások, amelyeket részletesen a következő ábrán mutatunk be.

A feltöltött CEL fájlra a rendszer a korábban alkalmazott kiértékelés valamennyi részlépését elvégzi, amely tartalmazza a MAS5 normalizálást, illetve a minőség ellenőrzését. Az ÖR/HER2 státusz számítását a korábban már leírt fix vágópont alkalmazásával (500 az ÖR esetében és 4800 a HER2 esetében) végzi el.

A „recurrence score” számítása már bonyolultabb, mivel az eredeti közleményben ⁶⁶ RT-PCR mérést végeztek. RT-PCR esetében a hírvivő RNS mennyiségét ciklusonként duplázva felszorzozzák. Eközben az általában magas kifejeződést mutató kontroll géneket is meghatározzák, majd az aktuálisan mért gének kifejeződéséből normalizálás során kivonják a háztartási gén esetében mért értéket. Ez tehát azt jelenti, hogy minél nagyobb a mért nyers génexpressziós érték, annál távolabb lesz a kontroll géntől, tehát a valódi génexpressziós szint annál kisebb lesz. Ezzel szemben gén chipok esetén a bekötődő mRNS-ek mennyisége arányos lesz a mért jelintenzitással, ami azt eredményezi, hogy magasabb érték esetén magasabb génexpresszió áll a mérés háttérében. Ezért ahhoz, hogy a gén chip adatokon az RT-PCR-re kifejlesztett algoritmust alkalmazni lehessen, az expressziós értékeken egy további transzformációt kellett végrehajtanunk. Ennek során a háztartási gén és az aktuálisan mért gén log expressziójának különbségét kivontuk 14-ből. A 14-es értéket „range top” paraméterként a rendszerben a felhasználó meg is tudja változtatni, de a legjobb összefüggéseket 14 esetén kaptuk, ezért a kiértékelés során mindenütt ezt az értéket használtuk fel. Az ezután megkapott génexpressziós értékeken a korábban leírt módon történik a „recurrence score” számítása, amelynek eredményeképpen a minták magas, közepes és alacsony rizikójú csoportokba kerülnek beosztásra.

A legerősebb gének azonosításához az előző fejezetben leírt rendszert használtuk fel, ezért az itt elvégzett számításokhoz az eredmények összehasonlíthatósága miatt egy további skálázó normalizálásra is szükség van. Ezután egyes klinikai csoportokra (összes beteg, nyirokcsomó pozitív betegek, nyirokcsomó negatív ösztrogén receptor pozitív betegek) elvégeztük az adatbázis összes génjére a medián érték vágópontként való alkalmazásával a Kaplan-Meier (Cox) elemzést. Az elemzés eredményeképpen a géneket sorba rendeztük. Ezután a legerősebb gént kombináltuk a második legerősebb génnel, majd újra elvégeztük a túlélési elemzést. Amennyiben a kombinált osztályozó nagyobb hatásosságot (alacsonyabb p érték és magasabb esélyhányados) eredményezett, akkor továbbléptünk, és még egy gént bevontunk az osztályozásba. Ezt addig ismételtük, ameddig az újabb gén bevonásával a hatásosság javult. A legerősebb génekkel a betegeket ezután két csoportba (jó és rossz

prognózis) osztottuk be annak alapján, hogy az adott mintában mért génexpressziós szintek a medián alá vagy felé kerültek.



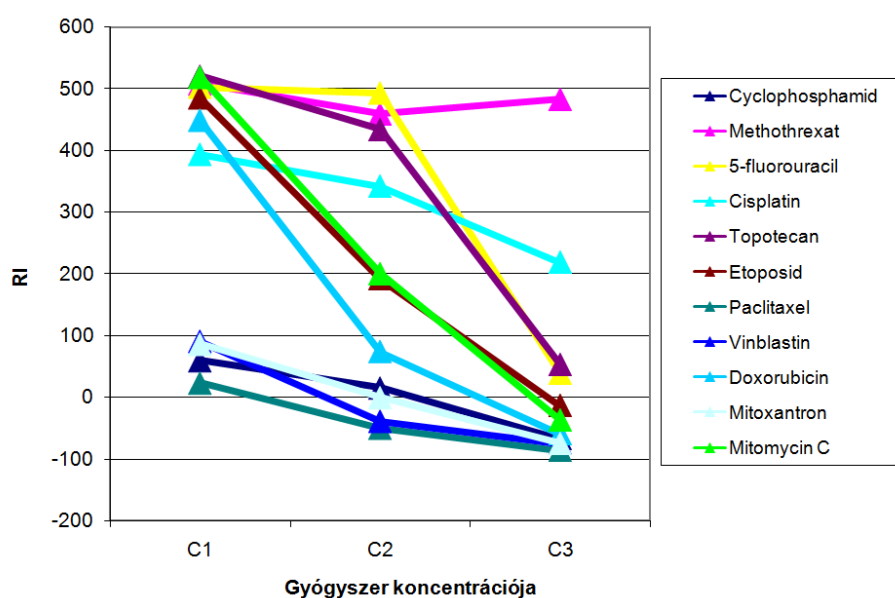
4. ábra. A recurrenceonline.com oldalon elvégzett számítások folyamatábrája. A feltöltött CEL fájl kiértékelése során három különböző előrejelzés készül, amelyekre a visszajelzést a felhasználó közvetlenül a holnapon kapja meg.

Az osztályozást az adatbázis valamennyi betegére elvégeztük, majd a kapott beosztást a klinikai (túlélési) adatokkal WinStat-ban és R-ben a „survival” csomag felhasználásával hasonlítottuk össze. Az egyes osztályozók hatásosságának egymással való összehasonlításához ROC elemzést végeztünk a MedCalc 11.6-os verziójának felhasználásával. Az Oncotype DX AUC értékeit irodalmi referenciákból vettük ¹⁹⁶.

4. EREDMÉNYEK

4.1. Kemoterápiás szerekkel szembeni rezisztencia-profil meghatározása

A vizsgálat során a következő gyógyszerekkel kezeltük a sejteket: 5-fluorouracil (C2-es koncentráció: 0,5 mikrogramm/ml), cisplatin (C2-es konc.: 0,5 mikrogr./ml), cyclophosphamid (C2-es konc.: 1,4 mikrogr./ml), doxorubicin (C2-es konc.: 0,05 mikrogr./ml), etopozid (C2-es konc.: 0,14 mikrogr./ml), methotrexat (C2-es konc.: 0,014 mikrogr./ml), mitomycin C (C2-es konc.: 0,05 mikrogr./ml), mitoxantron (C2-es konc.: 0,002 mikrogr./ml), paclitaxel (C2-es konc.: 0,025 mikrogr./ml), topotecan (C2-es konc.: 0,01 mikrogr./ml) és vinblastin (C2-es konc.: 0,01 mikrogr./ml). A három koncentrációra mért különböző értékeket az alábbi, az MDA-MB-231-es sejtvonalon készült grafikon mutatja.



5. ábra. Az MDA-MB-231-es sejtvonalon mért rezisztencia-értékek C1, C2 és C3 koncentráció mellett az egyes gyógyszerekre

Ezt elvégeztük valamennyi sejtvonalra, ami alapján minden gyógyszerre létre tudtunk hozni egy rezisztens és egy érzékeny sejteket tartalmazó panelt. A vizsgálat során eredetileg 12 gyógyszert mértünk le, azonban gemcitabin esetén mindössze egyetlen rezisztens vonalat azonosítottunk (és ezt a koncentráció váltás sem tudta növelni), ezért a további vizsgálatok során ezt a gyógyszert figyelmen kívül hagytuk.

Sejtvonal	Eredet	5-Fluorouracil	Cisplatin	Cyclophosphamid	Doxorubicin	Etoposid	Methotrexat	MitomycinC	Mitoxantron	Paclitaxel	Vinblastin	Topotecan @ C3	Mitoxantron @ C1
181/85p	hasnyálmirigy	É	R	É	K	R	K	R	É	É	É	K	R
257p	gyomor	R	R	É	É	R	K	R	K	É	É	R	R
A375	melanoma	K	R	R	K	R	K	K	K	É	É	R	R
BT20	emlő	É	É	R	É	É	É	É	É	É	É	É	É
C8161	melanoma	K	R	R	K	R	K	K	K	É	É	R	R
Colo699	tüdő	É	K	É	É	É	R	K	É	É	É	K	É
CX-2	colon	É	R	É	É	K	É	É	É	É	É	É	É
Du145	prosztata	R	R	R	É	R	R	É	É	É	É	K	K
DV-90	tüdő	É	É	É	É	É	É	É	É	É	É	É	É
ES-2	petefészek	É	R	É	R	R	R	R	R	R	R	R	É
FU-OV-1	petefészek	É	É	R	É	K	R	É	É	É	É	K	R
Hep3B	HCC	R	R	R	K	K	R	R	K	K	K	É	R
HRT-18	colon	É	R	É	É	K	R	É	É	É	É	É	K
HT-29	colon	É	R	É	É	R	K	É	É	É	É	É	R
MDA231	emlő	R	K	É	É	É	R	É	É	É	É	É	É
ME43	melanoma	É	R	É	É	É	R	É	É	É	É	É	É
MeWo	melanoma	É	R	É	É	É	R	É	É	É	É	É	K
OAW42	petefészek	K	É	R	K	R	R	É	K	É	É	R	R
OVKAR-3	petefészek	É	É	É	É	R	R	É	É	É	É	K	R
R103	emlő	É	É	É	É	É	É	É	K	É	É	É	É
R193	emlő	É	É	É	R	R	É	R	K	R	R	R	R
SKBR3	emlő	É	R	É	R	R	R	R	K	É	R	K	É
SKMel13	melanoma	É	K	K	K	K	É	É	K	É	É	É	K
SKMel19	melanoma	K	R	R	K	K	R	É	É	É	É	K	K
SKOV-3	petefészek	É	R	É	É	R	R	K	É	É	É	K	K
SNU182	HCC	K	R	R	K	R	R	R	K	É	É	É	K
SNU423	HCC	K	R	R	K	R	R	K	K	É	É	K	R
SNU449	HCC	R	R	R	R	R	R	R	É	R	R	R	R
SNU475	HCC	R	R	R	É	R	R	É	É	É	É	K	R
SW13	prosztata	R	R	R	É	R	K	K	É	É	É	K	É

2. táblázat. A C2 (illetve két gyógyszer esetében a C1 és a C3) koncentráció mellett a sejtvonalakat az RI értékek alapján rezisztensként (R) vagy érzékenyként (É) osztályoztuk. (K=közepes érték)

Ezután a sejtvonalak génexpressziós adatait gyógyszerenként különválasztva, két csoportot (rezisztens és érzékeny) képezve elemeztük tovább a fent leírt módon. Az egyes gyógyszerekkel szembeni rezisztenciával összefüggő gének jelen értekezés szempontjából nem annyira fontosak, mint azon gének, amelyek több gyógyszerrel szembeni rezisztenciával is összefügghetnek. Érdekes módon nem találtunk olyan univerzális rezisztencia-gént,

dc_718_13

amelyik minden gyógyszerrel szembeni érzékenységgel összefüggést mutatott volna. A gének közül azokat, amelyeket legalább 5 gyógyszerrel szembeni rezisztenciával összefüggésbe hoztunk, az alábbi táblázatban foglaltuk össze:

Gén	Gén teljes név	# gyógyszer	Vinblastin	Topotecan	Paclitaxel	MTX	Mitoxantron	Mitomycin C	Etoposide	Doxorubicin	Cyclophosphamide	Cisplatin	5-FU
TRIM2	tripartite motif-containing 2	8	X	X	X		X	X	X	X			X
APOBEC3B	apolipoprotein B mRNA editing enzyme	7	X	X	X		X	X	X	X			
BIRC2	baculoviral IAP repeat-containing 2	6	X	X	X			X	X	X			
TFPI2	tissue factor pathway inhibitor 2	6	X	X	X			X	X	X			
C10orf38	chromosome 10 open reading frame 38	6		X		X		X	X	X			X
ADD3	adducin 3 (gamma)	5	X	X	X					X			X
IGFBP7	insulin-like growth factor binding protein 7	5		X		X	X		X	X			
PRG1	proteoglycan 1, secretory granule	5	X	X	X			X		X			
SERPINE1	serine proteinase inhibitor, clade E, member 1	5	X	X	X			X		X			
DDIT4	DNA-damage-inducible transcript 4	5		X			X	X	X		X		
COX17	COX17 homolog	5	X	X	X			X		X			
SERPINB2	serine proteinase inhibitor, clade B , member 2	5	X	X	X			X		X			
LTBP2	latent transforming growth factor beta binding protein 2	5	X	X	X			X		X			
AOX1	aldehyde oxidase 1	5	X	X	X			X		X			
RAGE	renal tumor antigen	5	X	X	X			X		X			
FDFT1	farnesyl-diphosphate farnesyltransferase 1	5		X	X		X	X	X				
MAGED1	melanoma antigen, family D, 1	5		X		X	X				X		X
TFPI2	tissue factor pathway inhibitor 2	5	X	X				X	X	X			
211538_s_at	---	5		X			X		X		X		X
TFPI	tissue factor pathway inhibitor	5		X				X	X	X			X
DKK3	dickkopf homolog 3 (Xenopus laevis)	5	X	X				X	X	X			
GPSM2	G-protein signalling modulator 2	5	X	X				X	X	X			

3. táblázat. Azon gének listája, amelyek legalább 5 különböző gyógyszerrel szembeni rezisztenciával összefüggést mutattak a 30 sejtvonal génexpressziós mintázatának elemzése során

Az egyes gének előrejelzésének hatékonyságát ugyanezen sejtvonalakban teszteltük úgy, hogy a teljes statisztikai elemzést lefuttattuk egy sejtvonal kihagyásával (mindegyik gyógyszerre külön-külön), majd ezután ezt az egy sejtvonalat osztályoztuk az így azonosított génekkel. Ennek során összesen 220 helyes (ahol tehát a rezisztenciát vagy az érzékenységet

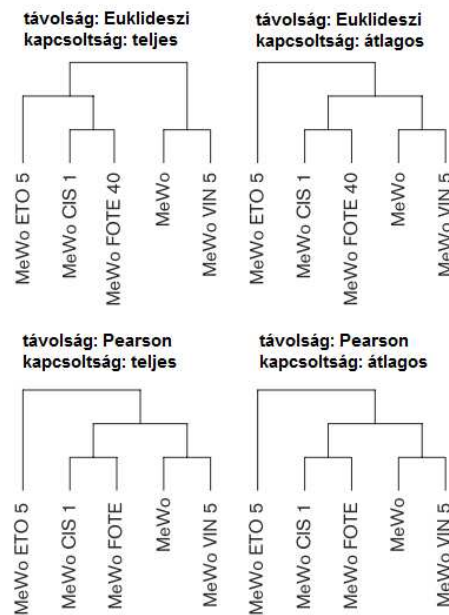
előre tudtuk jelezni) és 36 hibás osztályozás készült, amely eredmény a teljes kísérletre vetítve összesen 86%-os pontosságnak felel meg.

Végezetül az egyes gyógyszerekkel szemben meghatározott legfontosabb rezisztencia gének kifejeződését TaqMan RT-PCR-el is ellenőriztük. A microarray technológia hátránya, hogy olyan nagyszámú gént azonosítunk az elemzések során, hogy ezek mindegyikének pontos ellenőrzése már csak nagyon nehezen lenne megvalósítható. Ezért az RT-PCR során a legfontosabb 46 gén mérését tűztük ki célul. Az ellenőrzés során a kérdés azt volt, hogy a microarray-eken megfigyelt különbség az RT-PCR méréseken is szignifikáns lesz-e? Az RT-PCR kiértékelésnél ezért nem korrelációt számoltunk, hanem a sejtek rezisztencia-értékei alapján szétosztottuk az értékeket két csoportra (rezisztens és érzékeny), majd ezt a két csoportot hasonlítottuk össze t-próbával. Bár csak 46 gént ellenőriztünk, azonban olyan géneket választottunk, amelyek több gyógyszerrel szembeni rezisztencia-profilban is részt vettek, így összesen 76 gyógyszer-rezisztencia párt tudtunk ezzel a független módszerrel is igazolni $p < 0,05$ -ös szignifikancia szint mellett.

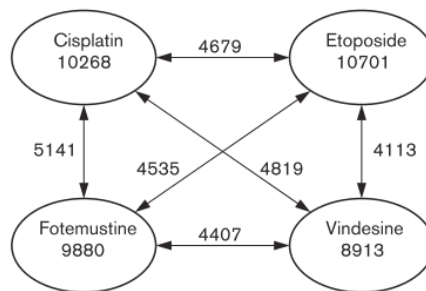
4.2 A kemorezisztencia génjei melanoma malignumban

Összesen négy, a melanoma kezelésében használt gyógyszerrel (cisplatin, etoposid, fotemustin és vindesin) szemben ellenálló sejtvonal génexpressziós mintázatát határoztuk meg, majd ezt hasonlítottuk a szülői sejtvonal mintázatához. Az eredmények alapján a gének 21-25%-a eltérő kifejeződést mutatott.

Négy különböző fürtelemzést végeztünk, amelynek célja volt, hogy meghatározzuk, hogy az egyes rezisztens vonalak milyen mértékben hasonlóak egymáshoz. A legkisebb különbséget a szülői és a vindesin rezisztens leányvonal között figyeltük meg. A cisplatin és a fotemustin rezisztens vonalak hasonló - bár a többiektől távoli - mintázatot mutattak. Valós idejű RT-PCR mérésekkel igazoltunk két gént, az SLC40A1-et ($p=0,0005$) és az LRAP-ot ($p=0,009$). A fürtelemzés eredményét és az egyes vonalak közötti hasonlóan kifejezett gének számát az alábbi két ábrán foglaltam össze.



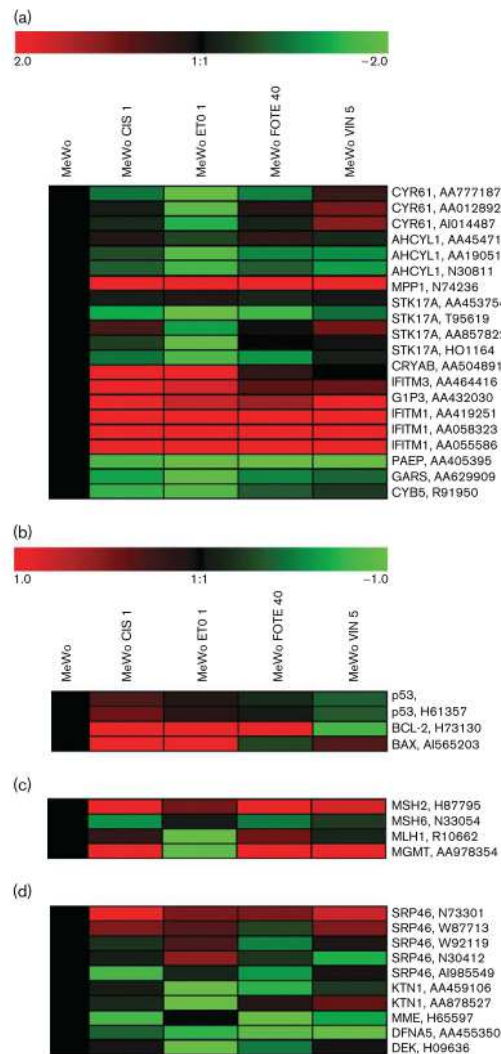
6. ábra. A teljes genomot mérő génexpressziós chip segítségével meghatározott mintázat alapján készített fürtelemzés mutatja, hogy a melanoma kezelésében felhasznált gyógyszerek közül melyekkel szembeni rezisztencia mutat hasonló génexpressziós mintázatot.



7. ábra. A különböző sejtvonalak közötti közösen felül/alul expresszált gének száma. A cisplatin és fotemustine közötti rezisztencia hasonlóságot a két vonal közötti nagyszámú közösen változott gén is alátámasztja.

A szakirodalomban korábban számos olyan gént azonosítottak, amelyek az egyes gyógyszerekkel szembeni rezisztenciával összefüggést mutattak ^{168, 197-199}, és amelyeket sikerült a gén chip adatok felhasználásával is igazolni (lásd következő ábra).

dc_718_13

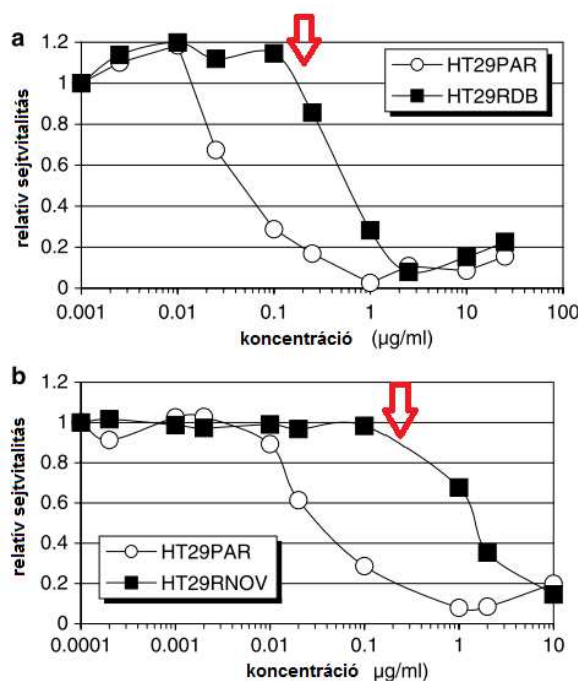


8. ábra. Az irodalomban korábban leírt gének expresszióváltozása az általunk vizsgált sejtvonalakban. Az egyes géneket Wittig (A), Kern (B), Lage (C) és Grottke (D) közleményeiből^{168, 197-199} választottuk ki.

4.3. Túlélés előrejelzése sejtkultúra modellben azonosított gének felhasználásával

A gén chip vizsgálatok előtt meghatároztuk az egyes sejtvonalak relatív sejtvitalitását a gyógyszerek folyamatosan emelkedő koncentrációja mellett. Az alábbi reprezentatív ábrán jól látszik, hogy a rezisztens leánysejtek szigmoid alakú sejtvitalitás-görbéje a logaritmus skálájú koncentrációs tengelyen a szülői sejtvonalakhoz képest jobbra tolódott. Meg kell jegyeznünk, hogy a rezisztencia mért értékei a korábbi irodalmi adatokat is igazolták, ezáltal alátámasztva, hogy a rezisztencia a sejtvonalak stabil, öröklődő tulajdonsága. A szülői és

leánysejtekre vonatkozó két görbe közötti hasadék lehetővé teszi, hogy meghatározzunk egy olyan koncentrációt, amelynél a szülői sejtvonalak már szinte egyáltalán nem mutatnak mérhető sejtvitalitást, miközben a rezisztens leányvonalak még alig mutatnak vitalitáscsökkenést. Ez az érték az, ahol a relatív vitalításban a legmagasabb az eltérés - ezen, az ábrákon piros nyíllal jelölt koncentrációk mellett végeztük a sejtekből az RNS izolálását a gén chip és az RT-PCR mérésekhez.



9. ábra. A HT29 sejtvonal (HT29PAR), valamint a daunorubicin (HT29RDB) és mitoxantron (HT29RNOV) rezisztens leányvonalai a daunorubicin (A) és mitoxantron (B) különböző koncentrációival kezelve eltérő sejtvitalitást mutatnak. A nyíllal jelölt koncentráció mellett határoztuk meg az egyes sejtvonalak génexpressziós mintázatát.

A sejtvonalakat két csoportra osztottuk, és a daunorubicin / doxorubicin illetve mitoxantron rezisztens leányvonalakat külön hasonlítottuk a szülői vonalakhoz. Valamennyi sejtvonalból készült gén chip mérés, a chipen az egyes géneket külön ellenőriztük, hogy valóban adtak-e mérhető jelet? Azon géneket, amelyek nem voltak legalább három rezisztens vonalban mérhetőek, kizártuk az elemzésből, így összesen 38 ezer próba maradt a doxorubicin rezisztens csoportban és 33 ezer próba a mitoxantron rezisztens csoportban. Ezután PAM elemzés elvégzésével rangsoroltuk az összes gént és azonosítottuk azokat a legfontosabb géneket, amelyek a küszöb alkalmazása után is szignifikánsak maradtak.

Eredményeink igazolták több, korábban már leírt gén eltérő kifejeződését a rezisztens vonalak esetében, mint a PGP²⁰⁰ és a BCRP²⁰¹.

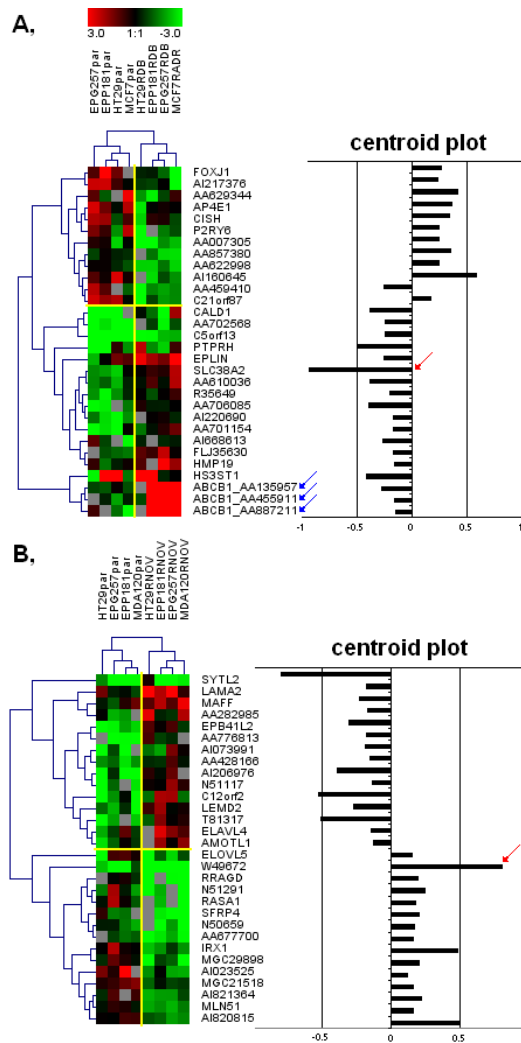
A következő ábrán a doxorubicin és mitoxantron rezisztenciával kapcsolatba hozott legerősebb gének szerepelnek. Az ábrán két okból is külön kiemeltém az ABCB1 gént. Egyrészt a gén chipek három különböző próbát is tartalmaztak a mérésére - ezen próbák expressziója teljesen együtt mozgott, másrészt nemcsak hogy mindhárom próba szignifikáns expresszióváltozást mutatott, de a fürtelemzés folyamán az egyes próbák egymás mellé is kerültek.

Ezután az ABCB1 gén kifejeződését RT-PCR-el is lemértük valamennyi sejtvonalban, amelynek során négy daunorubicin / doxorubicin rezisztens sejtvonal közül háromban a gén expresszió-különbségét sikerült igazolunk - meglepő módon a HT29 sejtvonalpáron belül a szülői sejtvonal mutatta a magasabb génexpressziót.

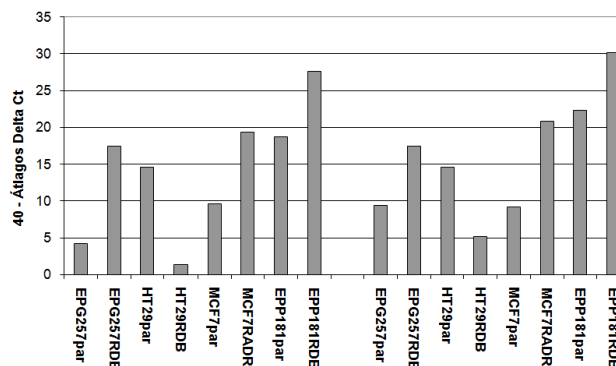
Ezután következett a vizsgálat legfontosabb része, a független klinikai mintákon való tesztelés. Ehhez Sorlie és munkatársai által korábban közölt adatokat⁵³ használtunk, amely vizsgálat összesen 44, doxorubicin monoterápiával kezelt beteget tartalmazott. Ezen vizsgálatban ugyanolyan gén chip platformot használtak, mint amivel mi a sejtvonalainkat elemeztük, az adatokat a Stanford Microarray Database-ből töltöttük le. Sorlie-ék mérése korábban történt, ezért az általuk alkalmazott gén chip még nem tartalmazta az összes általunk is vizsgált próbát (bár a technológia és az alkalmazott próbák azonosak voltak). Az általunk azonosított legjobb gének emiatt nem voltak meg mind a betegekre is.

Ezen okból kifolyólag nem az általunk azonosított legjobb géneket használtuk az osztályozáshoz, hanem egy alternatív kiértékelést végeztünk. Ennek során először kiszűrtük a sejtvonalakon készült gén chip adatokat, hogy csak azokat a géneket tartalmazzák, amelyekről klinikai mintákból is volt adat. A következő lépésben a PAM elemzést újra lefuttattuk, de a küszöböt 1-re állítottuk be. A PAM beépített osztályozó funkciójával ezután az összes betegre készítettünk egy előrejelzést. Az általunk rezisztensként osztályozott beteg osztályozásának alapja, hogy a „doxorubicin rezisztens sejtvonalakhoz mutatott nagyobb hasonlóságot”, míg az érzékenyként osztályozott betegek az „érzékeny vonalakhoz hasonlítottak jobban”.

dc_718_13

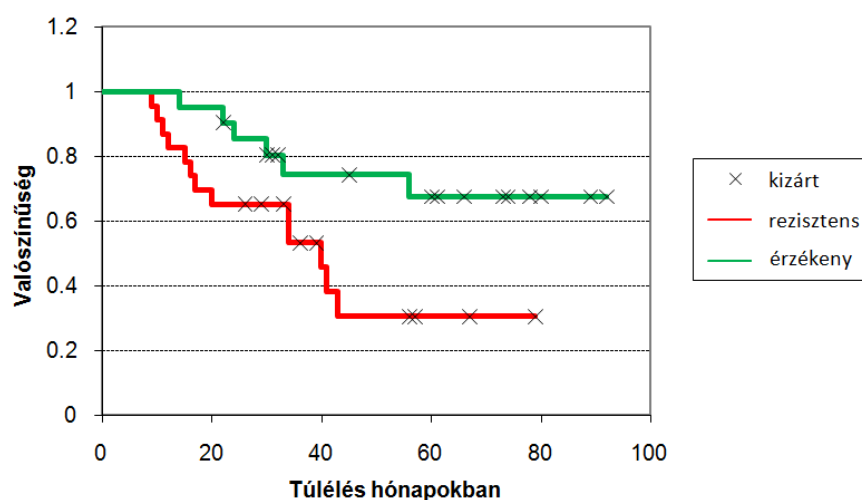


10. ábra. PAM elemzéssel azonosított legfontosabb gének centroid ábrája a négy doxorubicin-rezisztens (A) és a négy mitoxantron-rezisztens (B) leány sejtvonalnak a négy szülői vonallal történő összehasonlítása alapján. Az ábrák bal oldala ugyanezen gének hierarchikus fürtelemzésének ábráját mutatja. A piros nyilak a centroid plot legerősebb génjét jelölik, a kék nyilak az ismételt próbák reprodukálhatóságát illusztrálják.



11. ábra. Az ABCB1 gén kifejeződése valamennyi sejtvonalban. A HT29PAR sejtvonal az egyetlen, ahol a szülői sejtvonalban magasabb a gén kifejeződése

A betegek osztályozása eredményeképpen két csoportot kaptunk, amelyeket Kaplan-Meier elemzéssel hasonlítottunk össze (lásd következő ábra). A rezisztensként osztályozott betegek (akiknél tehát az előrejelzés alapján a doxorubicin monoterápia nem lesz hatásos, $n=23$) túlélése $32,9 \pm 18,7$ hónap volt, miközben az érzékeny betegek (akiknél tehát az előrejelzés alapján a doxorubicin monoterápia hatásos lesz, $n=21$) túlélése $49,7 \pm 26,1$ hónap volt. Hasonló elemzést a mitoxantron-rezisztencia vizsgálatára nem tudtunk végezni, mivel nem állt rendelkezésünkre ehhez felhasználható beteganyag.



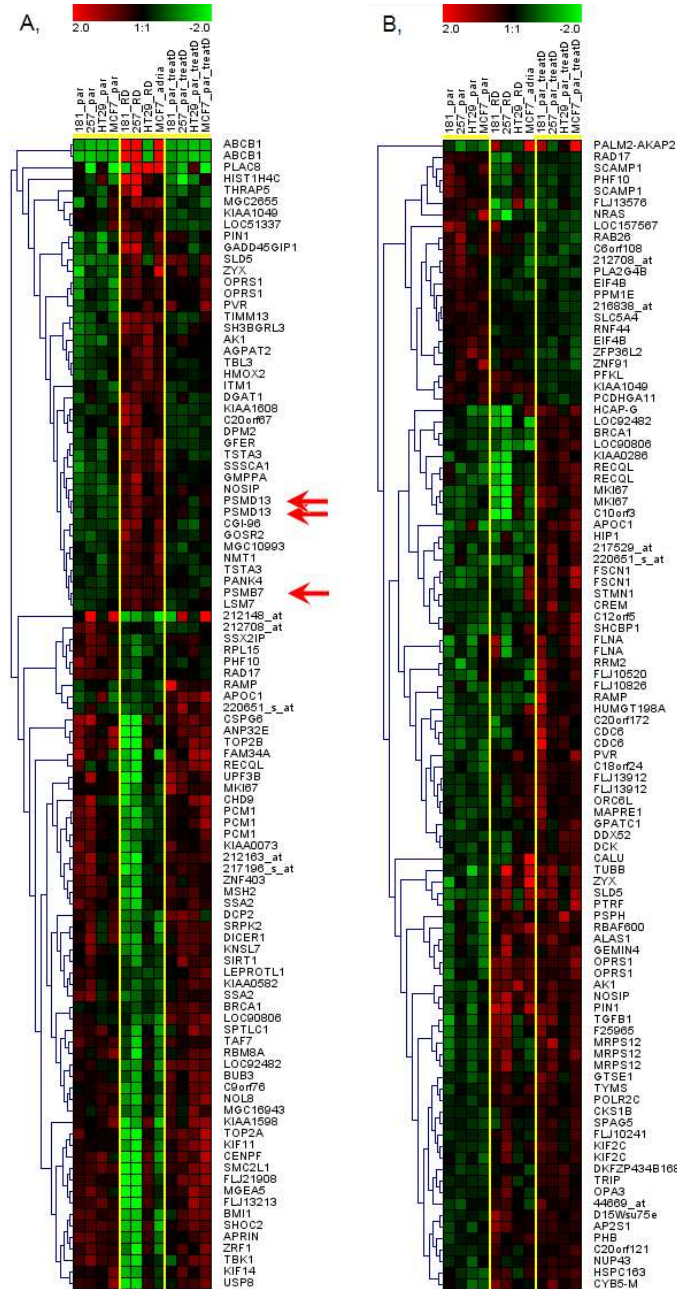
12. ábra. Doxorubicin monoterápiában részesült betegek túlélése gén chip adatok alapján rezisztensként és érzékenyként osztályozott betegek között szignifikánsan eltért ($p=0,0334$). Az osztályozás a rezisztens, illetve érzékeny sejtvonalakban mért génexpressziós mintázathoz való hasonlóság alapján történt.

4.4. A PSMB7 gén oki szerepének igazolása *in vitro* és *in silico*

A doxorubicin/daunorubicin rezisztencia génjeinek azonosításához rezisztens és érzékeny, valamint gyógyszerrel kezelt érzékeny sejtvonalak érzékenységet hasonlítottuk össze. A mérések egy olyan koncentráció mellett történtek, amely mellett a rezisztens vonalak az érzékeny vonalakkal szemben még nem mutatnak jelentős proliferáció-csökkenést. A rezisztenciával összefüggő gének közül azokat fogadtuk el szignifikánsnak, amelyek nem mutattak eltérő expressziót a gyógyszerrel való kezelés hatására, vagyis csak a rezisztens

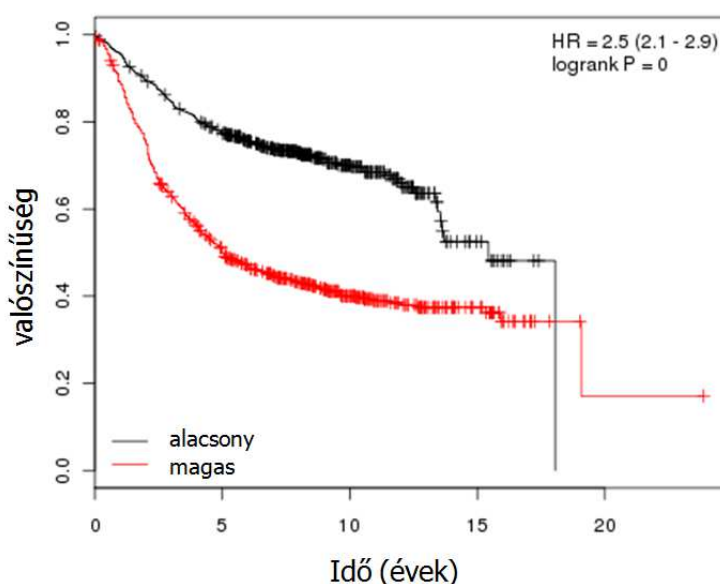
dc_718_13

vonatokra voltak jellemzőek. Az eltérően kifejeződő gének között összesen kettő proteaszóma-alegységet azonosítottunk, a PSMD13-at és a PSMB7-et (lásd következő ábra).



13. ábra. Hierarchikus fürtelemzés segítségével vizualizáltuk a mindhárom alcsoportot megkülönböztető (A) és a szülői és a kezelt szülői sejtvonalakat megkülönböztető (B) géneket. A piros nyilak a proteaszóma alegységeket jelölik. Piros szín: magasabb, zöld: alacsonyabb expresszió; par: szülői vonalak; RD: rezisztens leányvonalak; treadD: gyógyszerrel kezelt vonalak.

A GEO adatbázisaiból összesen 1592 beteg adatait töltöttük le, ezekben a betegekben a 64 ismert proteaszóma alegység expressziójának hatását vizsgáltuk a kiújulásmentes túlélésre a génexpressziók mediánja alapján. Összesen 12 proteaszóma alegység mutatott szignifikáns hatást, amelyek között a két fenti alegység közül a PSMB7 volt jelen. A PSMB7-et az Affymetrix gén chipen a 200786_at jelű próbával határoztuk meg. A PSMB7 $p < 1E-16$ mellett igen jelentős összefüggést mutatott a túléléssel - a mediánnál magasabb kifejeződést mutató betegek prognózisa rosszabb volt, mind a mediánnál alacsonyabb kifejeződést mutató betegeké. A rendelkezésre álló klinikai adatok hiányosságai miatt nem tudtuk vizsgálni, hogy a gén kifejeződése milyen hatással van a túlélésre a különböző protokollokkal kezelt betegek esetében. A további vizsgálatok során a többi alegységet nem vizsgáltuk, mivel azok nem mutattak összefüggést a túléléssel.

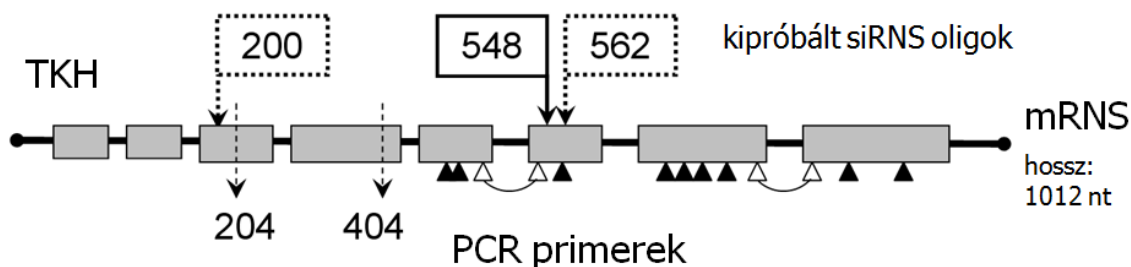


14. ábra. Kaplan-Meier túlélési elemzés 1592 emlőrákos betegben. A betegeket két csoportra osztottuk az alapján, hogy a PSMB7 gén (200786_at próba a gén chipen) expressziója a medián felett (magas) vagy alatt (alacsony) volt.

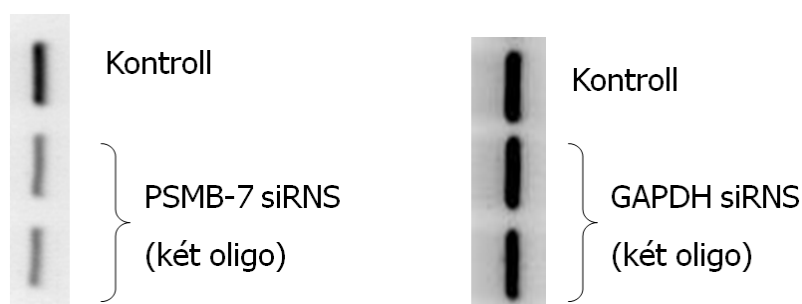
A PSMB7 funkcionális hatását RNS interferencia kísérlettel vizsgáltuk meg. A génre három különböző siRNS oligot terveztünk, amelyek a 200-as, 548-as és az 462-es pozíciókhoz kötöttek be. A PSMB7 csendesítésének mértékét PCR vizsgálattal igazoltuk, a PCR primerek úgy terveztük, hogy a két kötőhely közötti szakasz ne tartalmazzon

dc_718_13

kötőhelyet az siRNS oligokhoz, hogy ezáltal azt a hatást tudjuk mérni, amely a már aktivált dicer enzim által lebontott mRNS-ek mennyiségi meghatározását is tartalmazza.



15. ábra. A PSMB7 gén áttekintő képe. Az exonokat szürke dobozok, az Affymetrix microarray-eken található próbákat háromszögek jelölik. A fekete háromszöggel jelzett próbák egy-egy exonhoz kötődtek be, az üres háromszögek az exonok kapcsolódó végén kötnek be. Az ábra felső részén a kipróbált siRNS oligók pozícióit, az alsó részén a PCR primerek kötőhelyeit nyilak jelölik. TKH: transzkripció kezdőhely

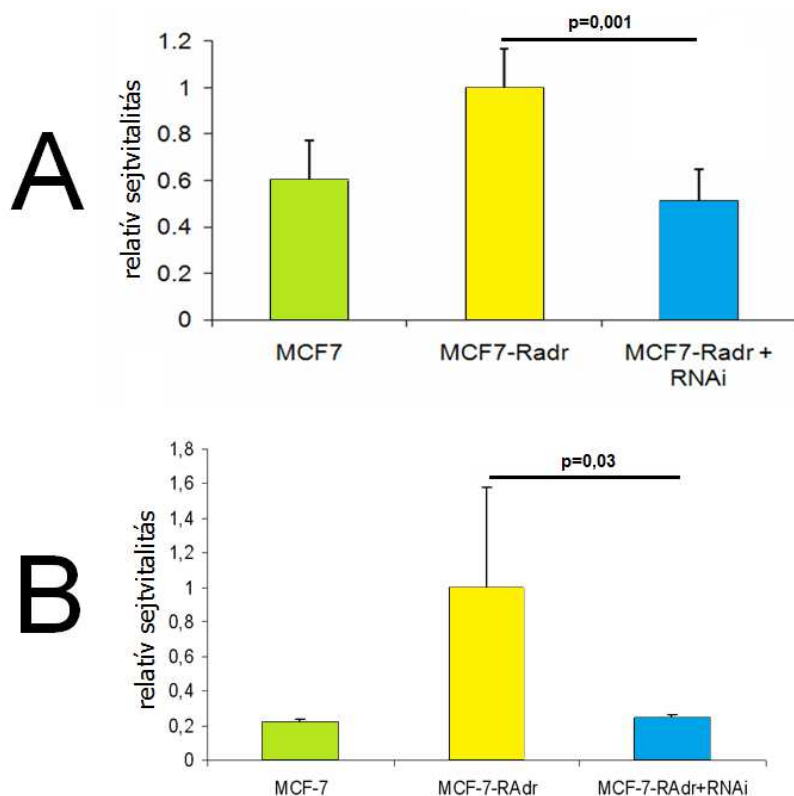


16. ábra. Az siRNS-el végzett génelcsendesítés hatása a génexpresszióra. A leghatásosabb (87%-os) PSMB7 elcsendesítést az 548-as pozícióban bekötődő siRNS molekulával értük el

A rezisztens sejtek 79,8%-a élt túl a doxorubicin kezelés mellett, amely 31%-ra csökkent az 548-as pozícióra bekötő siRNS oligoval végzett géncsökkentés mellett. Az siRNS kezelt és kezeletlen, doxorubicinnal kezelt sejtek között a különbség magasan szignifikáns volt ($p=0,001$). Az érzékeny sejtek 48%-a élt túl doxorubicin kezelés mellett, míg a negatív kontroll siRNS kezelés után a sejtek 73%-a maradt életben.

A rezisztenciára gyakorolt hatást paclitaxel kezelés mellett is megvizsgáltuk. Az érzékeny sejtek 43%-a élte túl a kezelést, míg a kombinált siRNS és gyógyszeres kezelés

hatására a rezisztens sejteknek mindössze 23%-a maradt életben. A két csoport között a különbség itt is szignifikáns volt ($p=0,03$).

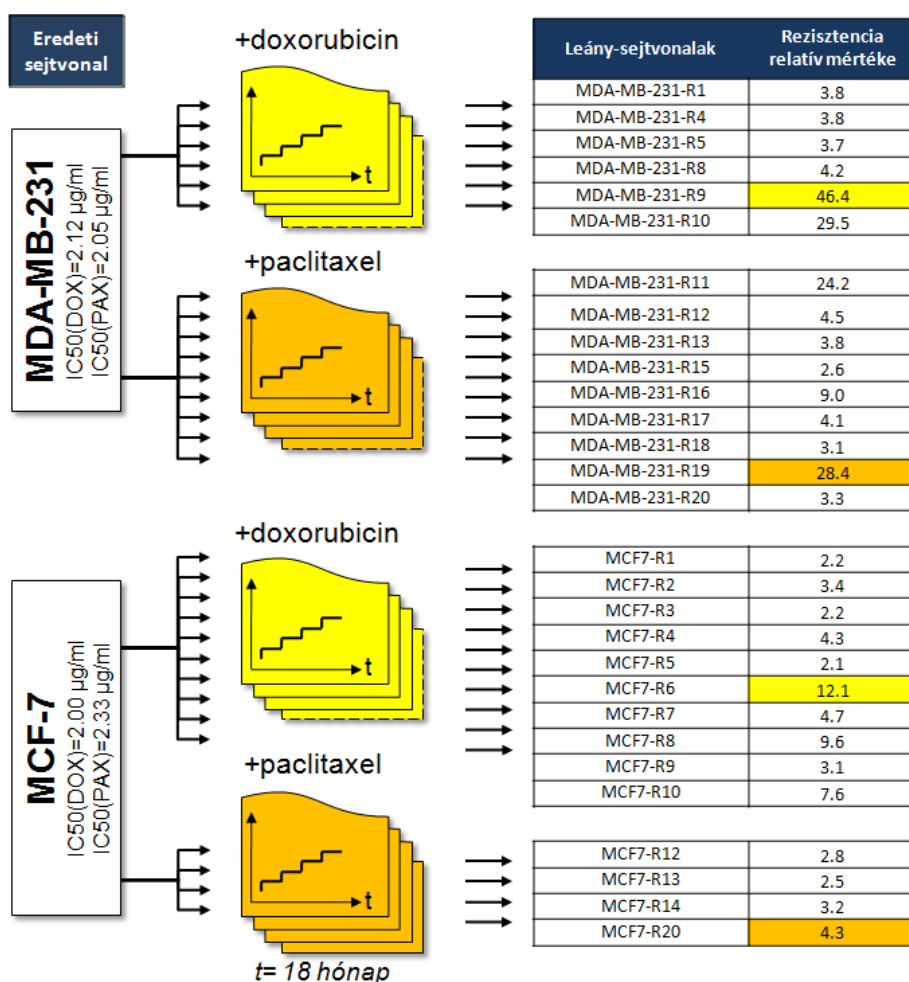


17. ábra. A PSMB7 csendesítésének hatása a sejtek relatív vitalitására gyógyszeres kezelés mellett a kezeletlen vonalakhoz viszonyítva doxorubicin (A) és paclitaxel (B) kezelés mellett. A rezisztens szülői vonal (zöld) nagymértékű pusztulást mutat, a rezisztens leányvonal (MCF-7-RAdr, sárga) vitalitás-csökkenése nem észlelhető. Az elcsendesített PSMB7 mellett a sejtek vitalitása az érzékeny szülői vonalakéval volt azonos.

4.5. Parallel evolúciós modell a párhuzamosan létrejövő rezisztencia modellezésére

A rezisztens vonalak létrehozása során a koncentrációemelések után előfordult, hogy egyes vonalak olyan mértékben kihaltak, hogy a további kezelés már nem volt lehetséges. Ezért a másfél éves kezelés eredményeképpen az induló 40 vonalból 27 rezisztens leányvonalat fejlesztettünk ki. Az IC50 értéke valamennyi sejtvonalban legalább a kétszeresére nőtt. Az egyes vonalakban a három másik gyógyszerrel szembeni keresztrezisztencia mérhető volt, azonban nem mutatott összefüggést a szelekcióhoz használt szerrel szembeni rezisztenciával. Az alábbi ábrán a rezisztencia létrehozásának

folyamatábráját, valamint az egyes vonalakban a használt gyógyszerrel szemben elért érzékenység-változás mértékét tüntettem fel.

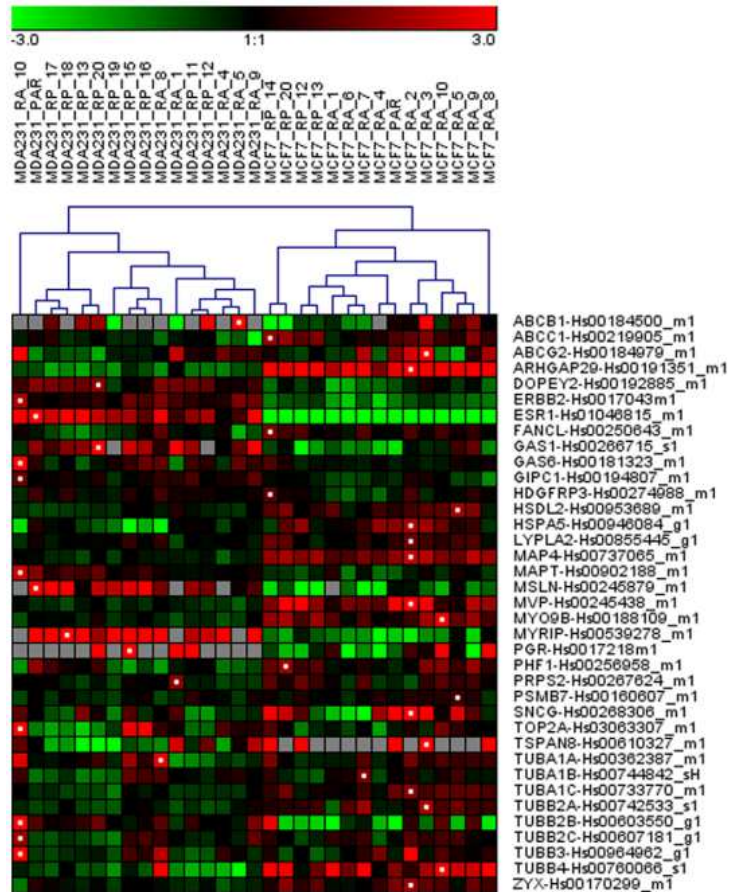


18. ábra. A rezisztencia párhuzamosan létrehozott modelljének áttekintő folyamatábrája és a rezisztencia egyes sejtvonalakban elért relatív szintje. Narancssárgával a paclitaxellel szemben legellenállóbb vonalakat, citromsárgával a doxorubicinnal szemben legellenállóbb vonalakat jelöltem meg.

Az RT-PCR-el vizsgált próbák a doxorubicin- és a paclitaxel-rezisztenciával kapcsolatba hozott gének mellett hormon receptorokat és túléléssel kapcsolatba hozott géneket is tartalmaztak. Amikor azt vizsgáltuk, hogy melyik az a mechanizmus, amelyik leginkább hozzájárul egy adott sejtvonalban a rezisztencia kialakulásához, akkor azt találtuk, hogy egy-egy sejtvonalban mindössze egy-két mechanizmus volt egyidejűleg aktív. Az

dc_718_13

áramlási citometriás mérésekkel nem figyeltünk meg egyértelmű összefüggést a rhodamin kiáramlás és a rezisztencia mértékek között.



19. ábra. Az RT-PCR-el mért gén expressziós értékek alapján színskóddal ellátott fűrdiagram, amelyben valamennyi sejtvonalat ábrázoltunk. A fehér pontok azt a mintát jelölik meg, ahol az adott gén a legmagasabb kifejeződést mutatta.

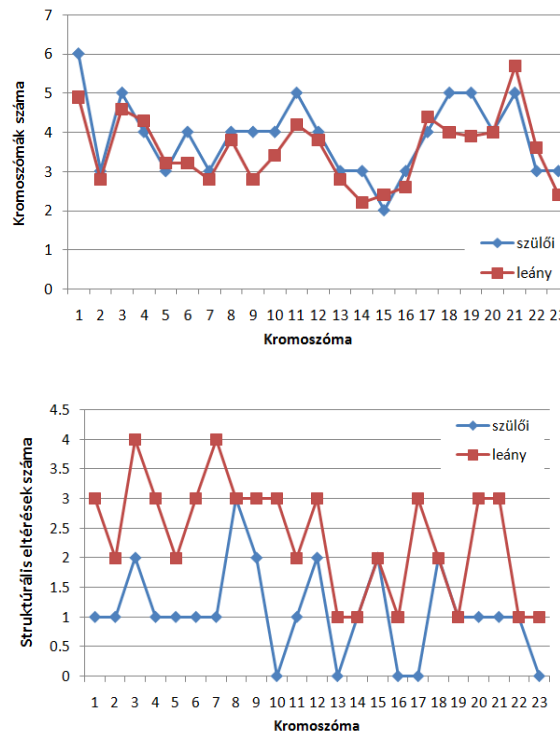
A citogenetikai elemzés azt mutatta, hogy az új sejtvonalak 60-110 kromoszómával rendelkeztek, amelyek közül az 1-es, 17-es és 21-es kromoszómák voltak a legtöbb esetben többszörözve. A G-sáv festés alapján a 3-as, 7-es, 17-es és a 21-es kromoszómák mutatták a legnagyobb változékonyságot. Leggyakrabban törölt a 9p21-es és a 18q21-es területek voltak. Az átlagos Shannon-féle diverzitási index 2,03 volt, a szülői vonal (MDA-MB-231) 2,05-öt ért el. A CIN pozitív vonalak átlagos rezisztenciája magasabb volt, mint a CIN negatív vonalaké, de a különbség nem volt statisztikailag szignifikáns.

RT-PCR próba	Gén	Gén neve	Sprmn	p- érték	Sprmn	p- érték
Doxorubicin rezisztenciával kapcsolatba hozott gének			vs. Doxorubicin rez. vonalak		vs. Paclitaxel rez. vonalak	
Hs00184500_m1	ABCB1	ATP-binding cassette, sub-family B (MDR), member 1	0.42	0.030	0.33	0.072
Hs00219905_m1	ABCC1	ATP-binding cassette, sub-family C (MRP), member 1	0.00	0.492	0.27	0.073
Hs00184979_m1	ABCG2	ATP-binding cassette, sub-family G, member 2	-0.25	0.085	0.01	0.470
Hs03063307_m1	TOP2A	topoisomerase (DNA) II alpha 170kDa	0.48	0.003	0.08	0.331
Hs00245438_m1	MVP	major vault protein	-0.26	0.077	0.42	0.009
Hs00946084_g1	HSPA5	heat shock 70kDa protein 5	0.20	0.142	0.37	0.019
Hs00160607_m1	PSMB7	proteasome unit beta-type 7	-0.14	0.221	-0.18	0.164
Paclitaxel rezisztenciával kapcsolatba hozott gének						
Hs00362387_m1	TUBA1A	tubulin, alpha 1a	-0.34	0.033	0.38	0.018
Hs00744842_sH	TUBA1B	tubulin, alpha 1b	-0.40	0.012	0.14	0.229
Hs00733770_m1	TUBA1C	tubulin, alpha 1c	-0.32	0.042	0.49	0.003
Hs00258236_m1	TUBB1	tubulin, beta 1	0.05	0.455	0.10	0.411
Hs00742533_s1	TUBB2A	tubulin, beta 2A	-0.35	0.026	0.42	0.009
Hs00603550_g1	TUBB2B	tubulin, beta 2B	-0.21	0.130	-0.21	0.127
Hs00607181_g1	TUBB2C	tubulin, beta 2C	-0.49	0.003	0.19	0.157
Hs00964962_g1	TUBB3	tubulin, beta 3	-0.45	0.006	0.42	0.010
Hs00893144_g1	TUBB4	tubulin, beta 4	-0.34	0.033	0.45	0.005
Hs00902188_m1	MAPT	microtubule-associated protein tau	0.28	0.066	-0.10	0.299
Hs00737065_m1	MAP4	microtubule-associated protein 4	-0.20	0.138	0.53	0.001
Emlőrák hormon receptorok						
Hs01046815_m1	ÖR	Ösztrogén receptor 1	0.11	0.273	-0.28	0.060
Hs01105519_m1	ÖR2	Ösztrogén receptor 2 (ER beta)	-0.01	0.489	-0.45	0.085
Hs00172183_m1	PGR	Progeszteron receptor	0.09	0.350	-0.13	0.297
Hs00170433_m1	ERBB2	ERBB2	-0.05	0.401	-0.26	0.075

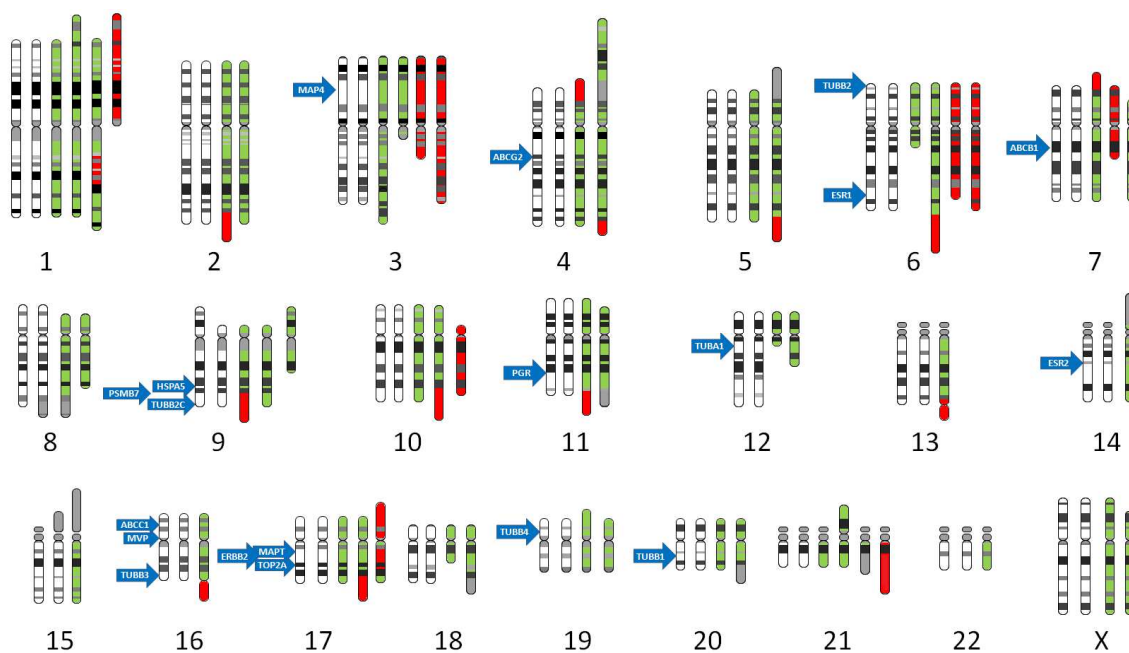
4. táblázat. Spearman rang korreláció az RT-PCR alapú gén expresszió és az MTT-vel mért IC50 értékek között a doxorubicin- és paclitaxel-rezisztens vonalakban. Félkövr: szignifikancia <0,01.

Sprmn: Spearman rang korrelációs koefficiens

dc_718_13



20. ábra. Az MDA-MB-231-es sejtvonalból származtatott vonalakban mért ploiditás (felső ábra) és az új strukturális eltérések száma (alsó ábra) kromoszómánként ábrázolva.



21. ábra. Citogenetikai eltérések az MDA-MB-231-es vonal leányvonaláiban. Fehér háttér: normális kromoszómakészlet; zöld háttér: a szülői sejtvonal kromoszómakészlete; piros háttér: a rezisztens vonalakban talált eltérések. Az RT-PCR-el mért gének elhelyezkedését nyilak jelölik a kromoszómákon.

4.6. Célzott terápiás szerekkel szembeni rezisztencia

A 45 sejtvonalban elvégzett rezisztencia-tesztek alapján az öt gyógyszerre sorba rendezett sejtvonalakat a következő táblázatban mutatom be.

Sejtvonal	Eredet	lapatinib		erlotinib		sorafenib		sunitinib		gefitinib	
CCRFCEM	ALL	SW948	0,26	NCI-H441	-0,11	NCI-H69	-1,53	A-375	-0,15	Hep-3b	-0,09
MOLT-4	ALL	K-562	0,43	HCT-15	-0,08	NCI-H441	-1,25	caOV3	0,07	HCT-15	0,06
K-562	Csontvelő	NCI-H358	0,48	caOV3	-0,02	A-375	-1,04	HCT-15	0,08	SW403	0,11
SK-N-AS	Agy	HCT-8	0,48	SW948	0,09	caOV3	-0,99	SW948	0,21	C-4II	0,18
BT-20	Emlő	HCT-15	0,49	A-375	0,22	Hep-3b	-0,91	NCI-H441	0,31	NCI-H358	0,22
MCF-7	Emlő	HT-1080	0,54	NCI-H1993	0,24	C-4I	-0,68	NCI-H358	0,33	NCI-H69	0,23
RAJI	Burkitt l.	NCI-H441	0,56	NCI-H1650	0,28	SW-620	-0,63	SW-620	0,35	WIDR	0,28
C-33A	Méhnyak	MCF-7	0,62	NCI-H358	0,30	HT-1080	-0,58	SW480	0,38	caOV3	0,29
C-4I	Méhnyak	caOV3	0,65	C-4II	0,31	HCT-15	-0,57	MCF-7	0,39	SNU-449	0,29
C-4II	Méhnyak	SNU-475	0,65	SW480	0,39	DMS 79	-0,53	HCT-8	0,41	HT-1080	0,29
HCT-15	Vastagb.	Hep-3b	0,66	NCI-H661	0,42	C-4II	-0,50	A427	0,41	A-375	0,30
HCT-8	Vastagb.	NCI-H1650	0,71	SNU-449	0,43	NCI-H1993	-0,45	Hep-3b	0,43	K-562	0,31
SW403	Vastagb.	A-375	0,73	C-4I	0,47	SW480	-0,41	C-4II	0,43	HCT-8	0,34
SW480	Vastagb.	HOS	0,73	K-562	0,48	SHP-77	-0,36	NCI-H1650	0,44	NCI-H661	0,36
SW-620	Vastagb.	BT-20	0,74	SNU-475	0,50	K-562	-0,34	HT-1080	0,51	MCF-7	0,41
SW948	Vastagb.	SHP-77	0,77	MCF-7	0,51	A427	-0,33	NCI-H82	0,51	C-4I	0,41
WIDR	Vastagb.	CCRF-CEM	0,78	Hep-3b	0,51	HCT-8	-0,31	ES-2	0,51	CCRFCEM	0,43
HT-1080	Fibrosarcoma	SNU-449	0,78	CCRF-CEM	0,55	NCI-H358	-0,30	CCRF-CEM	0,51	SNU-475	0,47
A-498	Vese	C-4II	0,78	NCI-H69	0,55	SW403	-0,28	SNU-475	0,51	SW-620	0,47
Hep-3b	Máj	NCI-H1993	0,82	SW403	0,55	SNU-475	-0,28	WIDR	0,53	NCI-H82	0,47
SNU-182	Máj	SW480	0,82	HOS	0,58	COLO-668	-0,25	NCI-H1993	0,54	SW480	0,48
SNU-423	Máj	C-33A	0,87	HT-1080	0,58	CCRF-CEM	-0,19	NCI-H69	0,56	ES-2	0,51
SNU-449	Máj	WIDR	0,88	C-33A	0,59	MOLT-4	-0,18	HOS	0,59	NCI-H1993	0,51
SNU-475	Máj	AN3_CA	0,89	AN3_CA	0,60	MCF-7	-0,16	SNU-449	0,60	NCI-H441	0,51
A427	Tüdő	SNU-423	0,90	ES-2	0,61	NCI-H1975	-0,14	C-4I	0,61	C-33A	0,53
COLO668	Tüdő	A-498	0,91	SW-620	0,61	SW948	-0,12	MOLT-4	0,61	DMS 79	0,54
DMS 114	Tüdő	NCI-H661	0,95	BT-20	0,67	SNU-423	-0,12	K-562	0,63	SNU-182	0,57
DMS 79	Tüdő	SW403	0,96	NCI-H82	0,68	SNU-449	-0,11	AN3_CA	0,63	SHP-77	0,64
NCI-H358	Tüdő	C-4I	0,98	WIDR	0,70	C-33A	-0,11	SNU-182	0,66	MOLT-4	0,65
NCI-H441	Tüdő	DMS 114	0,98	MOLT-4	0,73	WIDR	-0,11	BT-20	0,66	RAJI	0,65
NCI-H661	Tüdő	RAJI	0,99	A-498	0,73	NCI-H82	-0,09	COLO-668	0,71	NCI-H1650	0,67
NCI-H69	Tüdő	SNU-182	1,01	HEC-1-B	0,73	SNU-182	-0,08	A-498	0,72	A427	0,69

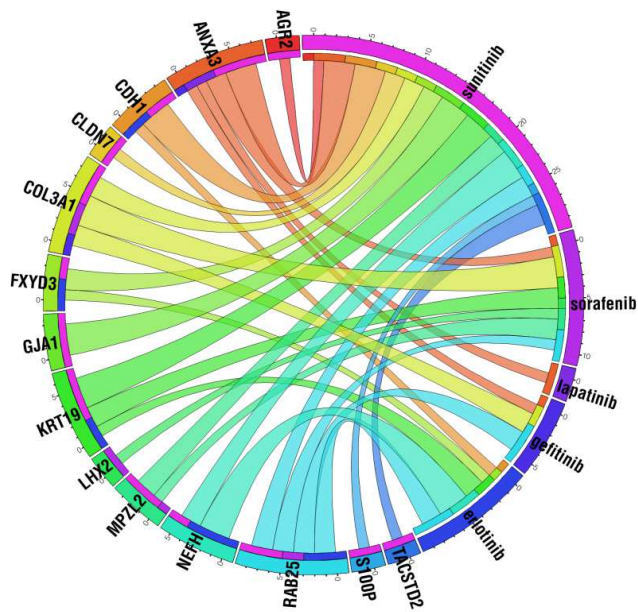
dc_718_13

NCI-H82	Tüdő	ChaGo-K-1	1,01	SNU-423	0,81	ChaGo-K-1	-0,07	SK-N-AS	0,72	BT-20	0,69
SHP-77	Tüdő	NCI-H82	1,02	NCI-H1975	0,82	RAJI	-0,05	RAJI	0,72	AN3_CA	0,71
ChaGOK1	Tüdő	MOLT-4	1,03	DMS 79	0,82	BT-20	-0,01	NCI-H1975	0,73	SNU-423	0,73
NCIH1650	Tüdő	NCI-H69	1,03	RAJI	0,82	DMS 114	0,00	ChaGo-K-1	0,78	NCIH1975	0,74
NCIH1975	Tüdő	ES-2	1,05	ChaGo-K-1	0,83	HOS	0,01	SNU-423	0,80	A-498	0,80
NCIH1993	Tüdő	NCI-H1975	1,08	DMS 114	0,89	NCI-H1650	0,14	NCI-H661	0,88	ChaGo-K1	0,81
RD	Izom	SK-N-AS	1,10	SNU-182	0,90	RD	0,20	DMS 79	0,89	HOS	0,87
CaOV3	Petef.	A427	1,10	COLO-668	0,91	SK-N-AS	0,20	SHP-77	0,90	DMS 114	0,93
ES-2	Petef.	RD	1,12	HCT-8	0,95	ES-2	0,21	C-33A	0,91	SK-N-AS	1,04
HOS	Szarkóma	COLO-668	1,13	A427	0,98	HEC-1-B	0,31	DMS 114	0,92	HEC-1-B	1,07
A375	Bőr	DMS 79	1,20	SK-N-AS	1,00	NCI-H661	0,32	RD	0,96	RD	1,28
AN3_CA	Uterus	HEC-1-B	1,31	RD	1,01	A-498	0,49	HEC-1-B	0,98	COLO668	1,34
HEC-1-B	Uterus	SW-620	1,31	SHP-77	1,04	AN3_CA	0,61	SW403	1,08	SW948	1,45

5. táblázat. A 45 kiválasztott sejtvonal az öt gyógyszerrel szembeni rezisztencia mértéke alapján sorba rendezve. A bekeretezett középső vonalakat osztályoztuk közepesként, és az alsó (rezisztens) és felső (érzékeny) csoportba sorolt vonalakat hasonlítottuk egymással össze.

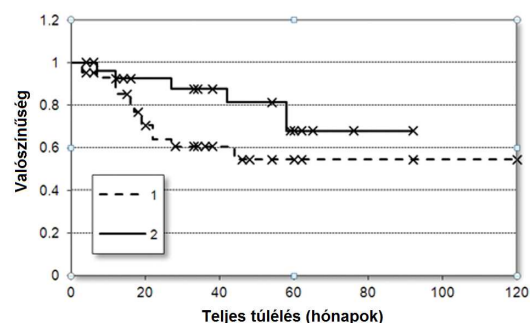
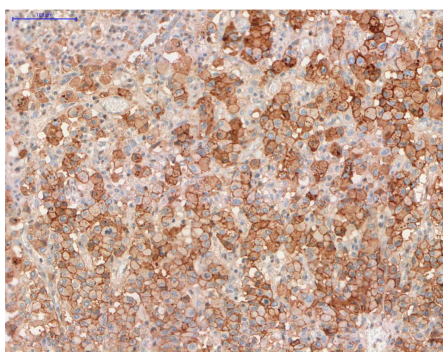
A fentebb ismertetett módon SAM-mel és rank products-al meghatározott géneket PAM-mel teszteltük a sejtvonalakban. Ennek során a sejtvonalak 92,8%-át helyesen tudtuk osztályozni az összes gén felhasználásával és a sejtvonalak 79%-át a legjobb 100 gén felhasználásával.

A legerősebb 63 gén kifejeződésében mutatott különbséget ezután RT-PCR felhasználásával teszteltük, amelynek eredményeképpen 45 gént igazoltunk $p < 0,05$ -ös szignifikancia mellett. A legszignifikánsabb gén erlotinib esetében az ITGB4, gefitinib esetében az IADA, sorafenib esetében a FAT4, lapatinib esetében a FURIN és az ME1, sunitinib esetében pedig a KRT18, az LGALS8, a RAB17 és a CD9 voltak. Néhány gén több gyógyszerrel szembeni rezisztenciával is összefüggést mutatott (ami nem meglepő, ha figyelembe vesszük a közös jelátviteli utak felépítését) - ilyen volt a COL3A1, az ANXA3, a RAB25 és a KRT19. Ezen több gyógyszerrel is összefüggő géneket az alábbi kördiagrammon ("circos" ábrán) mutatjuk be.



22. ábra. A többszörös rezisztenciával kapcsolatba hozott gének (bal oldal) és az egyes gyógyszerek (jobb oldal) kapcsolata. A szalagok vastagsága a korreláció mértékével arányos. A legtöbb gént a sunitinib-rezisztenciával hoztuk kapcsolatba, míg a legkevesebb kapcsolatot a lapatinib esetében találtuk.

A sejtvonalakban elvégzett igazolás után a legjobb géneket veserákos mintákban is megvizsgáltuk immunhisztokémiával. Összesen 39 betegtől gyűjtöttünk mintát, a medián teljes túlélés 14 hónap volt (az átlagos túlélés 20 hónap). A tumorokból készült TMA-kban az LGALS8, a RAB17, és az EPCAM festődési intenzitása illetve pozitív sejtek aránya szignifikánsan összefüggött a jobb túléléssel, de a CD9 nem mutatott összefüggést.



23. ábra. Az EpCAM immunhisztokémiai festése (bal oldal) alapján két csoportra osztott veserákos betegek túlélése szignifikánsan jobb volt (jobb oldal, Kaplan-Meier görbe, $p=0,01$).

4.7. A melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása

Összesen kilenc olyan vizsgálatot azonosítottunk, amelyek valamennyi feltételnek megfeleltek²⁰²⁻²¹⁰. A közlemények egymástól jelentősen eltértek, egy cikk csak sejtvonalakat használt, öt közlemény sejtvonalakat és tumormintákat, kettő közlemény csak tumormintákat, egy cikk pedig vérminták felhasználásával határozott meg génlistát.

A kilenc közleményben a melanoma pathogenezisével összefüggésben összesen 815 különböző gént azonosítottak. Ezek közül 43 gént több vizsgálat is leírt. Összesen öt gént (ERBB3, ADRB2, MERTK, SNF1LK és ITPKB) írtak le három vizsgálatban és egy gént (RAB33A) négy vizsgálatban.

A gén-ontológiai elemzés eredményeképpen a leginkább felülreprezentált kategóriák a membránon átívelő receptor tirozin kináz útvonalak ($p=8,4E-10$), a szervfejlődés ($p=9,4E-09$), a morfogenezis ($p=1,7E-08$), a szempigmentáció bioszintézise ($p=2,2E-08$) és a vírusokra adott válasz ($p=8,8E-08$) voltak.

4.8. Az emlőrák prognózisával kapcsolatba hozott génlisták kereszt-elemzése nyers adatok újrafeldolgozása után

Hét olyan közleményt azonosítottunk, ahol a nyers gén chip adatok is elérhetőek voltak, ezekben a közleményekben összesen 1470 betegből készült gén chipet közöltek le. Mivel a klinikai adatok alapján feltűnt az egyes vizsgálatok közötti nagyfokú hasonlóság, ezért az összes mintát sorba rendeztük az átlagos MAS5 normalizált génexpressziók alapján. Azoknál a mintáknál, ahol az átlag egyezett, az egyes gének szintjén is ellenőriztük, hogy valóban ugyanarról a mintáról van-e szó. Valamennyi olyan gén chip, ahol az átlag egyezett, az egyes gének szintjén is azonos volt. Összesen 640 redundáns gén chipet azonosítottunk, amelyek 251 eredeti többszörös közlését tartalmazták, azonban az egyes vizsgálatokban más és más klinikai adattal látták el az adott betegeket. Saját elemzésünk során ezeket a klinikai adatokat összegyűjtöttük, és ezután egy-egy betegre valamennyi adatot elérhetővé tettünk a további elemzésekhez.

A végső adatbázis az összesen 1470 gén chipből csak az 1079 egyedi gén chipet tartalmazta. A redundáns minták eltávolításával a klinikai adatok is jelentősen megváltoztak:

az átlagos kiújulás mentes idő két évvel hosszabb lett, a nyirokcsomó pozitív betegek aránya kevesebb mint a felére esett. A végső adatbázisban a minták 77%-a volt ösztrogén receptor pozitív, 13%-a nyirokcsomó-pozitív, az átlagos kiújulás-mentes idő 9.0 év volt, a betegek átlagos kora 56 ± 13 év, a tumorméret pedig $2,3 \pm 1,2$ mm volt.

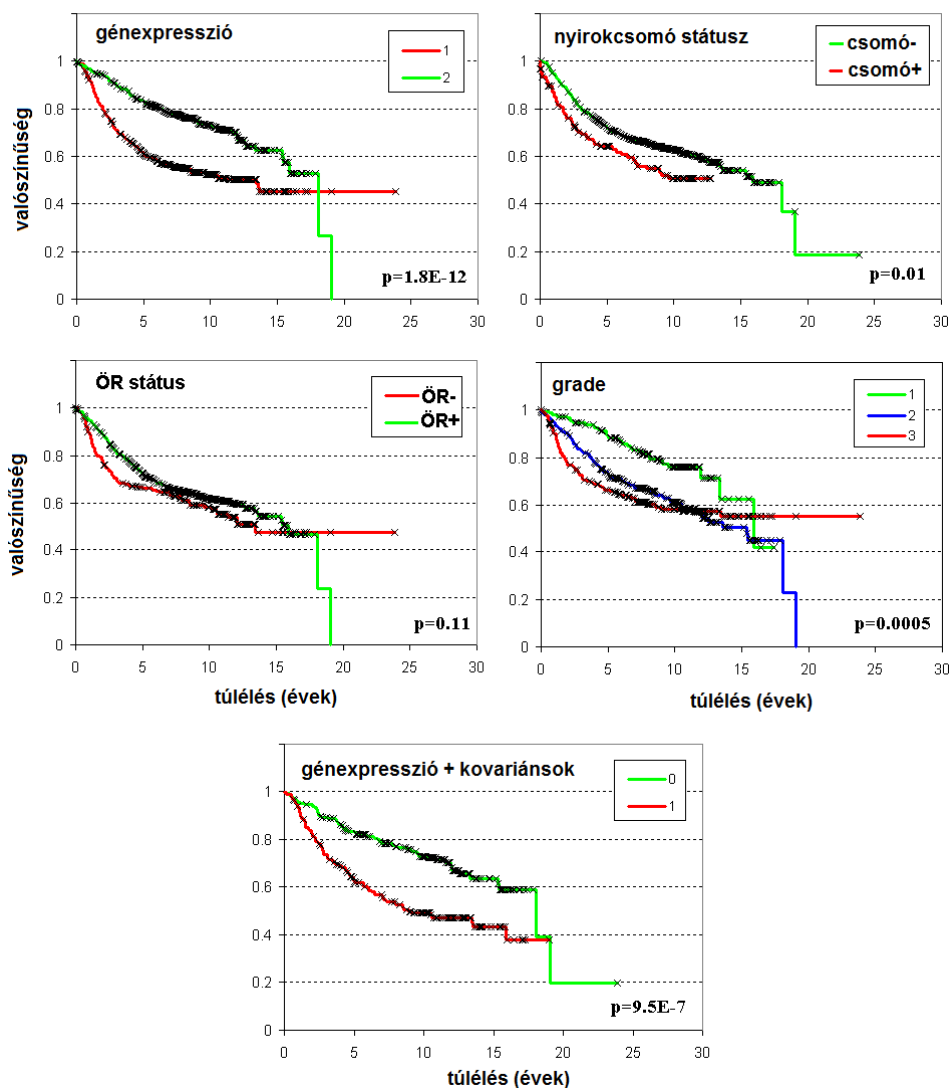
Eredeti közlemény					Elemzés	
Betegek száma	Technológia	A közölt gének száma	Osztályozási probléma	Referencia	A gének száma a GPL96 platformon	Cox p érték
89	Affymetrix HGU95Av2	56 metagén	Kiújulás előrejelzése	211	74 próba	0.017
159	Affymetrix HGU133A	64 gén	Jó és rossz prognózisú betegek elkülönítése	212	72 próba	0.00081 (GSE1456 nélkül*)
286	Affymetrix HGU133A	76 próba	Távoli áttét előrejelzése nyirokcsomó-negatív betegekben	213	76 próba	0.00001 (GSE2034 nélkül*)
244	<i>in silico</i>	7 gén	Prognózis előrejelzése nyirokcsomó-negatív betegekben	214	19 próba	0.033
162	cDNA chip 10,368 próbával	21 gén	Klinikai kimenet előrejelzése	215	30 próba	0.018
135	Agilent human 1A oligo chip	70 gén	Prognosztikus mintázat	216	91 próba	0.00071
65	Agilent teljes genom gén chipek	822 gén	Ösztrogén-szabályozott génekkel kimenet előrejelzése	217	331 próba	0.044
249	Affymetrix HGU133A+B	18 próba	Grade 1 és grade 3 megkülönböztetése	218	13 próba	0.00007

6. táblázat. Az 1079 egyéni gén chip mintában a tesztelt 20 gén-mintázat közül nyolc ért el statisztikai szignifikanciát. *A datasetek itt az eredeti közleményekben is felhasználásra kerültek, ezért a tanuló és a tesztalmaz átfedésének elkerülésére ezeket kizártuk az elemzésből.

A korábban közölt vizsgálatok között összesen 20 különböző génlistát azonosítottunk, amelyek közül a fenti táblán bemutatott nyolc génlista ért el statisztikai szignifikanciát $p < 0,05$ -ös küszöb mellett. A génlisták egyes génjeire elvégzett számítások alapján a

leggyengébben teljesítő génlisták is tartalmaztak olyan géneket, amelyek önmagukban képesek voltak a jó és rossz prognózisú betegeket elkülöníteni.

A gén-mintázatok mellett a klinikai paraméterek előrejelző értékét is kiszámoltuk, valamint a génexpressziós adatok alapján egy olyan új osztályozót is fejlesztettünk, amely az összes 1079 beteg adatai alapján készült. Amennyiben a génexpressziós adatokat a klinikai adatokkal kombináltuk, úgy 44 gén felhasználásával már magas szignifikanciát tudtunk elérni (lásd következő ábra).



24. ábra. Kaplan-Meier görbék a legjobb 376 gén felhasználásával, a nyirokcsomó és ÖR- státusz, valamint grade alapján. A 44 gén expresszióját klinikai változókkal (nyirokcsomó-státusz és grade) kombinálva kevesebb génnel is nagy osztályozási hatásosságot lehet elérni.

4.9. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének igazolása klinikai mintákban

A felépített petefészekrák adatbázis összesen 829 gén chip mintát tartalmazott, amelyek közül 806 petefészekrák, 23 pedig normál petefészek eredetű volt. A betegek közül 199-hez rendelkezünk túlélési adattal is. Elsőként 38 korábban közölt génlistát vizsgáltunk meg ebben az adatbázisban (a génlisták azonosítását egy másik közleményünkben mutattuk be ²¹⁹). Ezen mintázatok közül csak azokat vizsgáltuk, amelyek esetében a közölt gének legalább 50%-át az Affymetrix gén chippek is meghatározták ^{97, 99, 101, 103, 107-109, 111, 220-224}. Az adatbázisban megnéztük, hogy a génlisták mennyire képesek az egészséges és a rosszindulatú daganatos szövetek, illetve a szövettani altípusok egymástól való elkülönítésére *(ezeket lásd a következő táblázatban)*. Azt is megnéztük, hogy a túlélés előrejelzésére mennyire képesek ezek a mintázatok, azonban egyetlen lista sem volt statisztikailag szignifikáns.

A vizsgálat második részében a teljes adattábla felhasználásával a szövettani altípusok elkülönítésére és a túlélés előrejelzésére alkalmas géneket azonosítottuk, majd ezeket a géneket teszteltük az általunk gyűjtött betegekben.

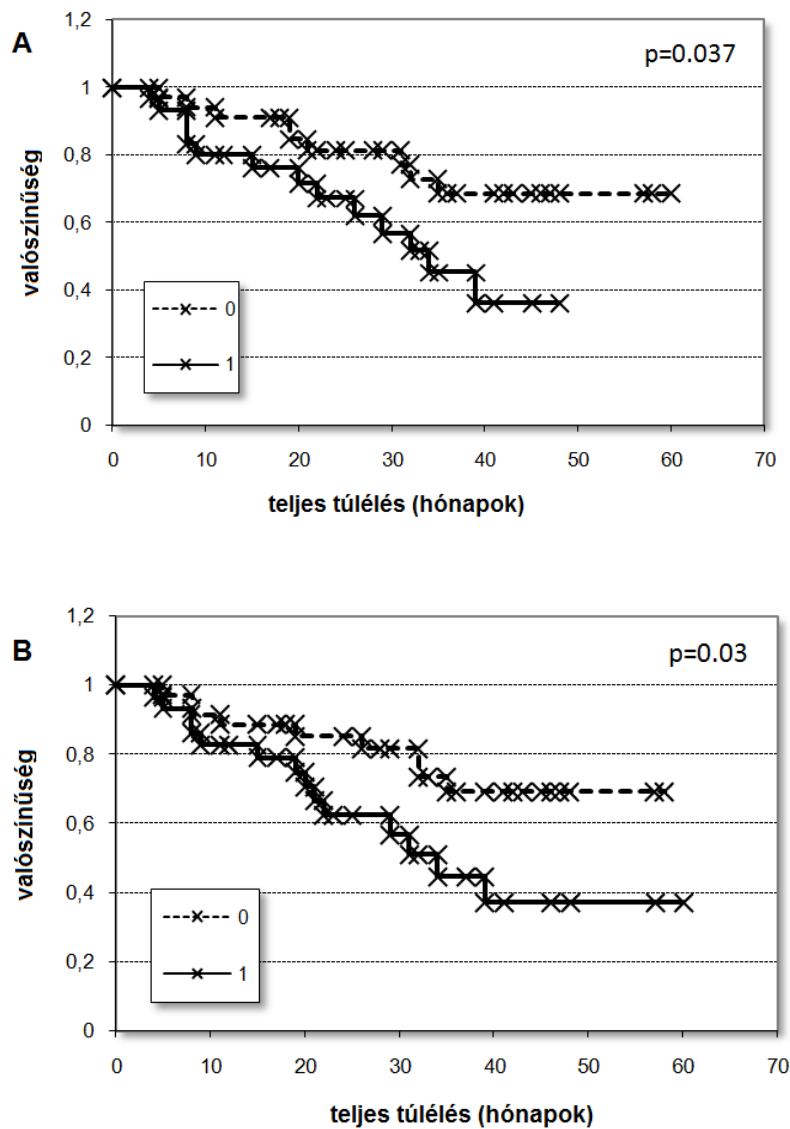
A klinikai mintagyűjtés során 64 petefészekrákos betegtől gyűjtöttünk tumormintát. A betegek medián kiújulás mentes túlélése 24,5, teljes túlélése pedig 29 hónap volt, négy betegben másodlagos emlő karcinóma is kialakult. Az ezen mintákból izolált RNS-ek felhasználásával három vizsgálatot végeztünk: a túléléssel, a szövettani altípusokkal, és az emlőrák kialakulásával kapcsolatba hozott géneket egymástól függetlenül vizsgáltuk. Az egyes szövettani altípusokon belül az alacsony mintaszám miatt csak a magas grade-ű szerózus tumorokat hasonlítottuk össze a többi altípus összevont csoportjával. A szignifikáns géneket a 8. táblázat tartalmazza. A túléléssel összefüggő gének alapján Kaplan-Meier túlélési görbéket is készítettünk, amelynek során a génexpresszió mediánja alapján osztottuk a betegeket két csoportra és ezeket hasonlítottuk egymással össze.

	Első szerző, év, referencia	Gének száma	p-érték
A, rákos és egészséges szövetek elkülönítésére képes			
	Bignotti et al, 2006 ⁹⁹	116	< 0.0001
	Donninger et al, 2004 ¹⁰¹	659	< 0.0001
	Fedorowicz et al, 2009 ²²⁰	28	< 0.0001
	Heinzelmann et al, 2006 ²²²	20	0.0006
	Warrenfeltz et al, 2004 ²²⁴	127	< 0.0001
	Welsh et al, 2001 ⁹⁷	17	< 0.0001
	Grisaru et al, 2007 ²²¹	68	0.0042
	Quinn et al, 2009 ²²³	71	0.0014
	Santin et al, 2004 ¹⁰⁹	4	0.195
	Zhang et al, 2007 ¹¹¹	7	0.071
	Klinck et al, 2008 ¹⁰³	37	0.011
	Park et al, 2008 ¹⁰⁸	26	0.234
B, szövettani altípusok megkülönböztetésére képes			
	Bignotti et al, 2006 ⁹⁹	116	< 0.0001
	Donninger et al, 2004 ¹⁰¹	659	< 0.0001
	Heinzelmann et al, 2006 ²²²	20	0.0007
	Welsh et al, 2001 ⁹⁷	17	0.0009
	Quinn et al, 2009 ²²³	71	0.0023
	Warrenfeltz et al, 2004 ²²⁴	127	< 0.0001
	Santin et al, 2004 ¹⁰⁹	4	0.0099
	Mougeot et al, 2006 ¹⁰⁷	53	0.0007
	Fedorowicz et al, 2009 ²²⁰	28	0.383

7. táblázat. Korábban közölt génlistákat használtunk osztályozóként 829 petefészekrákos beteg tumorán végzett gén chip adatokon. A rosszindulatú daganatos és egészséges szöveteket elkülönítő listák (A) és a szövettani altípusok megkülönböztetésére képes génlisták (B).

Azonosító	Gén szimbólum	Gén neve	SAM pont	q érték	Hol lett a gén eredetileg azonosítva
Teljes túlélés minden betegben					
Hs00172183_m1	PGR	progesteron receptor	1.62	<0.01	Hormonreceptor
Hs01105519_m1	ÖR2	Ösztrogén receptor 2 (ÖR beta)	1.55	<0.01	Hormonreceptor
Hs00610327_m1	TSPAN8	tetraspanin 8	1.54	<0.01	Altípus
Kiújulás mentes túlélés					
Hs00902188_m1	MAPT	microtubule-associated protein tau	-1.61	<0.01	Kemoterápiás válasz
Hs00268306_m1	SNCG	synuclein, gamma (breast cancer-specific protein 1)	-1.67	<0.01	Emlőrák specifikus gének
Magas grade szerózus az összes többi mintával szemben					
Hs00266715_s1	GAS1	growth arrest-specific 1	2.35	<0.01	Altípus
Hs01103751_m1	WT1	Wilms tumor 1	2.86	<0.01	Altípus
Hs00245879_m1	MSLN	mesothelin	1.74	<0.01	Altípus
Hs00418568_m1	NPR1	natriuretic peptide receptor A/guanylatecyclase A	2.37	<0.01	Altípus
Hs00610327_m1	TSPAN8	tetraspanin 8	-3.71	<0.01	Altípus
Hs00181323_m1	GAS6	growth arrest-specific 6	0.94	2.53	Altípus
Hs00191351_m1	ARHGAP29	Rho GTPase activating protein 29	1.38	<0.01	Altípus
Hs01065189_m1	MUC16	mucin 16, cell surface associated	1.73	<0.01	Altípus
Hs00170299_m1	ZYX	ESP-2, HED-2	1.53	<0.01	Altípus
Hs00188109_m1	MYO9B	myosin IXB	1.68	<0.01	Altípus
Hs00256958_m1	PHF1	PHD finger protein 1	0.86	2.53	Altípus
Hs00274988_m1	HDGFRP3	hepatoma-derived growth factor, related protein 3	0.94	2.53	Altípus
Hs00268306_m1	SNCG	synuclein, gamma (breast cancer-specific protein 1)	2.49	<0.01	Emlőrák specifikus gének
Hs01046815_m1	ÖR	ösztrogén receptor 1	0.87	2.53	Hormonreceptor
Hs00160607_m1	PSMB7	proteasome (prosome, macropain) subunit, beta type, 7	0.89	2.53	Kemoterápiás válasz
Hs00258236_m1	TUBB1	tubulin, beta 1	1.41	<0.01	Kemoterápiás válasz
Hs00362387_m1	TUBA1A	tubulin alpha 1a	0.96	2.53	Kemoterápiás válasz
Hs00737065_m1	MAP4	microtubule-associated protein 4	1.62	<0.01	Kemoterápiás válasz
Hs00742533_s1	TUBB2A	tubulin, beta 2A	1.05	2.53	Kemoterápiás válasz
Hs00744842_sH	TUBA1B	tubulin, alpha 1b	1.46	<0.01	Kemoterápiás válasz
Hs00893144_g1	TUBB4	tubulin, beta 4	1.13	2.53	Kemoterápiás válasz
Hs00902188_m1	MAPT	microtubule-associated protein tau	0.98	2.53	Kemoterápiás válasz

8. táblázat. A gén chip adatok elemzésével azonosított gének független tesztelése 64 általunk gyűjtött petefészekrákos mintán elvégzett RT-PCR mérés során.



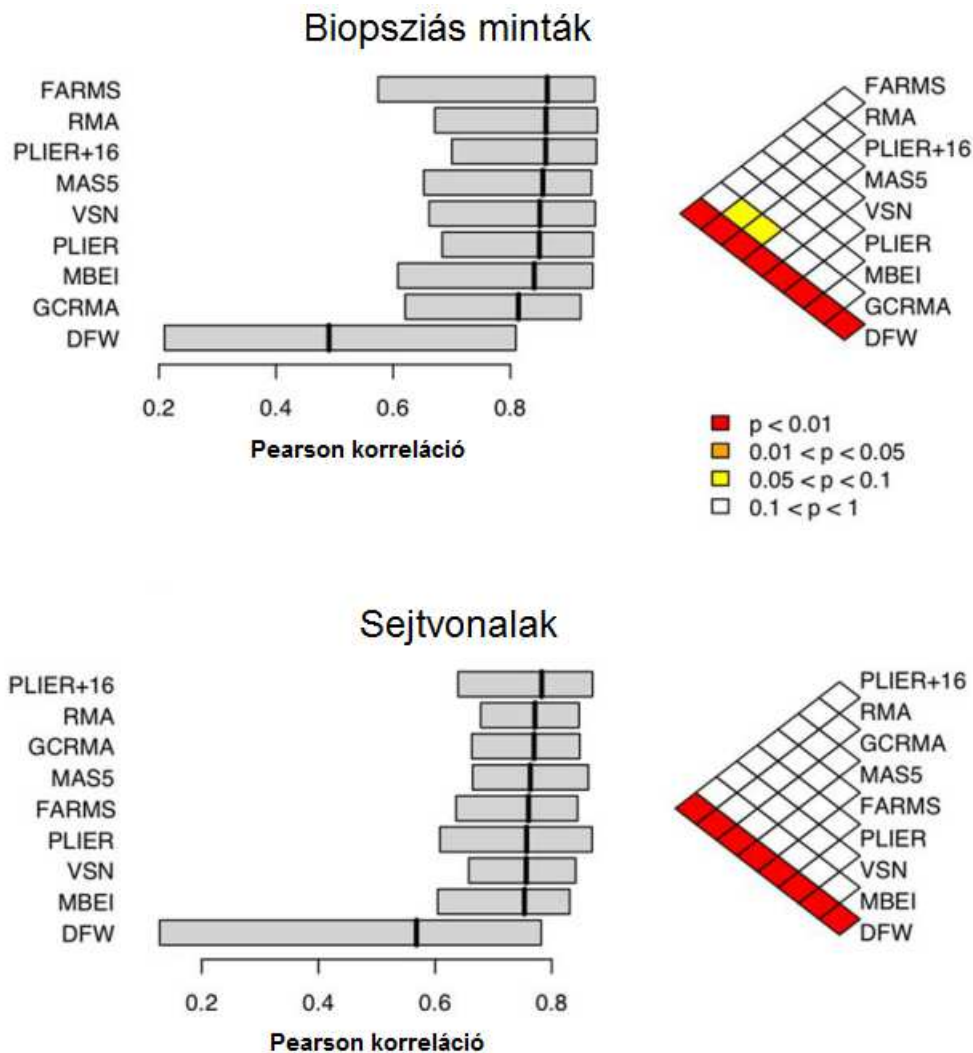
25. ábra. Kaplan-Meier elemzés az ÖR2 (A) és a PGR (B) génekre 64 petefészekrákos betegben. A gén expressziós értékeinek mediánja alapján osztottuk két csoportra a betegeket. (1: magas, 0: alacsony expressziót mutató minták)

4.10. Microarray adatok előfeldolgozása

A statisztikai számítások során valamennyi gént különálló mérésként kezeltünk. A $\log_2$ -es értékek között Pearson korrelációt számoltunk, majd ez egyes algoritmusok által

dc_718_13

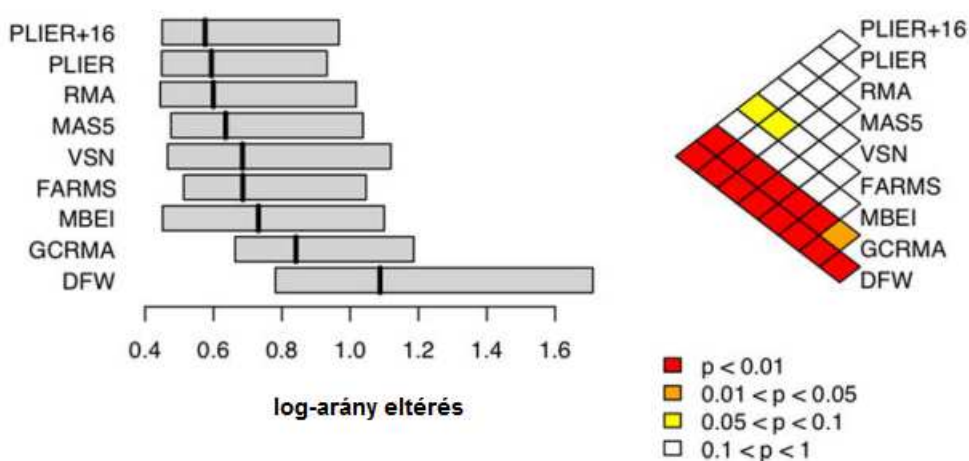
számított értékeket egymással Wilcoxon teszt segítségével hasonlítottuk össze. A medián Pearson értékek alapján a FARMS algoritmus volt a legjobb a biopsziás mintákban, a sejtvonalakban a PLIER+16. A legfontosabb algoritmusok között a különbség nem volt szignifikáns. A legrosszabb teljesítményt a DFW és a GCRMA algoritmusok adták. Az eredményeket az alábbi ábrán foglaltam össze.



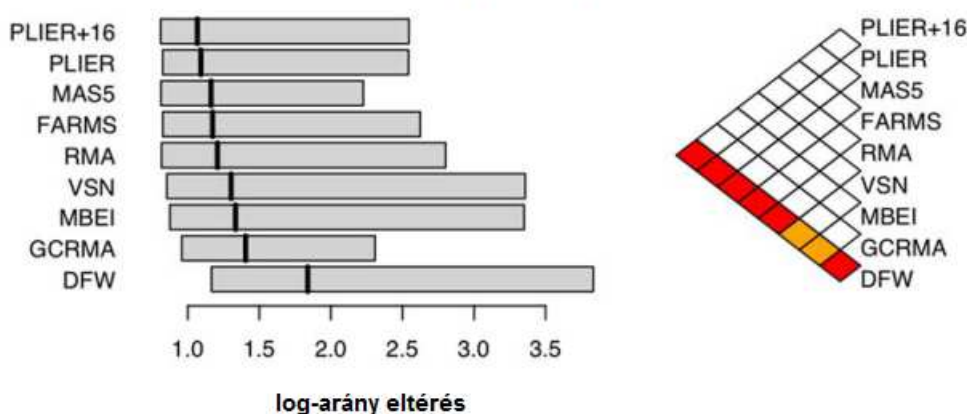
26. ábra. Kilenc előfeldolgozási algoritmus összehasonlítása vastagbélből vett biopsziás mintákban és sejtvonalakban. Az ábra bal oldala a gén chipen mért és az RT-PCR-el meghatározott, $\log_2$ -es génexpressziós értékek közötti korrelációt mutatja (nagyobb érték = jobb korreláció), az ábra jobb oldala az egyes algoritmusok egymáshoz viszonyított korrelációját. Az algoritmusokat hatásosság alapján sorba rendeztük, a legfelső a legjobb, a legalsó a legrosszabb.

A Pearson korreláció során az egyes értékeket egymással hasonlítottuk össze. Emellett azonban azt is célszerű meghatározni, hogy a két mérés mennyire egyezik egy adott gén esetén, másként fogalmazva, a mérések aránya mennyire tér el az egyes algoritmusok esetében. Erre az úgynevezett log-arány-eltérést („log-ratio discrepancy”, LRD) számítottuk ki ¹⁹³, majd a lenti ábrán az LRD eltéréseket a Pearson korrelációhoz hasonló módon ábrázoltuk, illetve az egyes algoritmusokat itt Wilcoxon teszt segítségével hasonlítottuk egymással össze. Az egyes algoritmusok között az eltérések így jelentősebbek lettek, azonban a legjobb eredményt itt is a PLIER+16, valamint a legrosszabb eredményt a DFW és a GCRMA algoritmusok adták.

Biopsziás minták



Sejtvonalak



27. ábra. Kilenc előfeldolgozási algoritmus összehasonlítása vastagbélből vett biopsziás mintákban és sejtvonalakban. Az ábra bal oldala a gén chipen mért és az RT-PCR-el meghatározott kifejeződés log-arány különbségét mutatja (kisebb érték=kisebb eltérés vagyis jobb korreláció), az ábra jobb oldala az egyes algoritmusok egymáshoz viszonyított korrelációját. Az algoritmusokat hatásosság alapján sorba rendeztük, a legfelső a legjobb, a legalsó a legrosszabb.

4.11. Bioinformatikai fejlesztés génexpresszió alapú biomarkerek tesztelésére

Létrehoztunk egy, a világhálón keresztül elérhető rendszert, amellyel egyes gének expressziójának túléléssel való összefüggését lehet vizsgálni. A rendszert 2011-ben kibővítettük, így a betegszám megnőtt 2344-re, majd 2012-ben megint bővítettük, akkor a betegszám 2978-ra nőtt. A rendszer fenntartása során további opciókkal egészítettük ki a szűrőket, mint például a korábbi verziók használata, a minőségellenőrzés, a részletes kezelési opciók alkalmazása. Az alábbi képen a honlap elemzőfelületének aktuális verziójáról készült képernyőkép látható.

The screenshot shows the Kaplan-Meier Plotter web interface. At the top, it says 'Kaplan-Meier Plotter' and 'Breast Cancer'. There are navigation links: 'KM plotter', 'Home', 'Download', 'Updates', and 'Contact'. A 'Switch to Ovarian Cancer' link is also present. The main form includes fields for 'Affy id/Gene symbols' (with a 'Use multigene classifier' button), 'Survival' (set to 'RFS (n=2880)'), 'Split patients by' (set to 'median'), 'Follow up threshold' (set to 'all'), and 'Compute median over entire dataset' (checked). There are 'Settings' and 'Draw Kaplan-Meier plot' buttons. A 'Probe set options' section has radio buttons for 'user selected probe set', 'all probe sets per gene', and 'only JetSet best probe set'. A 'Restrict analysis to subtypes...' section lists various clinical parameters like ER status, PR status, Lymph node status, Grade, and Molecular subtype, each with a dropdown menu. A 'Restrict analysis to selected cohorts...' section has radio buttons for 'systemically untreated patients', 'systemically treated patients', and 'patient cohort similar to SEER prevalences'. A 'Draw Kaplan-Meier plot' button is at the bottom. A green box highlights the number '2878' and the text 'patients'.

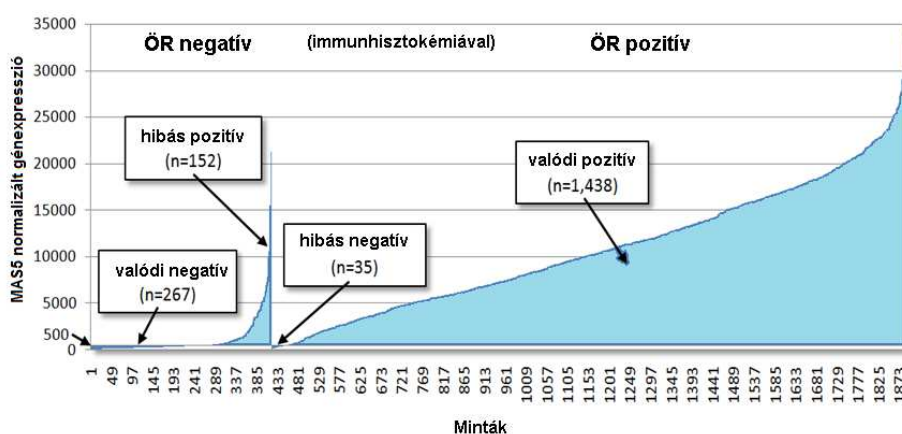
28. ábra. Az elemző rendszer kezelőfelülete. A legegyszerűbb elemzéshez elég egyetlen mezőt kitölteni (a vizsgálandó gén nevét), utána elindítani a számítást. Altípusokra és kezelési csoportokra alkalmazott szűrőkkel lehet részletesebb elemzéseket végezni. A rendszer folyamatosan mutatja, hogy az elemzésbe hány beteg kerül bele.

A rendszer működésének bemutatását eredetileg az ASCO által javasolt proliferációs biomarkerek vizsgálatával végeztük el. A megközelítés lehetőségeiből adódóan nem szükséges, hogy az általunk eredetileg elvégzett elemzés²²⁵ adataira szorítkozzunk, hiszen az adott géneket az adatbázis újabb verzióival is megvizsgálhatjuk. Az ezen frissített adatok alapján készített új táblázatot az alábbiakban beillesztettem. Az elemzés során a kiújulás mentes túléléssel való összefüggést vizsgáltuk az elérhető legjobb vágópont mellett. A táblázatban csak az általunk korábban a JetSet algoritmussal²²⁶ azonosított legjobb próbákat tüntettem fel.

Gén	Gén neve	Affy azonosító	HR	p érték
MKI67	A Ki-67 antitest által azonosított antigén	212021_s_at	1,51	1,5E-09
CCND1	Cyclin D1	208712_at	0,9	0,13
CCND2	Cyclin D2	200953_s_at	0,57	<1E-16
CCNE1	Cyclin E1	213523_at	1,55	6,5E-12
CCNE2	Cyclin E2	205034_at	1,95	<1E-16
CDKN1B	Cyclin-függő kináz 1B	209112_at	0,82	0,0085
CDKN1A	Cyclin-függő kináz 1A	202284_s_at	0,78	0,00027
TK1	Tymidin kináz 1	202338_at	1,37	8,5E-07
TK2	Tymidin kináz 2	204276_at	0,7	9E-08
TOP2A	Topoizomeráz II alfa	201292_s_at	2,11	<1E-16
TOP2B	Topoizomeráz II béta	211987_at	1,29	0,00019

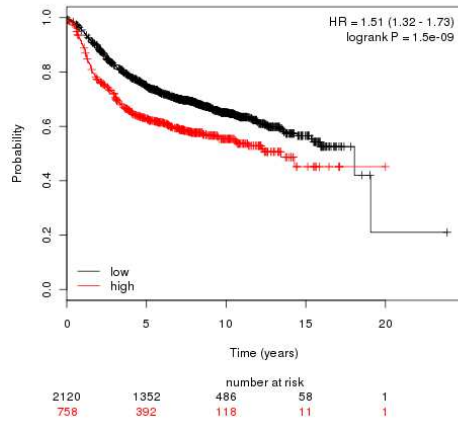
9. táblázat. Az ASCO által vizsgált proliferációs markerek teljesítménye az elemző rendszer által vizsgált 2978 betege tartalmazó adatbázisban kiújulás mentes túlélés előrejelzésére

Figyelembe véve, hogy a gén chip adatokkal a hormonreceptor-státusz független meghatározása is lehetséges, ezért a rendelkezésre álló adatok alapján összehasonlítottuk egymással a két platform által adott előrejelzést. Már korábbi vizsgálatok is igazolták az immunhisztokémiával meghatározott fehérjeszint és az mRNA szint közötti pozitív összefüggést ^{227, 228}. A jelen vizsgálat keretein belül a gén chip adatok felhasználásával elvégzett összehasonlítás alapján az mRNA alapú hormonreceptor-státusz mérése 90% feletti pontosságot mutatott (lásd következő ábra). Meg kell jegyeznünk, hogy a kétszeres normalizálás alkalmazása miatt a korábbi 500-as vágóponthoz ¹⁹⁵ viszonyított eltérés az egyes betegek esetében megváltozhatott. Azonban ez a kettős normalizálás tette lehetővé, hogy különböző platformokat egymással is kombináljunk, hogy ezáltal a végső betegszámot jelentősen növelni tudjuk.

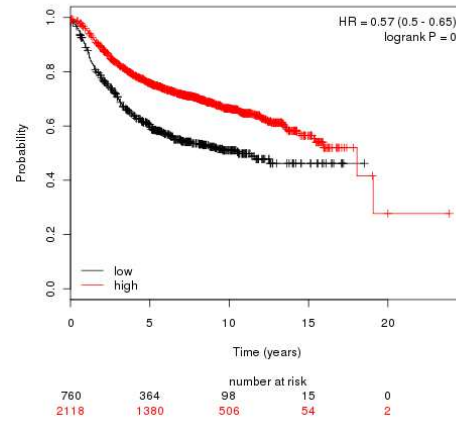


29. ábra. A gén chip alapú és az immunhisztokémia alapú ösztrogén státusz kapcsolata 1892 beteg adatai alapján. A vágópontként alkalmazott 500-as értéket az Y tengelyen nyíl jelöli.

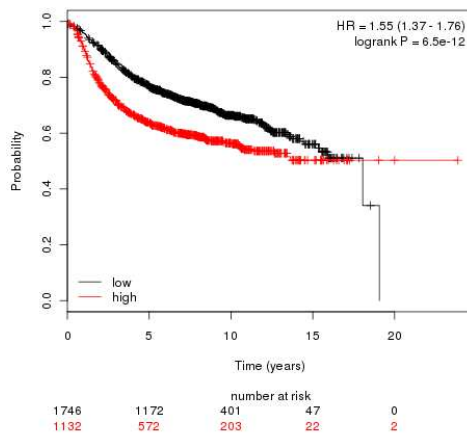
MKI67



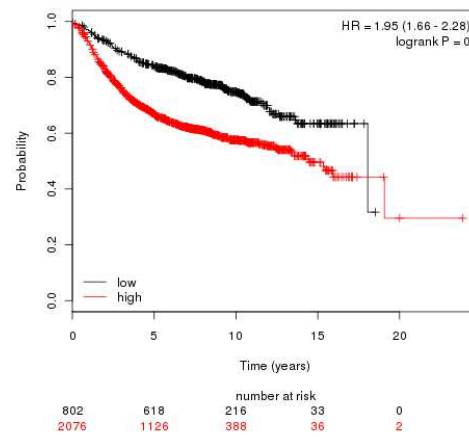
CCND2



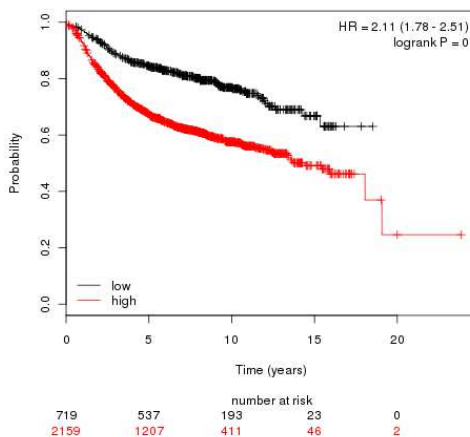
CCNE1



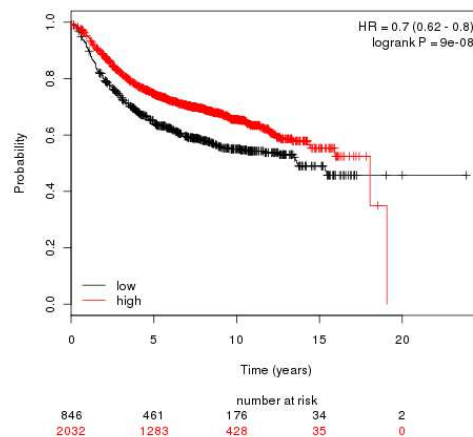
CCNE2



TOP2A



TK2



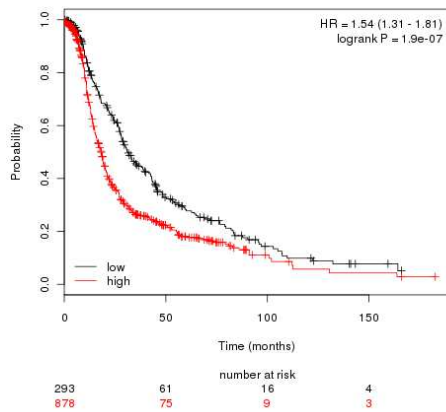
30. ábra. A www.kmplot.com/breast oldal felhasználásával készült ábrák az ASCO által vizsgált proliferációs marker-jelöltek közül az optimalizált vágópont alkalmazásával legjobban teljesítő 6 gén segítségével osztályozott betegek Kaplan-Meier túlélési görbéjét mutatják. Az előrejelzés a kiújulás mentes túlélés idejére vonatkozik.

4.12. Fejlesztések többgénés osztályozók és optimalizált vágópont számítására

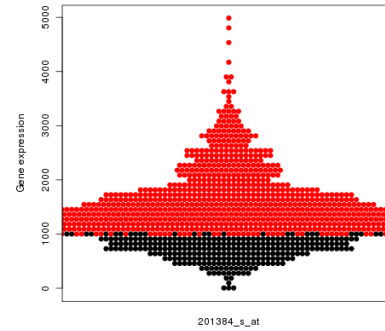
A létrehozott rendszer a világhálón a <http://www.kmplot.com/ovar> cím alatt érhető el. A beépített új opciók tartalmazzák több gén egyidejű mérésének lehetőségét, a legerősebb vágópont kiszámítását, a génexpressziós értékek eloszlásának egydimenziós ábrázolását, valamint további, a petefészekrákra specifikus szűrők alkalmazását. A program fejlesztése során létrehozott újabb algoritmusokat beépítettük a korábban elindított emlőtumoros mintákat felhasználó rendszerbe is. Az online felületet, valamint a több gén beadására létrehozott ablakot az alábbi ábrákon mutatom be.

31. ábra. Az világhálón keresztül elérhető elemző felület (felül), amely tartalmazza az új számítási lehetőségeket. A „use multigene classifier” gombra kattintva a felhasználó elé felugrik egy ablak, aminek segítségével több gént is be lehet egyidejűleg adni (alul).

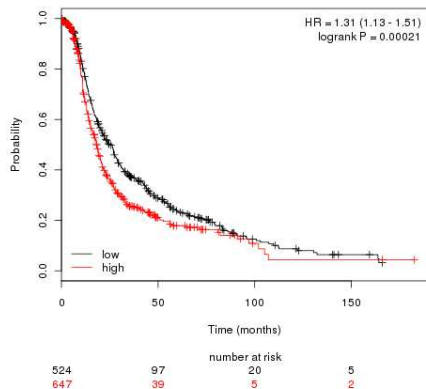
CA125



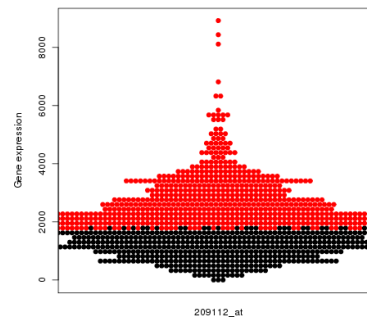
CA125



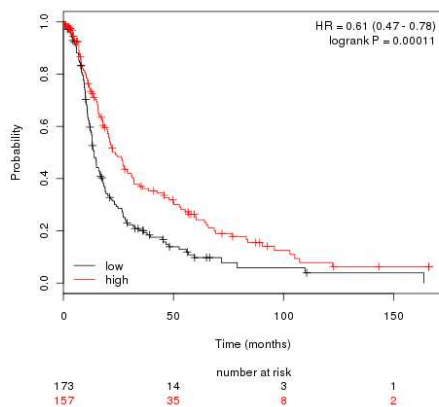
CDKN1B



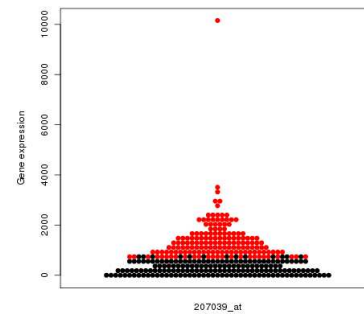
CDKN1B



P16



P16



32. ábra. Három kiválasztott gén expressziójának összefüggése a túléléssel (bal oldali ábrák), valamint ezen gének expressziójának egydimenziós eloszlása (jobb oldali ábrák). A P16-os gént csak a szuboptimális sebészeti utáni túlélés esetén vizsgáltuk, a többi gén vizsgálatába valamennyi beteg belekerült.

A petefészekrák túlélésével kapcsolatba hozott egyes gének részletes elemzése túlmutat jelen értekezés keretein. A számítások során a legjobban teljesítő gének a CA125 (HR=1,54, $p=1,9E-07$), a KLK6 (HR=0,79, $p=0.002$), az IFNG (HR=0,81, $p=0,004$), a P15 (HR=1,3, $p=0,0005$), a P16 (HR=0,61, $p=0,00011$), a CDKN1B (HR=1,31, $p=2,1E-04$) és a BIRC5 (HR=0,75, $p=0,00017$) voltak. Ezek közül a fenti ábrán bemutatom a CA125, a CDKN1B, és a P16 hatásosságát és az expressziós értékek egydimenziós eloszlását.

4.13. Online diagnosztikai rendszer fejlesztése

A létrehozott diagnosztikai rendszer a világhálón a <http://www.recurrenceonline.com> címen érhető el. Figyelembe véve, hogy a rendszerbe a gén chipeket egyesével kell feltölteni, ez időigényes lehet nagy mintaszám esetén, ezért elkészítettünk egy külön lefuttatható automatizált parancsfájlt is, amivel nagy mintamennyiséget is ki lehet egyszerre értékelni. A rendszer bemutatása utáni első év során *(vagyis a közlés és jelen értekezés megírása között eltelt idő alatt)* három megkeresés érkezett, amikor 100-nál nagyobb betegszámon akarták a számítások eredményét tudni.

A rendszer első verziója lehetővé tette, hogy a kiértékeléshez szükséges egyes paramétereket a felhasználó állítsa be. Ezek közé tartozott az alkalmazhatósági teszt (amire azért volt szükség, mivel a „recurrence score” csak nyirokcsomó negatív és ösztrogén receptor pozitív betegek esetén használható), a „range top” paraméter, ami az RT-PCR - gén chip átalakítás során fontos, és a próbák számításának módja. Az Affymetrix chippek egy gént több próbával is mérnek, ezért lehetőség nyílik a gént mérő próbák átlagos jelintenzitását vagy a legerősebb próba jelintenzitását használni, amelyet a felhasználók saját maguknak is beállíthattak. Azonban a rendszer első éve alatt lényegében elenyésző olyan elemzés történt, ahol ezen változóknál az előre beállított paramétereket megváltoztatták, ezért a rendszer újabb verzióiból kivettük ezt a lehetőséget.

Az elemzés végeredménye egy jelentés formájában készül el, ami PDF fájl formájában letölthető a honlapról. A fájl grafikus módon is ábrázolja a várható túlélést a számított pontértékek függvényében, illetve a receptor-státusz meghatározása során egy színkóddal ellátott csík mentén megjelöli a mért értéket, aminek segítségével a felhasználó a vágóponttól való eltérés mértékére kap visszajelzést.

INPUT

Patient ID *:

CEL file (HGU133A or HGU133+2 array) location*:

Lymph Node Status: *

(* = mandatory fields)

COMPUTE

☒ **Array quality control**
Check the CEL file for biases acquired during sample processing, hybridization and array scanning.

Recurrence Score

☒ **Risk category using strongest genes**
The selection of the best gene set is based on the lymph node status. Lymph node positive, ER negative and patients with unknown lymph node status are also eligible.

☒ **21-gene score**

Advanced parameters:
(not recommended to change)

☒ Eligibility check

Mean Mode:
Best Probe Set

Range Top: *
14

☒ **ER Status**

☒ **HER2 Status, based on**

☐ Bimodal distribution

☐ Immunohistochemistry

The results will be displayed in a pop-up window. If you would like to receive a copy, please enter your email address:

Execute

Patient ID: XXX

Age: 62

Lymph node status: Negative

Grade: 2

2010-09-28

Microarray platform: HGU133A

Eligibility Test: passed.

Array Quality Test: passed.

Computed Recurrence Score: 14 (low risk)

The recurrence score can be used to predict the probability of distant recurrence in patients with breast cancer who have estrogen-receptor positive tumors with no lymph nodes involved:

expected RFS

ER status: Positive. (MAS5 normalized gene expression: 9108)

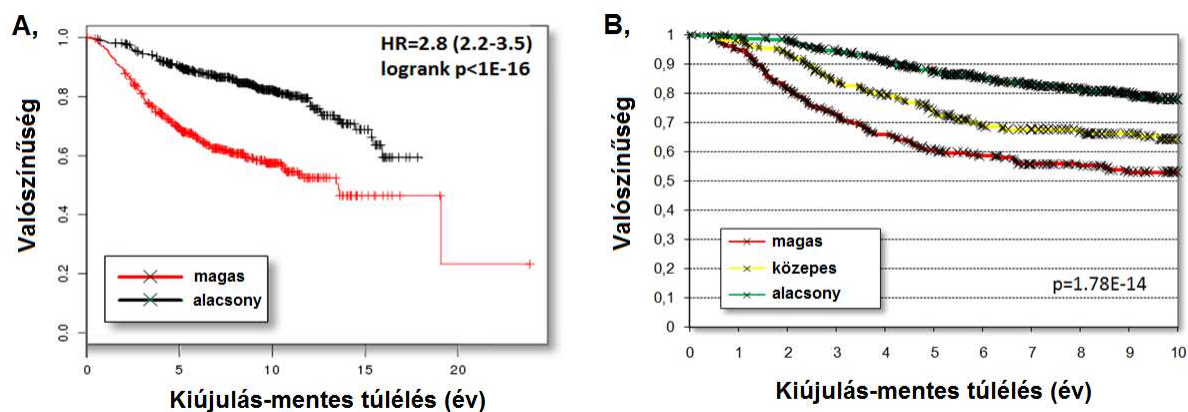
X									
62.5	125	250	500	1000	2000	4000	8000	16000	64000
ER negative				ER positive					

HER2 status: Positive. (MAS5 normalized gene expression: 1266)

X								
143.8	287.5	575	1150	2300	4600	9200	18400	36800
HER2 negative				HER2 positive				

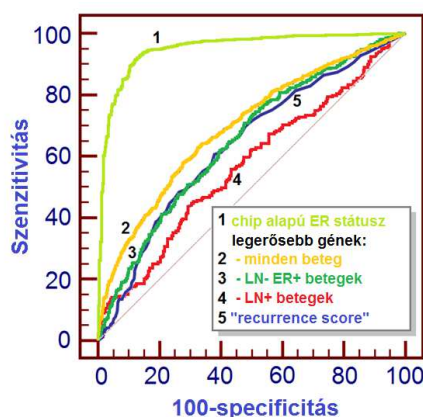
33. ábra. A világhálón keresztül elérhető kezelőfelület a részletes elemzési lehetőségekkel (bal oldalon), és egy visszaadott jelentés, amelyen három kiválasztott teszt készült el (jobb oldalon) (megjegyzés: a továbbfejlesztések miatt a honlap és a jelentés aktuális képe ettől eltérhet)

A kutatás második lépése a létrehozott rendszer elemzése volt nagy betegszámon. Ehhez 2472 gén chip adatait használtuk fel, amik közül 1509 olyan nyirokcsomó negatív és ösztrogén receptor pozitív beteg volt, akiknél a „recurrence score” számítása nem ütközött semmilyen akadályba. A betegeket mind a „recurrence score” algoritmus ($p < 1,78E-14$), mind a legerősebb gének alkalmazásával ($p < 1E-16$) magas hatásossággal tudtuk csoportokba osztani. A molekuláris markerek alapján végzett osztályozás meghaladta a klinikai paraméterek által elérhető beosztás szignifikanciáját (ösztrogén receptor: $p = 0,0002$, grade: nem szignifikáns). A rendszer alapjául szolgáló algoritmus ösztrogén-receptor meghatározó képességét egy korábbi fejezetben részletesen is bemutattam.



34. ábra. A legjobb gének (A) és a „recurrence score” (B) felhasználásával csoportosított nyirokcsomó-negatív, ösztrogén receptor pozitív betegek kiújulás-mentes túlélése

Az osztályozók teljesítményének összehasonlítását ROC elemzés segítségével végeztük. Az ROC végzése során egy fix vágópont alkalmazása nélkül, a valóban pozitív, álpozitív, valóban negatív és álnegatív betegek arányainak változása alapján lehet készíteni egy görbét, amely a szenzitivitást mutatja az 1-specificitás függvényében. A görbe alatti terület (AUC) 0,5 és 1 között változhat, ahol 0,5 a teljesen használhatatlan teszt (a véletlenszerű osztályozás), az 1 pedig egy tökéletes teszt osztályozó képességét mutatja. A legerősebb gének alkalmazásával AUC=0,695-öt, a „recurrence score” számításával AUC=0,637-et, az ösztrogén-státusz meghatározásánál pedig AUC=0,809-et ért el a rendszer.



35. ábra. Az egyes osztályozók relatív teljesítményének összehasonlítása ROC elemzéssel. A várható túlélést előrejelző osztályozók esetében az ötéves túlélésig bekövetkezett eseményt használtuk fel végpontként.

5. MEGBESZÉLÉS

5.1. Tizenegy kemoterápiás szerrel szembeni rezisztencia-profil meghatározása 30 rosszindulatú daganatos sejtvonal felhasználásával

Az elsőként bemutatott kutatásunk során tizenegy kemoterápiás szerrel szembeni rezisztencia-profilt határoztunk meg 30 rosszindulatú daganatos sejtvonal felhasználásával. A vizsgálatunk eredményei ²²⁹ alapján a következőket állapíthatjuk meg:

- sejtvonalak felhasználásával sikeresen tudtunk olyan géneket és génexpressziós mintázatokat azonosítani, amelyek a gyógyszerekkel szembeni rezisztenciát előre tudják jelezni;
- a DNS microarray technológia alkalmas arra, hogy az érzékenység és rezisztencia transzkriptom szintű markereit azonosítani tudjuk;
- sikerült azonosítanunk a gének egy olyan csoportját (pld TRIM2, APOBEC3B), amelyek több gyógyszerrel szembeni rezisztenciában is fontos szereppel bírhatnak.

A kutatás eredményei közül az egyik legérdekesebb lelet az irodalomban korábban intenzíven vizsgált multidrog-rezisztencia gének (MDR1, MRP1) kifejeződése volt. Bár azon sejtvonal, ahol az MDR1 kifejeződése a legmagasabb (SNU449), az összes vizsgált gyógyszerrel szemben ellenálló volt, de az MDR1 univerzális szerepe az összes sejtvonalra nézve már nem igazolódott. Ezzel eredményeink alátámasztották a fenotípusos különbségek mögött a tumorok nagyfokú genetikai heterogenitását. Ennek eredményeképpen a különböző sejtvonalak különböző rezisztencia-mechanizmusokat tudnak kifejleszteni (*lásd még 4.5 és 5.5. fejezet*).

Bár készültek más, sejtvonalakat felhasználó vizsgálatok is, például az NCI-60 sejtvonalak felhasználásával ^{54, 230, 231}, azonban az általunk végzett kísérlet jelentősen különbözik ezektől. Mi ugyanis nem az úgynevezett GI50 értéket határoztuk meg, hogy a sejtvonalakat két csoportra osszuk. A GI50 érték azt mutatja meg, hogy egy gyógyszerből melyik az a koncentráció, amely a szaporodást 50%-ban gátolni tudja. Azonban a GI50 akár olyan gyógyszerkoncentráció mellett is bekövetkezhet, amely már klinikailag nem alkalmazható vagy a farmakokinetikai korlátok, vagy pedig a gyógyszerek toxikussága miatt. Ezért mi a korábbi vizsgálatokban megállapított klinikai koncentrációkat ²³² alkalmaztuk, amelyek már tartalmazzák a gyógyszer hígulását, kiürülését, bomlási idejét stb. Az optimális

koncentráció kiválasztása mellett vizsgálatainkba olyan gyógyszereket vettünk be, amelyek a rosszindulatú daganatos betegek kemoterápiás kezelése során a legelterjedtebbek.

A vizsgálat során három olyan további kutatási kérdés jelentkezett, amelyek a következőkben elvégzett kutatások alapkérdéseként jelentek meg:

- lehetséges volna-e a sejtvonalakban azonosított gének segítségével klinikai betegekben is előre jelezni a kemoterápiával szembeni rezisztenciát? *(lásd 5.3. fejezet)*
- a vizsgálat során a sejtvonalak gyógyszer nélküli génexpressziós mintázatát használtuk fel - lehetne-e célzott terápiás szerekkel szemben is rezisztencia-géneket azonosítani hasonló *in vitro* vizsgálatokkal? *(lásd 5.6. fejezet)*
- a statisztikai kiértékelés során több előfeldolgozási algoritmust is használtunk (pld MAS5, RMA, VSN), azonban ezek eredményessége nem feltétlenül egyezik - melyik lehet a legmegfelelőbb normalizálási módszer? *(lásd 5.10. fejezet)*

5.2 A kemorezisztencia génjei melanoma malignumban

A kemorezisztencia génjeit melanoma malignumban vizsgáló kutatásunk eredményei ²³³ alapján a következőket állapíthatjuk meg:

- rezisztens és érzékeny sejtvonalpárok felhasználásával lehetséges volt olyan géneket azonosítani, amelyek különbözőképpen fejeződnek ki az egyes vonalakban;
- fürtelemzés segítségével az egyes vonalak közötti hasonlóságot tudtuk azonosítani;
- az irodalomban korábban leírt géneket (pld STK17A, CPNAS) sikeresen lehet igazolni gén chip mérések felhasználásával.

Az 5.1 és 5.2. fejezetekben leírt projektek eredményei egymást is megerősítik, amennyiben a kérdés a gén chipek felhasználásával génexpresszióbeli különbségek azonosítása.

Hasonló vizsgálatok alapján lehetséges további, egy vagy több kiválasztott gént felhasználni, a kemorezisztenciával összefüggő biomarkerek azonosítása más tumortípusok esetén is. Meg kell jegyeznünk, hogy ezen tanulmány eredményeivel szemben egy későbbi vizsgálatunkban a génexpressziós mintázatok alapján a melanoma áttétképző potenciáljával kapcsolatba hozott génlistákat nem sikerült igazolnunk ²³⁴ *(lásd 5.7. fejezet)*.

5.3. A várható túlélés előrejelzése sejtkultúra modellben azonosított gének felhasználásával antraciklin monoterápiával kezelt betegekben

A biomarkerek sejtkultúra alapú azonosítása során fontos kérdés, hogy az eredmények mennyire alkalmasak klinikai mintákban a terápiás válasz vagy túlélés előrejelzésére. Az ezen kérdés megválaszolására végzett kutatás során doxorubicin monoterápiával kezelt betegek várható túlélését becsültük meg sejtkultúra modellben azonosított gének felhasználásával. A kutatásunk eredményei ²³⁵ alapján a következőket állapíthatjuk meg:

- Rezisztens és érzékeny sejtvonalpárok felhasználásával két gyógyszerrel (doxorubicin és paclitaxel) szemben azonosítottuk a rezisztencia legfontosabb génjeit.
- Egy független, 44 beteget tartalmazó betegmintán alkalmazva a génexpressziós mintázatuk alapján általunk doxorubicin kemoterápiára érzékenyként osztályozott betegek túlélése több mint 50%-al hosszabb volt, mint a rezisztensként osztályozott betegek túlélése.
- Vizsgálatunk igazolta az *in vitro* rendszerek használhatóságát klinikailag is alkalmazható előrejelző tesztek kifejlesztésére.

Az elemzés során a PAM program osztályozó funkciójával készítettük el a klinikai előrejelzést az összes betegre - de nem a legjobb gének, hanem a teljes génexpressziós mintázat felhasználásával. Ezáltal összességében nem a legjobb gének hatékonyságát néztük és nem egy többgénos osztályozót igazoltunk, hanem az általunk elvégzett elemzéssel a minták hasonlóságát állapítottuk meg. Ez fontos, koncepcionális kérdés, amivel tehát elsősorban az általunk felállított modell rendszer használhatóságát igazoltuk, nem pedig egy konkrét prediktív tesztet állítottunk elő.

A vizsgálat folyamán különböző szöveti eredetű sejtvonalakat használtunk. Az eltérő sejtvonalak alkalmazása lehetővé tette, hogy az egyes szövettípusokra nem specifikus rezisztencia géneket tudjunk azonosítani. Ezáltal a korábbi vizsgálatok szöveti különbözősége visszavezethető technikai gyengeségeitől mentesülhettek az eredményeink.

Az általunk felhasznált klinikai minták egy korábbi vizsgálatból származtak ⁵³. Sorlie és munkatársai azt is leírták, hogy az általuk azonosított úgynevezett luminális A altípusú betegek várható prognózisa jobb, mint a luminális B illetve a többi altípusba sorolt betegeké.

Hogy az eredményeiket össze tudjuk hasonlítani az általunk készített osztályozással, ezért az általuk ugyanezen betegekre közölt molekuláris altípus beosztást összehasonlítottuk az általunk számolt várható kemoterápiás válasszal. A bazális, a HER2 pozitív és a luminális B altípus szinte kizárólag rezisztens betegeket tartalmazott, miközben a luminális A altípusba került a legtöbb érzékeny beteg. Ezek alapján a mi eredményeink Sorlie-ék vizsgálatait nemcsak megerősítik, de lényegében meg is tudjuk az ő eredményeiket magyarázni. A több kemoterápiára érzékeny beteg későbbi kezelése után várható jobb terápiás válasz ugyanis közvetlenül is megjelenhet jobb prognózis formájában. Amit még ki kell emelni, hogy a Sorlie-ék által bevezetett molekuláris altípus beosztás a betegek egy részét nem tudta osztályozni, miközben az általunk létrehozott kemorezisztencia-előrejelzés ezekben a betegeken is működött. Mindezen eredmények a sejtvonalmodellek alkalmazásának lehetőségeit támasztják alá.

A doxorubicinnal szembeni rezisztenciára vonatkozó eredményeinket két korábbi vizsgálattal is összehasonlítottuk. Suganuma és munkatársai gyomorrákban vizsgálták a rezisztencia génjeit négy gyógyszerrel szemben ²³⁶, azonban ők nem közöltek olyan konkrét génlistát, amely a doxorubicin kezeléssel szembeni rezisztencia génjeit tartalmazta volna. Kang és munkatársai szintén gyomorrákos sejtekkel dolgoztak ²³⁷, és 250 olyan gént azonosítottak, amelyek a doxorubicin-rezisztenciával összefügghetnek. Bár nem találtunk olyan gént, amelyik az általuk közölt génlistán és a mi sejtvonalainkban is érintett lett volna, azonban ez nem zárja ki azt az elképzelést, amely szerint egy adott rezisztenciát több különböző mechanizmus (génlista) egymástól függetlenül is létrehozhat ²³⁸.

Végezetül meg kell jegyeznünk, hogy szerencsés helyzetben voltunk, amikor az általunk vizsgált betegek egy monoterápiával kezelt csoportból származtak. A ma is alkalmazott klinikai protokollok esetén ez már szinte soha nem realitás, a mindennapi gyakorlat számos szer egyidejű vagy szekvenciális alkalmazását tartalmazza. Ezért egy-egy gyógyszer izolált hatásának kiszűrése, illetve fordítva, az egy gyógyszerrel szembeni érzékenység és rezisztencia biomarkereinek azonosítása ma sokkal nehezebb feladat.

5.4. Genom-szintű adatokkal azonosított PSMB7 gén oki szerepének igazolása *in vitro* körülmények között

A fentebb leírt vizsgálatok során egyrészt igazoltuk, hogy a gén chipek segítségével azonosítani lehet az eltérő kifejeződést mutató géneket, másrészt azt is bizonyítottuk, hogy az ezen gének alapján készített mintázatokkal akár klinikai betegekben is előre lehet jelezni a várható prognózist. Ezek alapján azonban nem lehet eldönteni, hogy egy adott gén „csak egy biomarker”, amely a vizsgált kórfolyamattal együtt utazó potyautas („passenger event”), vagy pedig olyan gén, amelyik a kórfolyamat létrejöttében oki szereppel is bír („driver event”).

A technikai korlátok miatt nem lehetséges az összes szóba jövő gén oki szerepének egyidejű vizsgálata, ezért érdemes a rendelkezésre álló tudás, vagyis a már korábban közölt irodalom alapján azokat géneket, vagy géncsaládot kiválasztani, amelynek funkcionális szerepét feltételezhetjük. A sejtvonalakon végeztünk egy ilyen vizsgálatot, amelynek során a RAS jelátvitel gátlásának szerepét vizsgáltuk a rezisztencia létrehozásában ²³⁹.

Egy további ilyen szóba jövő géncsalád a proteaszómákat felépítő géneket tartalmazza. A proteaszóma több katalitikus hellyel rendelkező fehérjelebontó komplex, amely a szükségtelenné vált fehérjék ATP-függő lebontásában vesz részt. A proteaszómák részt vesznek a sejtciklus szabályozásában és az érújdonképződésben is. A fehérje több alegységből áll össze, amelyek között alfa és béta alegységeket, valamint még ezeken belül számmal jelölt további kisebb részeket is tartalmaznak. A katalitikus alegységek a PSMB5, PSMB7 és a PSMB6 egységeken helyezkednek el ²⁴⁰.

A proteaszómák vizsgálata azért különösen érdekes, mivel rendelkezésünkre állnak proteaszóma gátlószerek, amelyek a fehérjék lebontását megállítják. A gátlószerek tumorelles hatását már korábban leírták, a sejtciklus befolyásolásán keresztül programozott sejthalált indukálnak²⁴¹. Az endoplazmatikus retikulumból stressz hatására induló programozott sejthalál beindítása ²⁴² adja az első proteaszóma-gátlószer, a myeloma multiplexben alkalmazott bortezomib ²⁴³ hatásmechanizmusának alapját is. Ezek mellett olyan vizsgálatot is közöltek, ahol leírták a bortezomib és a doxorubicin kezelés egymás hatásosságát kölcsönösen erősítő hatását ²⁴⁴. Azonban a doxorubicin rezisztenciára gyakorolt hatását, illetve maguknak a proteaszóma alegységeknek a szerepét mindeközéig még nem vizsgálták.

A 4.4. fejezetben ismertetett projekt során egy proteaszóma alegység génjének oki szerepét igazoltuk sejtkultúramodell, transzkriptomikus adatok, RNS interferencia-alapú génelcsendesítés és publikus betegminták felhasználásával. A vizsgálat eredményei ²⁴⁵ alapján a következő megállapításokat tehetjük:

- A kezelés és a rezisztencia közötti genom-szintű különbségeket felhasználva azonosítottuk a proteaszóma fehérjekomplex PSMB7 alegységét, mint a doxorubicinnel szembeni rezisztencia egyik oki („driver”) génjét.
- A vizsgálat során igazoltuk, hogy a genom-szintű adatokkal azonosított célgének kísérletes körülmények között is magas relevanciával rendelkeznek.
- A nagy klinikai mintán is igazolt eredmények alapján a rossz prognózissal kapcsolatba hozható PSMB7 gén ígéretes terápiás célpont lehet.

Már korábban is tudtuk, hogy egy-egy gén magas kifejeződése összefügghet a kemorezisztencia kialakulásával, amire a legtöbbször leírt példák az ABCB1 ^{246, 247} és a TOP2A ^{248, 249} gének. Az siRNS-el való csendesítése alapján a PSMB7 gén használhatósága is túlmutat a sima biomarker jelöltségen.

Meg kell jegyeznünk, hogy az siRNS hatás mérése kizárólag sejtkultúrák körülmények között történt. A homogén körülmények között tartott, könnyen transzfektálható sejtek lehetővé teszik, hogy egy-egy gén izolált hatását tudjuk mérni. Ehhez az RNS interferencia alapú géncsökkentés az irodalomban bevett, széles körben elterjedt módszer. Azonban ezek az eredmények a módszer korlátai miatt klinikailag nem alkalmazhatóak azonnal - ennek fő oka az *in vivo* transzfekció megoldatlan volta. Sejtekben számos különböző hatékony technológiát leírtak már, mint a kationos lipidek, nanorészecskék, polietilénemin, cyclodextrin, chitosan és a kollagén alkalmazása ²⁵⁰. Azonban *in vivo* az RNS interferencia működéséhez szükséges körülményeket nem lehet reprodukálható és hatékony módon létrehozni. Emellett további probléma, hogy a hatás leszűkítése a rosszindulatú daganatos sejtekre szintén nem megoldott.

A PSMB7 gént korábban alig vizsgálták rosszindulatú daganatos betegekben. Magas expresszióját dokumentálták vastagbél-carcinómában ²⁵¹. A proteaszóma-gátlószerekkel történő kezelés után a PSMB5, PSMB6 és PSMB7 alegységek expressziója megnőtt egerek máj, tüdő, vékony- és vastagbelében. Jelenleg még rendkívül hiányos ismereteink ellenére

úgy tűnik, hogy a magas proteaszóma expresszió szerepet játszhat a nagy fehérje-komplexek által közvetített betegségekben is ²⁵².

Jelen vizsgálatunkban a PSMB7 nemcsak sejt kultúrák körülmények között volt hatásos, de a klinikai mintákon is. Ezen vizsgálat értékét csökkenti, hogy a kezelési adatok nem álltak rendelkezésre, ezért nem tudtuk megítélni, hogy az 1592 beteg közül mennyien részesültek valóban doxorubicin-kezelésben.

5.5. Parallel evolúciós modell alkalmazása a párhuzamosan létrejövő kemorezisztencia modellezésére

A fenti vizsgálatok során alkalmazott sejt kultúrák modellek alapja az a feltételezés, hogy a kísérletes körülmények között tartott sejtek a való világhoz hasonló rezisztenciamechanizmusokat fognak „alkalmazni”. Ezen hasonlóság vizsgálatára a fentiektől eltérő módon párhuzamosan hoztunk létre rezisztens vonalakat, majd ezeket egymással hasonlítottuk össze. Eredményeink ²⁵³ alapján megállapíthatjuk:

- A sejtvonalakban történő párhuzamos kezelés révén sikerült 27 sejtvonalban létrehozni a párhuzamos evolúció modelljét egy konkrét szelekciós nyomással szemben.
- Több, korábban már leírt különféle mechanizmus aktiválódását (pld TOP2A, TUBB3, ABCB1) tudtuk igazolni a rezisztencia létrejötte során.
- Mindössze 1-2 mechanizmus aktiválódása már elégséges volt ahhoz, hogy a rezisztencia létrejöjjön.

A kutatás egyik legfontosabb kérdése, hogy egyáltalán klinikailag releváns-e az általunk alkalmazott megközelítés. Az RT-PCR és az áramlási citometriás mérések során olyan mechanizmusokat és géneket vizsgáltunk, amelyeket az irodalomban már többen is leírtak, ezeket a mechanizmusokat sikeresen tudtuk a létrehozott vonalokban is kimutatni. Ezek alapján sejtvonalmodellünk alkalmas lehet arra, hogy ezzel a megközelítéssel új mechanizmusokat is azonosítsunk.

Érdekes megfigyelés, hogy nem találtunk olyan vonalat, amiben egy gyógyszer hatására keresztrezisztencia jött volna létre a többi vizsgált szerrel (doxorubicin, paclitaxel, cisplatin, 5-fluorouracil) szemben is. Ezen eredményeink alapján a valódi „multidrog-

rezisztencia” egy igen ritka esemény - ez egyezik azon klinikai megfigyelésekkel, amelyek szerint egy szerrel szemben ellenállóvá vált tumorokat második vonalban más gyógyszerrel még hatásosan lehet kezelni ²⁵⁴.

Az egyes párhuzamosan fejlesztett sejtvonalak rezisztenciája eltért, ami felveti, hogy klinikailag is létrejöhet hasonló jelenség. Ezt az egy tumor több részéből (vagy akár a távoli áttétekből) vett többszörös mintavétel alapján lehetne *in vivo* is vizsgálni. Alternatív megközelítés lehet a tumor heterogenitásának vizsgálata egyetlen mintából ²⁵⁵, vagy a keringő tumorsejtek többszörös összehasonlító vizsgálata ²⁵⁶. Hasonló vizsgálatok klinikai értéke azonban továbbra is kérdéses, hiszen a létrejövő rezisztenciáért akár a tumorsejtek egy olyan kis hányada is felelőssé tehető, amelyeket többszörös mintavétellel sem tudnánk azonosítani.

5.6. Öt célzott terápiás szerrel szembeni rezisztencia és érzékenység vizsgálata 45 sejtvonalban

A fentiek során a citotoxikus gyógyszerekkel szembeni rezisztenciát tanulmányoztuk. Egy újabb kutatás során öt célzott terápiás gyógyszerre (lapatinib, erlotinib, sorafenib, sunitinib és gefitinib) fókuszáltunk, ezeken az 5.1. fejezethez hasonló kísérletet végeztünk. Vizsgálatunk 4.6. fejezetben ismertetett eredményei ²⁵⁷ alapján megállapíthatjuk:

- Az alkalmazott, sejtkultúra és transzkriptom szintű elemzések kombinációjára épülő megközelítés alkalmasnak bizonyult arra, hogy célzott terápiás szerekkel szembeni relatív rezisztenciát előrejelző biomarkereket azonosítsunk.
- Veserákos mintákban igazoltuk a RAB17, az LGALS8 és az EpCAM gének összefüggését a sunitinib-rezisztenciával.
- A legtöbb keresztrezisztenciával összefüggő gént a sunitinib-rezisztenciával kapcsolatba hozott génlista tartalmazta.

Felvetődik a kérdés, hogy miért a sunitinib-rezisztenciával összefüggő génlista tartalmazta a legtöbb olyan gént, amely egy másik listában is szerepelt? Ez a megfigyelés ellentétben áll a rendelkezésre álló irodalmi adatokkal, ahol a sunitinib-rezisztenciával sokkal kevesebb gént hoztak kapcsolatba, mint a másik négy gyógyszerrel szembeni rezisztenciával. Emiatt a sunitinib-rezisztencia génjeire koncentráltunk, és ezekben végeztük el a klinikai tesztelést. Bár eredményeink igazolták a gének egy részét, azonban a vizsgált betegminták

alacsony száma miatt (39 beteg) a talált összefüggéseket mindenképpen további, jóval nagyobb klinikai vizsgálatoknak kell igazolniuk.

A gyógyszerek hasonló hatásmechanizmusát ismerve meglepő, hogy egyetlen olyan gént sem találtunk, amely mind az öt vizsgált szerrel szembeni rezisztenciával összefüggést mutatott volna. Négy gyógyszerrel szembeni rezisztenciával is mindössze kettő gén függött össze. Az ANXA3 részt vesz a növekedéssel összefüggő jelátvitelben ²⁵⁸, korábban platina-szerekkel szembeni rezisztenciával is kapcsolatba hozták ²⁵⁹, valamint az EGFR mutáció által létrehozott génexpressziós változások is érintik ²⁶⁰. A RAB25 a RAS onkogének családjába tartozik, és bár vastagbélrákkal ²⁶¹ és a receptor-negatív emlőrákkal ²⁶² már összefüggésbe hozták, de nincs adat a tirozin kináz gátlószerekkel való kapcsolatáról.

Az elvégzett kísérlet nagyban hasonlít az elsőként bemutatott vizsgálatunkra. Ez nem véletlen, mivel ebben a projektben felhasználtuk az előző vizsgálatban gyűjtött pozitív tapasztalatokat. Ezek között van a rezisztencia mértékének meghatározása három koncentráció mellett, a sok különböző eredetű sejtvonal alkalmazása a tumortípustól független gének azonosításához, a klinikailag alkalmazott koncentráció használata és a többlépcsős statisztikai kiértékelés. Jelen projekt azonban felül is múlta a korábbi vizsgálatunkat, mivel több sejtvonalat használtunk, a gén chip mérések háromszoros ismételtsében készültek, és a legjobb géneket független klinikai vizsgálatban is ellenőriztük.

5.7. Korábbi közleményekben a melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása

Az 5.2. fejezetben már bemutattuk, hogy saját sejtkultúra alapú vizsgálataink során a melanoma kemorezisztenciájával összefüggő géneket tudtunk azonosítani. Kérdésként vetődik fel, hogy az irodalomban azonosított génexpressziós mintázatok mennyire reprodukálhatóak független betegmintákban is. Az ezen kérdés megválaszolására végzett, a 4.7. fejezetben összefoglalt projektünk eredményei ²⁶³ alapján a következő megállapításokat tehetjük:

- A melanoma génexpressziós mintázatok alapján generált génlisták különbözősége a tumorok nagyfokú heterogenitását mutatja.
- A génexpressziós mintázatok generáló experimentális és klinikai vizsgálatok reprodukálhatósága alacsony.

A kutatás alaphipotézise alapján azonosított melanoma-specifikus gén mindössze egyetlen egy volt (a RAB33A), amely Ras-asszociált gén ²⁶⁴. A Ras jelátviteli útvonal génjeit egy másik közleményünkben alaposan elemeztük ²⁶⁵.

Több oka lehet annak, ami miatt az egyes vizsgálatok eredményeinek 95%-a nem reprodukálható. Az egyes cikkekben felhasznált szövetek eredete jelentősen eltért egymástól (tumorok, áttétek, sejtvonalak). A klinikai minták tisztaságát nem ellenőrizték, ami miatt jelentős lehet a szennyeződés vérrel vagy kötőszövettel. A tanulmányokban vizsgált minták klinikai paraméterei is eltértek egymástól.

A vizsgálatok során több esetben nagyon alacsony esetszámokkal dolgoztak, amely a megbízható biomarkerek azonosítását lehetetlenné teszi. Az általunk fentebb már részben említett, de a következő fejezetekben részletesen is elemzett kutatásaink során pontosan az volt az egyik fő célunk, hogy a különböző eredetű adatok integrálásával egy olyan adatbázist hozzunk létre, amely lehetővé teszi, hogy az egyes vizsgálatokban elért statisztikai hatásosságot jelentősen meghaladó elemzéseket tudjunk végezni.

Meg kell jegyeznünk, hogy a vizsgált tanulmányokban alkalmazott technológia is jelentősen eltért. Ez egyrészt a fizikális méréshez használt gén chipok különböző voltát tartalmazza, de az adatok statisztikai kiértékelése során is eltérő algoritmusokat vetettek be. Mindezek a különbségek felvetik, hogy csak abban az esetben fogunk tudni megbízható és reprodukálható transzkriptomikus vizsgálatokat végezni, ha a kísérletek körülményeit a lehető legnagyobb mértékben standardizálni tudjuk.

Összességében tehát azt állapíthatjuk meg, hogy az összehasonlítható és reprodukálható eredmények generálásához szükséges lesz a kutatási programok során a mintagyűjtés, a minták előkészítésének, feldolgozásának és a statisztikai kiértékelésnek standardizálása.

5.8. Korábbi közleményekben emlőrák prognózisával kapcsolatba hozott génlisták keresztelemzése nyers adatok újrafeldolgozása után

Az előző fejezetben tárgyalt, általunk végzett első keresztelemzési vizsgálat egy jelentős gyengesége, hogy nyers adatok hiányában az eredeti szerzők által már feldolgozott adatok alapján készített génlistákat hasonlítottuk egymással össze. Mindezek alapján felvetődött a kérdés, hogy lehetséges lenne-e az eredeti közlemények nyers adatait kombinálni

és teljesen újra feldolgozni a fenti iránymutatások figyelembevétel, valamint hogy ebben az integrált adatbázisban ellenőrizve az egyes gének vagy génexpressziós mintázatok valódi teljesítménye mit fog mutatni?

Az ezen kérdés mentén végzett, a saját eredményeinket a 4.8. fejezetben bemutató kutatásaink ⁷⁷ alapján a következő megállapításokat tehetjük:

- Korábban közölt adatok felhasználásával létrehoztunk egy integrált, klinikai és génexpressziós adatokat tartalmazó adatbázist, amely a prognosztikus biomarkerek keresztelemzését lehetővé teszi. Ezen adatbázist később további vizsgálatok során is fel tudtuk használni ²⁶⁶.
- Az irodalomban közölt mintázatok többsége nem volt képes a jó és rossz prognózisú betegeket egymástól elkülöníteni. Az általunk szignifikánsnak talált génlisták (n=8) hatásosságát eközben egy független platformon igazoltuk.
- Az adatbázis felhasználásával azonosított gének közül kevesebb is elég egy új osztályozóba, ha a klinikai paramétereket is figyelembe vesszük.
- Nagyszámú redundáns gén chipet közöltek korábbi közleményekben (n=640).

A redundáns gén chipek közlése egy igen érdekes szokással társult. Időben az első közleményt Miller és munkatársai hozták nyilvánosságra GSE3494-es GEO azonosítószám alatt ²⁶⁷. Amint ezt a közleményt elfogadták, a rákövetkező napon adták fel a következő, Sotiriou és munkatársai által fémjelzett cikket ⁷¹, amelyhez az adatokat GSE2990-es számon regisztrálták. Ezután ugyanezt a módszert folytatták tovább Ivshina és munkatársai (GSE4922) ²¹⁸, illetve később Loi és munkatársai (GSE6532) ²⁶⁸. Érdekes kérdés, hogy ez a módszer vajon nem etikátlan-e? Elsősorban a kutatóknak jelentős költségmegtakarítást eredményezett, hogy azonos befektetett munkával több jó közleményt is el tudtak fogadtatni. Azonban lényegében arról van szó, hogy az első közlemény során a rendelkezésre álló adatok egy részét visszatartották, ami a független igazolást, illetve az adatok felhasználását más vizsgálatokban jelentős mértékben megnehezíti.

Ezen túlmenően azonban van egy sokkal nagyobb probléma is a redundáns közléssel. Egyes vizsgálatokban hozzánk hasonlóan mások is törekednek a független tesztelésre. Amennyiben egy ilyen független tesztelésbe két függetlennek tűnő adatbankot vonnak be, amelyek azonban nagyfokú hasonlóságot mutatnak (pld GSE3493 és GSE4922), akkor a független tesztelés hatásosságát jelentősen lecsökkenti a pár többszörösen közölt minta.

Ebben az esetben ugyanis lényegében nem más történik, mint hogy ugyanazon minta súlya a végső osztályozó hatásosságára óriási mértékben megnő - ennek háttere mögött az áll, hogy legtöbb esetben az alkalmazott statisztikai tesztek a t-próbára épülnek. Azonban a t-próba során a szórást is bele vesszük a statisztikai számításba. Értelemszerűen azonos minták alkalmazása a szórást rendkívüli mértékben csökkenti (a nullához közelíti), ezért valamennyi statisztikai eredmény erősen megkérdőjelezhető. A fenti, véleményünk szerint hibás megközelítésre a legjobb példa a Breastmark rendszer³. Mindezek az eredmények mutatják, hogy a nyilvánosan is elérhető adatok magasabb kontrolljára lenne szükség. *(Megjegyezzük, hogy vizsgálatunk közzéje után a GEO-ban kiegészítették a megfelelő GSE számokat egy utalással, amely a többi hasonló közleményre utal. Optimális esetben azonban nem megtartani, hanem teljesen törölni kellene a redundáns adatokat.)*

Gén chip vizsgálatok alapján azonosított biomarkerek független igazolása minden esetben elengedhetetlenül a vizsgálat része kell hogy legyen²⁶⁹. Egyes szerzők már korábban is felvetették, hogy ezt egyetlen (ugyanazon) platform felhasználásával kell a másik (vagy nagyobb) beteganyagon elvégezni²¹⁶. Lényegében ez az, amit mi ebben a vizsgálatban a korábbi chip adatok alapján elvégeztünk.

Kritikaként merülhet fel az előző pozitívum, hogy a vizsgálat során a géneket egy platformra vetítettük, ugyanis ezzel egyes, nem teljesen tükrözhető gének kiesnek, illetve más géneket eltérő hatásossággal tudunk lemérni. Ezen művelet során azt a korábbi megfigyelést vettük alapul, hogy a gén chippek egy-egy konkrét gén expresszióját megbízható és reprodukálható módon tudják meghatározni¹⁶⁰. Emiatt az egyes platformok használata kevésbé kritikus, amennyiben az általunk választott géneket mérő próbák nem kötnek be több génhez is. Egyrészt a gén chipeket tervező cégek törekedtek arra, hogy ez ne fordulhasson elő, másrészt viszont több esetben is független technológia *(például RT-PCR - lásd következő fejezet)* segítségével is igazolták egyes kiválasztott gének eltérő kifejeződését.

³ http://cancerres.aacrjournals.org/cgi/content/short/72/24_MeetingAbstracts/P3-04-12?rss=1

5.9. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének igazolása klinikai mintákban

A következő kutatásunk során nemcsak a génlistákat azonosítottuk, de általunk gyűjtött független beteganyagon is ellenőriztük az egyes gének eltérő kifejeződését. Saját eredményeink²⁷⁰ alapján a következő megállapításokat tehetjük:

- Az általunk elvégzett kiértékelés során azonosított, a túléléssel legjobban összefüggő gének a PGR, az ÖR2 és a TSPAN8 voltak.
- Az általunk gyűjtött klinikai mintákban RT-PCR alkalmazásával a gén chipen azonosított gének több mint 90%-át sikeresen tudtuk igazolni.
- A korábban közölt génlisták közül nyolc volt képes a szövettani altípus elkülönítésére, de egy sem tudta a túlélést előre jelezni.

A korábbi vizsgálatokban elért alacsony statisztikai erő, illetve az eredmények rossz reprodukálhatósága miatt még nem készült olyan transzkriptom szintű vizsgálat, amely a klinikai döntéshozatalt is befolyásolhatta volna. Jelen vizsgálatban nemcsak a nagy betegszám járult hozzá a statisztikai erő növeléséhez, de a független technológia felhasználásával elvégzett igazolás is az eredményeket erősíti.

A legérdekesebb eredmény az ÖR2 és a PGR összefüggése a túléléssel. Irodalmi adatok alapján az ösztrogén receptor fokozott kifejeződése inkább az alacsony malignitású tumorokra volt jellemző²⁷¹. Jóllehet a kutatásunk során az ÖR-t nem találtuk meg a legerősebb gének között, azonban az ÖR2 azonosítása teljesen váratlan lelet volt. Mindezek az eredmények azt sugallják, hogy a petefészekrák kezelése során a hormonterápiának is lehet helye, ha a megfelelő beteganyagot ki tudjuk szelektálni.

A kutatás során létrehozott adatbázis további felhasználási lehetősége, hogy független klinikai teszteléseket végezzünk újonnan publikált többgénes vizsgálatok eredményeinek igazolására. Ennek során lehet az egyes vizsgálatokat külön-külön igazolni, ahogy fent több vizsgálat során már bemutattuk. Egy alternatív megközelítés, ha ezt a lehetőséget minden kutatónak a kezébe adjuk, ami a világhálón keresztül alkalmazott programok fejlesztése révén ma megvalósítható cél - ezt fogom bemutatni a következő fejezetekben.

5.10. Microarray adatok előfeldolgozása

A 4.1., 4.4., 4.6., 4.8. és 4.9. fejezetekben leírt kutatásaink során a mérési adatokat gén chippek (más néven génexpressziós microarray-ok) felhasználásával generáltuk. A technológia alkalmazása során az eredményt egy leolvasó („scanner”) segítségével határoztuk meg, amely egy képfájl formájában menti le a mért adatokat. Ezen képfájlban az egyes képpontokat elhelyezkedés alapján azonosítja a program, hozzárendeli az adott pozícióban elhelyezett próbához, majd a kapott intenzitási értékeket egy következő fájlban tárolja el. Az egyes génekre vonatkozó kifejeződést ezen fájl további feldolgozásával tudjuk kiszámolni, amely folyamatot előfeldolgozásnak („pre-processing”) nevezzük, mivel ezt még az egy kísérleten belül elvégzett több mérés összehasonlítása előtt le kell futtatni.

Az előfeldolgozás három további részlépésre osztható, mint a háttér korrekció (amelynek során az a célunk, hogy az esetleg mérhető háttérzajt kiszűrjük), a normalizálás (melynek során célunk, hogy a különböző minták esetében kapott génexpressziós értékek egymással összehasonlítható skálán legyenek) és a próbaszintű adatok összesítése (melynek során célunk, hogy az egy gén kifejeződését mérő több próba adatait egyetlen értékre alakítsuk át). Erre a folyamatra számos algoritmust fejlesztettek ki, amelyek azonban egymástól jelentősen el is térhetnek. Figyelembe véve, hogy kutatásaink során egyrészt mi magunk is jelentős számú microarray mérést végeztünk, másrészt a különböző vizsgálatok keresztelemzése, illetve a később tárgyalt további fejlesztések előfeltétele az adatok összehasonlíthatósága, ezért szükséges volt ezen előfeldolgozási algoritmusok részletes elemzése.

Az algoritmusok összehasonlítására elméletben több lehetőség is van. A korábban alkalmazott módszerek között vezető helyen szerepelt az ismert koncentrációjú RNS hozzáadása a mintákhoz („spike-in”), majd ezen RNS koncentrációjának összehasonlítása az algoritmusok által generált értékekkel ²⁷². Egy másik megközelítés során két különböző mintát különböző arányban keverték össze, majd az egyes méréseket egymással összehasonlították ¹⁶⁰. Bár ezen vizsgálatok lehetővé tették olyan keresztelemzéseket, ahol az egyes módszereket egymáshoz hasonlították ²⁷³, azonban a teljesen ellenőrzött, laboratóriumi megközelítés nem tette lehetővé, hogy a klinikai környezetben várható variáció melletti teljesítményt meg lehessen pontosan becsülni.

Mivel klinikailag lehetetlen teljesen előre definiálni egyes gének várható értékét, ezért ilyen esetben egy független módszer alkalmazása lehet a megoldás. Ennek során a kérdéses gén kifejeződését ezzel a módszerrel is meghatározzuk, majd az egyes algoritmusok által számított eredménnyel összevethetjük. Meg kell jegyeznünk, hogy ebben az esetben nem tudjuk az egyes gének kifejeződését abszolút értelemben meghatározni, azonban az egyes algoritmusok egymáshoz viszonyított teljesítménye kiszámolható.

A 4.10. fejezetben bemutatott saját vizsgálatunk eredményei ¹⁹³ alapján megállapíthatjuk:

- a legjobban teljesítő előfeldolgozási algoritmus a PLIER+16 volt,
- a legelterjedtebb algoritmusok hasonlóan jól teljesítettek.

Az irodalomban az Affymetrix gén chip adatok felhasználása során leggyakrabban alkalmazott algoritmusok között a MAS5, az MBEI és az RMA szerepelnek ¹⁶². Ennek háttérében a korai megjelenésük, a felhasználóbarát programok (MAS5 és MBEI), illetve a gyors számítási idő (RMA) szerepelhetnek. Eredményeink során ezen algoritmusok mindegyike jól teljesített.

Ki kell emelnünk a MAS5 algoritmust, amely abban különbözik valamennyi többi programtól, hogy akár egyetlen gén chip feldolgozására is felhasználható. Fontos hangsúlyozni, hogy az egyetlen mérésen belüli feldolgozás lehetővé teszi, hogy különböző vizsgálatokba bevonjuk ugyanazt a mintát, amely ezen vizsgálatokban mindig ugyanazzal az értékkel fog szerepelni. Bár ezt nem lehet elmondani a többi algoritmusról, azonban azoknál megvan az az elméleti előny, hogy a mintaszám növelésével az algoritmus pontossága is növekedni fog. A MAS5-re vonatkozó korábbi összehasonlító vizsgálatok között volt olyan, ahol nem ért el jó hatásosságot ¹⁹¹, míg más vizsgálatban szintén jól teljesített ²⁷⁴.

A keresztelemzések és rendszerfejlesztések során célunk volt az adatbázisok folyamatos bővítése és továbbfejlesztése, ezért a jól teljesítő algoritmusok között a MAS5-öt választottuk, amely egyetlen gén chipre is elvégezhető. Valamennyi, ezen statisztikai projekt után elvégzett gén chip feldolgozás során a MAS5 algoritmust használtuk.

5.11. Online rendszer fejlesztése az emlőrák prognózisát előrejelző génexpresszió alapú biomarker-jelöltek tesztelésére

A részben a 4.8. fejezet eredményeire épülő, a 4.11. fejezetben részletesen ismertetett bioinformatikai fejlesztéseink²²⁵ alapján a következő megállapításokat tehetjük:

- Létrehoztunk egy, a világhálón keresztül elérhető rendszert, amely lehetővé teszi a rendelkezésre álló transzkriptomikus adatok keresztelemzésével 22277 gén biomarker-szerepének valós időben történő vizsgálatát.
- A lekérdező felületet kiegészítettük szűrőkkel, amelyek alkalmazásával az elemzést egy-egy kiválasztott klinikai csoportban is el lehet végezni.
- A rendszer felhasználásával elvégeztük a korábban leírt legfontosabb túléléssel kapcsolatba hozott biomarkerek keresztelemzését.

A Kaplan-Meier rajzoló első verzióját emlőrákos minták felhasználásával fejlesztettük ki. Ennek az volt az oka, hogy ebben a tumortípusban rendkívül nagyszámú gén chip adat állt rendelkezésre, amely a keresztelemzés alapját adó adatbázis létrehozásához kulcsfontosságú volt.

Általánosságban elmondható, hogy egy biomarker-jelölt azonosítása után különböző lehetőségek állnak rendelkezésre, hogy az azonosított jelölt szerepét igazolni tudjuk. Bár erre az egyes gének direkt vizsgálata, például immunhisztokémia vagy RT-PCR az általánosságban alkalmazott „arany standard”, azonban sokszor a biomarker-jelöltek nagy száma miatt ez nem lehetséges. Ezekben az esetekben szükség van egy előzetes szűrésre, hogy ki tudjuk zárni azokat a géneket, amelyek várhatóan egyáltalán nincsenek a vizsgált betegséggel kapcsolatban. Az általunk kifejlesztett rendszer lényegében pontosan erre készült.

A számítások során Kaplan-Meier elemzés készül az adott génnel. A Kaplan-Meier elemzés nagy előnye, hogy nem csak azokat a betegeket lehet bevonni a számításba, akiknél a vizsgált esemény (kiújulás vagy halál) bekövetkezett. Ehelyett úgynevezett proporcionális túlélést lehet számolni, ahol azon betegek arányát számoljuk, akiknél az adott időpillanatig még nem következett be esemény. Amennyiben mindenkinél van ismert esemény, akkor a Kaplan-Meier görbe nem tér el az empirikus eloszlástól. Klinikai körülmények között azonban ez szinte soha nincs így, hiszen egyes betegek kikerülnek egy vizsgálatból, más

okból is bekövetkezhet az esemény, vagy egyéb olyan tényező jelentkezhethet, ami miatt egy-egy beteget egy adott időpillanatban ki kell zárni. Mindezek mellett a Kaplan-Meier görbe, amely az idő függvényében ábrázolja a követett betegek számának arányát, általánosan is könnyen érthető. Ezek miatt döntöttünk a Kaplan-Meier görbe alkalmazása mellett.

Azonban továbbra is nyitott kérdés volt, hogy pontosan milyen vágópontot alkalmazzunk, amelynek felhasználásával a betegeket két csoportra lehet osztani. Alapból az irodalomban legelterjedtebb megoldást, a medián alkalmazását építettük be a rendszerbe. További vágópontok lehetnek az alsó és a felső negyedek, illetve az alsó és felső harmadok. A vágópont további optimalizálását a következő fejlesztésünkről szóló 4.12. fejezetben részletesen is tárgyaljuk.

Meg kell jegyeznünk a rendszer egyik gyenge pontját, ez pedig a betegek szelektált kiválogatása az eredeti közleményekben. Amennyiben a szerzők egy-egy konkrét csoportot elemeztek, akkor a többi betegre elsősorban költségcsökkentő okból nem végeztek el a gén chip méréseket. Ennek tipikus példája a kezeletlen betegek beválogatása: az adatbázisban a bazális altípusba tartozó betegek majdnem fele nem kapott kemoterápiás kezelést, miközben az aktuális klinikai ajánlások ezen betegcsoport esetében egyértelműen kemoterápiát javasolnak. Bár a betegek ilyen beválogatását meg lehet indokolni (például a kezeletlen betegek vizsgálata jobban lehetővé teszi a betegség lefolyásának megértését, illetve sok olyan vizsgálat volt, amikor a betegek gyűjtése még a jelenleg alkalmazott protokollok bevezetése előtt történt), azonban ez számunkra felvetette a „generalizálhatósági” kérdést. Ezt úgy oldottuk meg, hogy létrehoztunk egy szűrőt, amelybe 500 véletlenszerűen kiválasztott olyan beteget vontunk be, akik esetében az egyes klinikai paraméterek (hormonreceptor-státusz, grade, stb.) megegyeztek az aktuális SEER-ben mért USA-beli prevalenciákkal.

A rendszer használatának bemutatása során az ASCO proliferációs biomarker-jelölteket teszteltük. Az eredmények alapján a TOP2A, a CCND2 és a CCNE2 rendelkeznek a legjobb előrejelző képességgel. Külön ki kell emelnünk, hogy a két cyclin egymáshoz viszonyítva fordított összefüggést mutat a kiújulás mentes túléléssel. Más szóval, míg a CCND2 esetében a magasabb kifejeződés jobb prognózissal társul, addig a CCNE2 esetében ez pont fordítva van. Bár az általunk elvégzett vizsgálatok nem teszik lehetővé, hogy a géneket direkt összehasonlítsuk egymással, ezen eredmények arra utalnak, hogy a két

sejtciklust befolyásoló gén ellentétes hatása mögött ugyanazon szabályozó rendszer kölcsönös befolyásolása állhat.

5.12. A világhálón keresztül elérhető keresztelemző rendszer továbbfejlesztése többgénes osztályozók és optimalizált vágópont számítására

A rendszer továbbfejlesztésének 4.12. fejezetben részletezett eredményei alapján ²⁷⁵ a következő megállapításokat tehetjük:

- A világhálón keresztül elérhető, túléléssel kapcsolatba hozható biomarkerek elemzésére létrehozott rendszert kiegészítettük optimalizált vágópont és többgénes osztályozók alkalmazását lehetővé tevő algoritmusokkal.
- A rendszert elkészítettük petefészekrák biomarkerek tesztelésére is, és az algoritmus felhasználásával elvégeztük 37, az irodalomban leírt biomarker keresztelemzését.

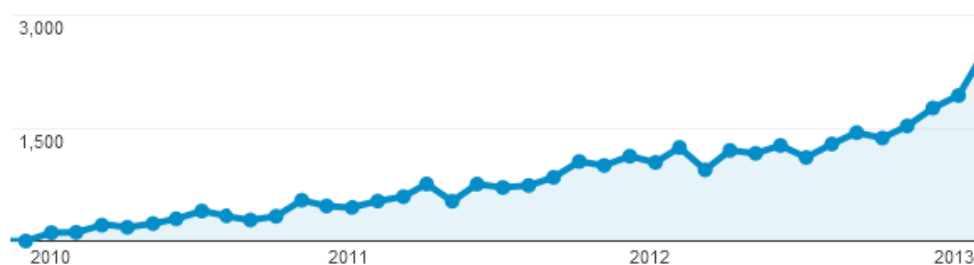
A petefészekrák-adatbázis létrehozása során az 1287 betegből 505 beteg egyetlen adatbázisból (a TCGA-ból ²⁷⁶) származott. A TCGA esetében az eredeti kutatócsoport is vizsgálta egyes gének összefüggését a túléléssel. Az általunk létrehozott rendszerben is lehetővé tettük, hogy a teljes elemzést egy-egy különálló közlemény adatain külön le lehessen futtatni, ezáltal a TCGA-ban azonosított gének direkt keresztelemzése is kivitelezhető.

A rendszer felépítése lehetővé teszi, hogy folyamatosan adjunk az adatbázishoz további betegeket. Az eredetileg közölt 1287 beteget tartalmazó adatbázist is tovább bővítettük, és a jelen értekezésben elkészített ábrák már az újabb, 1464 beteget tartalmazó adatbázis felhasználásával készültek.

A rendszer továbbfejlesztésének következő lépése lehet a klinikai változók szűrőként való alkalmazása helyett kibővített multivariáns Cox regressziós számítások végzése. Jóllehet egy ilyen elemzés során csak azokat a betegeket lehet bevonni a statisztikai számításokba, akiknél minden adat rendelkezésre áll, azonban ezen megközelítés jelentős előnye, hogy az okozati összefüggések mélyebb megértését teszi lehetővé. Bár az értekezés elkészítésének idején még nincs közölt ilyen rendszerünk, már dolgozunk ezen platform létrehozásán.

A adatbázis folyamatos növelésének és az új fejlesztéseknek köszönhetően a Kaplan-Meier rajzoló felhasználtsága folyamatosan növekedett az elmúlt három év folyamán. A

legutolsó mért hónapban (2013 május) az egyedi látogatók száma nagyobb volt, mint az első év folyamán összesen (lásd alábbi ábra).



36. ábra. A www.kmplot.com rendszer havi egyedi látogatói számának alakulása az elmúlt 3 év folyamán a Google Analytics adatai alapján.

5.13. A világhálón keresztül elérhető diagnosztikai rendszer

Az eddig bemutatott bioinformatikai számítások és fejlesztések során a biomarkerek azonosítása volt a cél. A biomarkerek alkalmazására elvégzett, 4.13. fejezetben bemutatott bioinformatikai projektünk eredményeit⁹ a következőkben foglalhatjuk össze:

- A betegek osztályozása gén chip mérések alapján lehetséges és a korábbi diagnosztikai tesztek hatásosságánál jobb.
- A létrehozott, világhálón a www.recurrenceonline.com címen keresztül elérhető rendszer rendkívül rövid elemzési idő mellett (10 perc) lényegében 100%-os reprodukálhatóságot kínál.
- Több különböző diagnosztikai mérés kombinálása egy rendszerbe technikailag megvalósítható.

A többgénes tesztek terjedésének alapja, hogy az egy gént vizsgáló tesztekhez képest a rosszindulatú daganatos sejteket jobban tudják jellemezni. Azonban jellemzően ezen tesztek független igazolása nem lehetséges, mivel egyrészt a betegmintákat nem lehet egy külső csoporthoz is eljuttatni, másrészt pedig az alkalmazott tesztek egy rutin immunhisztokémiához viszonyítva túl bonyolultak ahhoz, hogy alap patológiai intézetben a mérést el lehessen végezni. Az általunk elvégzett fejlesztés során nemcsak egy eszközt

hotzunk létre a gén chipok kiértékelésére, de nagy mintaszámon el is végeztük a rendszer tesztelését.

Az általunk létrehozott kiértékelő rendszer további előnye, hogy maga a számítás mindössze néhány percet vesz igénybe. A hagyományos megközelítés során a betegből vett mintát egy központi laborba kell beküldeni, ahol a mérés és a kiértékelés megtörténik, majd a beteget kezelő belgyógyász illetve onkológus visszakapja a diagnózist. Ez jól működik a receptor-státuszok meghatározása során, azonban a többgénés tesztek esetében a teljes folyamat tipikusan hónapos nagyságrendű időbe kerül, amely alatt a beteg nem kaphatja meg a megfelelő kezelést. Az általunk létrehozott rendszer koncepcióján arra épül, hogy a gén chip mérés azonnal a beteget kezelő klinikán (vagy a legközelebbi genomikus központban) megtörténik, ami a műtét után 24-48 órán belül lehetséges. Ezután az elkészült nyers fájlokat a honlapra feltöltve a diagnózis lényegében további idővesztés nélkül azonnal elkészül.

Ki kell emelni, hogy a létrehozott program hatásossága meghaladja az RT-PCR-re épülő 21 génés teszt hatásosságát. E mögött a gén chipok elvégzéséhez szükséges friss minta megbízhatóbb génexpressziós profilja állhat. Az RT-PCR-hez az RNS-t FFPE mintákból kell izolálni, amelyekben azonban az RNS töredezett, részben lebomlott állapotban van jelen. A műtét során készített friss fagyasztás az eredeti RNS sokkal jobb megőrzését teszi lehetővé.

A rendszer magasabb hatásosságának egy másik forrása lehet a jóval nagyobb betegszám alkalmazása. Miközben a korábbi tesztek 64-233 beteget használtak az eredeti közleményekben^{59, 66, 68, 71}, addig a www.recurrenceonline.com rendszert több mint 2400 gén chip felhasználásával teszteltük. A statisztikai tesztek során sok esetben nagyobb statisztikai erőt lehet elérni nagyobb mintaszám esetén.

Végezetül meg kell említeni, hogy a rendszer szerkezetéből adódóan (vagyis hogy egy program) a továbbfejlesztése viszonylag egyszerű. Figyelembe véve, hogy az emlőrák genetikailag heterogén betegség, ezért további, egyes alcsoportokra fejlesztett osztályozók alkalmazása is hatásosabb lehet, mint az olyan teszt, amelynek felhasználása során lényegében csak két korábbi csoport közül azt vizsgáljuk meg, hogy melyikhez áll a mi konkrét betegünkben kivett rosszindulatú daganatos szövet „genetikai ujjlenyomata” közelebb.

6. A SAJÁT KUTATÁSI EREDMÉNYEK ÖSSZEFOGLALÁSA

1. A kutatások első fő csoportjában a legjobb biomarker-jelölteket azonosítottuk *in vitro* kísérletekben transzkriptom szintű adatok alkalmazásával.

- 1.1. Tizenegy kemoterápiás szerrel szembeni rezisztencia-profilt határoztunk meg 30 rosszindulatú daganatos sejtvonalban, majd gén chip adatok felhasználásával sikeresen tudtunk olyan géneket és génexpressziós mintázatokat azonosítani, amelyek a gyógyszerekkel szembeni rezisztenciát előre tudják jelezni. Azonosítottunk 22 olyan gént, amelyek a multidrog-rezisztenciában is fontos szereppel bírhatnak.
- 1.2. A kemorezisztencia génjeit melanoma malignumban rezisztens sejtvonalpárok felhasználásával azonosítottuk. Fürtelemzés segítségével az egyes vonalak közötti hasonlóságot tudtuk azonosítani, valamint igazoltuk, hogy az irodalomban korábban leírt géneket eredményesen lehet igazolni gén chip mérések felhasználásával.
- 1.3. Rezisztens és érzékeny emlőrák-sejtvonal párok felhasználásával két gyógyszerrel szemben azonosítottuk a rezisztencia legfontosabb génjeit. Egy független betegmintán alkalmazva a génexpressziós mintázatuk alapján általunk kemoterápiára érzékenyként osztályozott, doxorubicin monoterápiával kezelt betegek túlélése több mint 50%-al hosszabb volt, mint a rezisztensként osztályozott betegek túlélése.
- 1.4. A transzkriptomikus adatokkal azonosított proteaszóma fehérjekomplex egyik alegységének oki szerepét igazoltuk kísérletes körülmények között RNS interferencia alkalmazásával. A kezelés és a rezisztencia közötti genom-szintű különbségeket felhasználva azonosítottuk és klinikai mintákban is igazoltuk a PSMB7 gént, mint a doxorubicinnel szembeni rezisztencia egyik oki („driver”) génjét.
- 1.5. A sejtvonalakban történő párhuzamos doxorubicin és paclitaxel kezelés révén sikerült létrehozunk a párhuzamos evolúció modelljét egy konkrét szelekciós nyomással szemben. A 27 sejtvonalban több, korábban leírt mechanizmus aktiválódását tudtuk igazolni a rezisztencia létrejötte során (pld TOP2A és ABCB1) és bizonyítottuk, hogy mindössze 1-2 mechanizmus aktiválódása már elégséges volt ahhoz, hogy a rezisztencia létrejöjjön.
- 1.6. Öt célzott terápiás szerrel szembeni érzékenységet vizsgálatuk 45 sejtvonalban és a relatív rezisztenciát előrejelző biomarkereket azonosítottunk. Vesorákos mintákban igazoltuk az LGALS8, a RAB17 és az EPCAM gének összefüggését a sunitinib-rezisztenciával.

2. A következő kutatások során a cél a korábban azonosított biomarkerek igazolása volt transzkriptom szintű adatok keresztelemzésével és klinikai mintákon való teszteléssel.

- 2.1. A korábbi közleményekben a melanoma pathogenezisével kapcsolatba hozott génlisták összehasonlítása által igazoltuk nagyfokú heterogenitását génexpressziós szinten.
- 2.2. Korábban közölt emlőrákos adatok felhasználásával létrehoztunk egy integrált, 1079 beteg klinikai és génexpressziós adatait tartalmazó adatbázist, amely korábbi prognosztikus biomarkerek keresztelemzését lehetővé teszi.
- 2.3. Petefészekrákok szövettani altípusait meghatározó és a várható túlélést előrejelző gének azonosítására építettünk fel adatbázist 829 betegminta felhasználásával. Az általunk készített gén chip adatbázis kiértékelésével azonosított, a túléléssel legjobban összefüggő gének a PGR, az ÖR2 és a TSPAN8 voltak, amelyek szerepét 64 saját betegben is igazoltuk.

3. A kutatásaink harmadik fő csoportjában expressziós gén chip adatok feldolgozását lehetővé tévő bioinformatikai rendszereket fejlesztettünk.

- 3.1. Először a gén chip adatok előfeldolgozására elérhető algoritmusokat hasonlítottuk össze. A legelterjedtebb algoritmusok (MAS5, PLIER, RMA) hasonlóan jól teljesítettek.
- 3.2. Létrehoztunk egy, a világhálón keresztül elérhető rendszert, amely lehetővé teszi 2977 emlőrákos minta transzkriptomikus adatának keresztelemzésével 22277 gén biomarker-szerepének valós időben történő vizsgálatát. A rendszer felhasználásával elvégeztük a korábban leírt 11 legfontosabb biomarker keresztelemzését.
- 3.3. Ezt a rendszert elkészítettük petefészekrák biomarker-jelöltek tesztelésére is. A program felhasználásával elvégeztük az irodalomban leírt 37 biomarker keresztelemzését.
- 3.4. A biomarkerek alkalmazására elkészítettünk egy világhálón keresztül elérhető diagnosztikai rendszert, amely emlőrákos betegek várható túlélését tudja előre jelezni, valamint ösztrogén- és HER2 receptor státusz meghatározására is alkalmas. A betegek osztályozása gén chip mérések alapján a korábbi diagnosztikai tesztekhez képest magasabb hatásosságot és reprodukálhatóságot kínál.

HIVATKOZÁSOK JEGYZÉKE

1. Levi F, Lucchini F, Negri E, La Vecchia C. Trends in mortality from major cancers in the European Union, including acceding countries, in 2004. *Cancer* 2004;101:2843-50.
2. Kasler M, Otto S. [European and Hungarian national tasks in oncology]. *Magy Onkol* 2008;52:21-33.
3. Gottesman MM, Fojo T, Bates SE. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* 2002;2:48-58.
4. Cuzick J, Forbes JF, Sestak I, Cawthorn S, Hamed H, Holli K, Howell A. Long-term results of tamoxifen prophylaxis for breast cancer--96-month follow-up of the randomized IBIS-I trial. *J Natl Cancer Inst* 2007;99:272-82.
5. Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, McGale P, Pan HC, Taylor C, Wang YC, Dowsett M, Ingle J, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* 2011;378:771-84.
6. Swain SM. Tamoxifen for patients with estrogen receptor-negative breast cancer. *J Clin Oncol* 2001;19:93S-7S.
7. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 1998;351:1451-67.
8. Layfield LJ, Goldstein N, Perkinson KR, Proia AD. Interlaboratory variation in results from immunohistochemical assessment of estrogen receptor status. *Breast J* 2003;9:257-9.
9. Gyorffy B, Benke Z, Lanczky A, Balazs B, Szallasi Z, Timar J, Schafer R. RecurrenceOnline: an online analysis tool to determine breast cancer recurrence and hormone receptor status using microarray data. *Breast Cancer Res Treat* 2012;132:1025-34.
10. Rakha EA, Reis-Filho JS, Ellis IO. Combinatorial biomarker expression in breast cancer. *Breast Cancer Res Treat* 2010;120:293-308.
11. Dunnwald LK, Rossing MA, Li CI. Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. *Breast Cancer Res* 2007;9:R6.
12. Schroth W, Goetz MP, Hamann U, Fasching PA, Schmidt M, Winter S, Fritz P, Simon W, Suman VJ, Ames MM, Safgren SL, Kuffel MJ, et al. Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA* 2009;302:1429-36.
13. Higgins MJ, Stearns V. Pharmacogenetics of endocrine therapy for breast cancer. *Annu Rev Med* 2011;62:281-93.
14. Visvanathan K, Chlebowski RT, Hurley P, Col NF, Ropka M, Collyar D, Morrow M, Runowicz C, Pritchard KI, Hagerty K, Arun B, Garber J, et al. American society of clinical oncology clinical practice guideline update on the use of pharmacologic interventions including tamoxifen, raloxifene, and aromatase inhibition for breast cancer risk reduction. *J Clin Oncol* 2009;27:3235-58.
15. Dawood S, Broglio K, Buzdar AU, Hortobagyi GN, Giordano SH. Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: an institutional-based review. *J Clin Oncol* 2010;28:92-8.
16. Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell* 2000;103:211-25.
17. Olayioye MA, Neve RM, Lane HA, Hynes NE. The ErbB signaling network: receptor heterodimerization in development and cancer. *The EMBO journal* 2000;19:3159-67.

18. Valabrega G, Montemurro F, Aglietta M. Trastuzumab: mechanism of action, resistance and future perspectives in HER2-overexpressing breast cancer. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2007;18:977-84.
19. Bazley LA, Gullick WJ. The epidermal growth factor receptor family. *Endocr Relat Cancer* 2005;12 Suppl 1:S17-27.
20. Wong AL, Lee SC. Mechanisms of Resistance to Trastuzumab and Novel Therapeutic Strategies in HER2-Positive Breast Cancer. *International journal of breast cancer* 2012;2012:415170.
21. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001;344:783-92.
22. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Jr., Davidson NE, Tan-Chiu E, Martino S, Paik S, Kaufman PA, Swain SM, Pisansky TM, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *The New England journal of medicine* 2005;353:1673-84.
23. Brufsky A. Trastuzumab-based therapy for patients with HER2-positive breast cancer: from early scientific development to foundation of care. *American journal of clinical oncology* 2010;33:186-95.
24. Nahta R, Esteva FJ. HER2 therapy: molecular mechanisms of trastuzumab resistance. *Breast cancer research : BCR* 2006;8:215.
25. Scaltriti M, Rojo F, Ocana A, Anido J, Guzman M, Cortes J, Di Cosimo S, Matias-Guiu X, Ramon y Cajal S, Arribas J, Baselga J. Expression of p95HER2, a truncated form of the HER2 receptor, and response to anti-HER2 therapies in breast cancer. *J Natl Cancer Inst* 2007;99:628-38.
26. Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K, Linn SC, Gonzalez-Angulo AM, Stemke-Hale K, Hauptmann M, Beijersbergen RL, Mills GB, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 2007;12:395-402.
27. Nagata Y, Lan KH, Zhou X, Tan M, Esteva FJ, Sahin AA, Klos KS, Li P, Monia BP, Nguyen NT, Hortobagyi GN, Hung MC, et al. PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* 2004;6:117-27.
28. Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T, Aprile G, Kulikov E, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010;376:687-97.
29. Qi J, McTigue MA, Rogers A, Lifshits E, Christensen JG, Janne PA, Engelman JA. Multiple mutations and bypass mechanisms can contribute to development of acquired resistance to MET inhibitors. *Cancer Res* 2011;71:1081-91.
30. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, Lindeman N, Gale CM, Zhao X, Christensen J, Kosaka T, Holmes AJ, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007;316:1039-43.
31. Bean J, Brennan C, Shih JY, Riely G, Viale A, Wang L, Chitale D, Motoi N, Szoke J, Broderick S, Balak M, Chang WC, et al. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci U S A* 2007;104:20932-7.

32. Andre F, Berrada N, Desmedt C. Implication of tumor microenvironment in the resistance to chemotherapy in breast cancer patients. *Curr Opin Oncol* 2010;22:547-51.
33. Weinstein IB, Joe AK. Mechanisms of disease: Oncogene addiction--a rationale for molecular targeting in cancer therapy. *Nat Clin Pract Oncol* 2006;3:448-57.
34. Suda K, Tomizawa K, Osada H, Maehara Y, Yatabe Y, Sekido Y, Mitsudomi T. Conversion from the "oncogene addiction" to "drug addiction" by intensive inhibition of the EGFR and MET in lung cancer with activating EGFR mutation. *Lung Cancer* 2012;76:292-9.
35. Cohen AJ, Mankey J, Wendt W. Response: the claims of public health and public safety. *Sci Pract Perspect* 2003;2:15-7.
36. Gomez HL, Doval DC, Chavez MA, Ang PC, Aziz Z, Nag S, Ng C, Franco SX, Chow LW, Arbushites MC, Casey MA, Berger MS, et al. Efficacy and safety of lapatinib as first-line therapy for ErbB2-amplified locally advanced or metastatic breast cancer. *J Clin Oncol* 2008;26:2999-3005.
37. Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Siebels M, Negrier S, Chevreau C, Solska E, Desai AA, Rolland F, Demkow T, et al. Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med* 2007;356:125-34.
38. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, Oudard S, Negrier S, Szczylik C, Kim ST, Chen I, Bycott PW, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med* 2007;356:115-24.
39. Cohen MH, Williams GA, Sridhara R, Chen G, Pazdur R. FDA drug approval summary: gefitinib (ZD1839) (Iressa) tablets. *Oncologist* 2003;8:303-6.
40. Cameron D, Casey M, Press M, Lindquist D, Pienkowski T, Romieu CG, Chan S, Jagiello-Gruszfeld A, Kaufman B, Crown J, Chan A, Campone M, et al. A phase III randomized comparison of lapatinib plus capecitabine versus capecitabine alone in women with advanced breast cancer that has progressed on trastuzumab: updated efficacy and biomarker analyses. *Breast Cancer Res Treat* 2008;112:533-43.
41. Geyer CE, Forster J, Lindquist D, Chan S, Romieu CG, Pienkowski T, Jagiello-Gruszfeld A, Crown J, Chan A, Kaufman B, Skarlos D, Campone M, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *N Engl J Med* 2006;355:2733-43.
42. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Oudard S, Negrier S, Szczylik C, Pili R, Bjarnason GA, Garcia-del-Muro X, Sosman JA, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2009;27:3584-90.
43. Choueiri TK, Plantade A, Elson P, Negrier S, Ravaud A, Oudard S, Zhou M, Rini BI, Bukowski RM, Escudier B. Efficacy of sunitinib and sorafenib in metastatic papillary and chromophobe renal cell carcinoma. *J Clin Oncol* 2008;26:127-31.
44. Marks JL, Broderick S, Zhou Q, Chitale D, Li AR, Zakowski MF, Kris MG, Rusch VW, Azzoli CG, Seshan VE, Ladanyi M, Pao W. Prognostic and therapeutic implications of EGFR and KRAS mutations in resected lung adenocarcinoma. *J Thorac Oncol* 2008;3:111-6.
45. Dave B, Migliaccio I, Gutierrez MC, Wu MF, Chamness GC, Wong H, Narasanna A, Chakrabarty A, Hilsenbeck SG, Huang J, Rimawi M, Schiff R, et al. Loss of phosphatase and tensin homolog or phosphoinositol-3 kinase activation and response to trastuzumab or lapatinib in human epidermal growth factor receptor 2-overexpressing locally advanced breast cancers. *J Clin Oncol* 2011;29:166-73.
46. Di Nicolantonio F, Martini M, Molinari F, Sartore-Bianchi A, Arena S, Saletti P, De Dosso S, Mazzucchelli L, Frattini M, Siena S, Bardelli A. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol* 2008;26:5705-12.

47. Sartore-Bianchi A, Martini M, Molinari F, Veronese S, Nichelatti M, Artale S, Di Nicolantonio F, Saletti P, De Dosso S, Mazzucchelli L, Frattini M, Siena S, et al. PIK3CA mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. *Cancer Res* 2009;69:1851-7.
48. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q, Ladanyi M, Zakowski MF, Heelan RT, Kris MG, Varmus HE. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2005;2:e17.
49. Jackman DM, Miller VA, Cioffredi LA, Yeap BY, Janne PA, Riely GJ, Ruiz MG, Giaccone G, Sequist LV, Johnson BE. Impact of epidermal growth factor receptor and KRAS mutations on clinical outcomes in previously untreated non-small cell lung cancer patients: results of an online tumor registry of clinical trials. *Clin Cancer Res* 2009;15:5267-73.
50. Huang F, Reeves K, Han X, Fairchild C, Platero S, Wong TW, Lee F, Shaw P, Clark E. Identification of candidate molecular markers predicting sensitivity in solid tumors to dasatinib: rationale for patient selection. *Cancer Res* 2007;67:2226-38.
51. Coldren CD, Helfrich BA, Witta SE, Sugita M, Lapadat R, Zeng C, Baron A, Franklin WA, Hirsch FR, Geraci MW, Bunn PA, Jr. Baseline gene expression predicts sensitivity to gefitinib in non-small cell lung cancer cell lines. *Mol Cancer Res* 2006;4:521-8.
52. Akcakanat A, Zhang L, Tsavachidis S, Meric-Bernstam F. The rapamycin-regulated gene expression signature determines prognosis for breast cancer. *Mol Cancer* 2009;8:75.
53. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418-23.
54. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, Pergamenschikov A, Lee JC, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 2000;24:227-35.
55. Szakacs G, Annereau JP, Lababidi S, Shankavaram U, Arciello A, Bussey KJ, Reinhold W, Guo Y, Kruh GD, Reimers M, Weinstein JN, Gottesman MM. Predicting drug sensitivity and resistance: profiling ABC transporter genes in cancer cells. *Cancer Cell* 2004;6:129-37.
56. Hofmann WK, de Vos S, Elashoff D, Gschaidmeier H, Hoelzer D, Koeffler HP, Ottmann OG. Relation between resistance of Philadelphia-chromosome-positive acute lymphoblastic leukaemia to the tyrosine kinase inhibitor STI571 and gene-expression profiles: a gene-expression study. *Lancet* 2002;359:481-6.
57. Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, Kazemier KM, Pei D, Cheng C, Pui CH, Relling MV, Janka-Schaub GE, Pieters R, et al. Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. *N Engl J Med* 2004;351:533-42.
58. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.
59. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530-6.
60. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869-74.

61. Kihara C, Tsunoda T, Tanaka T, Yamana H, Furukawa Y, Ono K, Kitahara O, Zembutsu H, Yanagawa R, Hirata K, Takagi T, Nakamura Y. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. *Cancer Res* 2001;61:6474-9.
62. Zembutsu H, Ohnishi Y, Tsunoda T, Furukawa Y, Katagiri T, Ueyama Y, Tamaoki N, Nomura T, Kitahara O, Yanagawa R, Hirata K, Nakamura Y. Genome-wide cDNA microarray screening to correlate gene expression profiles with sensitivity of 85 human cancer xenografts to anticancer drugs. *Cancer Res* 2002;62:518-27.
63. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;100:10393-8.
64. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Bernards R, Friend SH. Expression profiling predicts outcome in breast cancer. *Breast Cancer Res* 2003;5:57-8.
65. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC, O'Connell P. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 2003;362:362-9.
66. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817-26.
67. Ma XJ, Salunga R, Dahiya S, Wang W, Carney E, Durbecq V, Harris A, Goss P, Sotiriou C, Erlander M, Sgroi D. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. *Clin Cancer Res* 2008;14:2601-8.
68. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;27:1160-7.
69. Tutt A, Wang A, Rowland C, Gillett C, Lau K, Chew K, Dai H, Kwok S, Ryder K, Shu H, Springall R, Cane P, et al. Risk estimation of distant metastasis in node-negative, estrogen receptor-positive breast cancer patients using an RT-PCR based prognostic expression signature. *BMC Cancer* 2008;8:339.
70. Sanchez-Navarro I, Gamez-Pozo A, Pinto A, Hardisson D, Madero R, Lopez R, San Jose B, Zamora P, Redondo A, Feliu J, Cejas P, Gonzalez Baron M, et al. An 8-gene qRT-PCR-based gene expression score that has prognostic value in early breast cancer. *BMC Cancer* 2010;10:336.
71. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006;98:262-72.
72. Mansel RE, Goyal A, Douglas-Jones A, Woods V, Goyal S, Monypenny I, Sweetland H, Newcombe RG, Jasani B. Detection of breast cancer metastasis in sentinel lymph nodes using intra-operative real time GeneSearch BLN Assay in the operating room: results of the Cardiff study. *Breast Cancer Res Treat* 2009;115:595-600.
73. Liu R, Wang X, Chen GY, Dalerba P, Gurney A, Hoey T, Sherlock G, Lewicki J, Shedden K, Clarke MF. The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* 2007;356:217-26.
74. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, Brown PO. Gene expression signature of fibroblast serum

response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* 2004;2:E7.

75. Davis LM, Harris C, Tang L, Doherty P, Hrabec P, Sakai Y, Bocklage T, Doeden K, Hall B, Alsobrook J, Rabinowitz I, Williams TM, et al. Amplification patterns of three genomic regions predict distant recurrence in breast carcinoma. *J Mol Diagn* 2007;9:327-36.

76. Ring BZ, Seitz RS, Beck R, Shasteen WJ, Tarr SM, Cheang MC, Yoder BJ, Budd GT, Nielsen TO, Hicks DG, Estopinal NC, Ross DT. Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;24:3039-47.

77. Gyorffy B, Schafer R. Meta-analysis of gene expression profiles related to relapse-free survival in 1,079 breast cancer patients. *Breast Cancer Res Treat* 2009;118:433-41.

78. Mamounas EP, Tang G, Fisher B, Paik S, Shak S, Costantino JP, Watson D, Geyer CE, Jr., Wickerham DL, Wolmark N. Association between the 21-gene recurrence score assay and risk of locoregional recurrence in node-negative, estrogen receptor-positive breast cancer: results from NSABP B-14 and NSABP B-20. *J Clin Oncol* 2010;28:1677-83.

79. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE, Jr., et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006;24:3726-34.

80. Oratz R, Paul D, Cohn AL, Sedlacek SM. Impact of a commercial reference laboratory test recurrence score on decision making in early-stage breast cancer. *J Oncol Pract* 2007;3:182-6.

81. Asad J, Jacobson AF, Estabrook A, Smith SR, Boolbol SK, Feldman SM, Osborne MP, Boachie-Adjei K, Twardzik W, Tartter PI. Does oncotype DX recurrence score affect the management of patients with early-stage breast cancer? *Am J Surg* 2008;196:527-9.

82. Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. *Nat Rev Cancer* 2004;4:309-14.

83. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006;98:1183-92.

84. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;365:488-92.

85. Whittaker S. Adjuvant therapy in melanoma. *Clin Exp Dermatol* 2000;25:497-502.

86. McGovern VJ, Shaw HM, Milton GW. Histogenesis of malignant melanoma with an adjacent component of the superficial spreading type. *Pathology* 1985;17:251-4.

87. Middleton MR, Lorigan P, Owen J, Ashcroft L, Lee SM, Harper P, Thatcher N. A randomized phase III study comparing dacarbazine, BCNU, cisplatin and tamoxifen with dacarbazine and interferon in advanced melanoma. *Br J Cancer* 2000;82:1158-62.

88. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57-70.

89. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 2000;406:536-40.

90. Shih Ie M, Kurman RJ. Molecular pathogenesis of ovarian borderline tumors: new insights and old challenges. *Clin Cancer Res* 2005;11:7273-9.

91. Bonome T, Lee JY, Park DC, Radonovich M, Pise-Masison C, Brady J, Gardner GJ, Hao K, Wong WH, Barrett JC, Lu KH, Sood AK, et al. Expression profiling of serous low malignant potential, low-grade, and high-grade tumors of the ovary. *Cancer Research* 2005;65:10602-12.

92. Zorn KK, Bonome T, Gangi L, Chandramouli GV, Awtrey CS, Gardner GJ, Barrett JC, Boyd J, Birrer MJ. Gene expression profiles of serous, endometrioid, and clear cell subtypes of ovarian and endometrial cancer. *Clin Cancer Res* 2005;11:6422-30.
93. Bayani J, Brenton JD, Macgregor PF, Beheshti B, Albert M, Nallainathan D, Karaskova J, Rosen B, Murphy J, Laframboise S, Zanke B, Squire JA. Parallel analysis of sporadic primary ovarian carcinomas by spectral karyotyping, comparative genomic hybridization, and expression microarrays. *Cancer Research* 2002;62:3466-76.
94. Mok SC, Chao J, Skates S, Wong K, Yiu GK, Muto MG, Berkowitz RS, Cramer DW. Prostin, a potential serum marker for ovarian cancer: identification through microarray technology. *J Natl Cancer Inst* 2001;93:1458-64.
95. Ono K, Tanaka T, Tsunoda T, Kitahara O, Kihara C, Okamoto A, Ochiai K, Takagi T, Nakamura Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Research* 2000;60:5007-11.
96. Tonin PN, Hudson TJ, Rodier F, Bossolasco M, Lee PD, Novak J, Manderson EN, Provencher D, Mes-Masson AM. Microarray analysis of gene expression mirrors the biology of an ovarian cancer model. *Oncogene* 2001;20:6617-26.
97. Welsh JB, Zarrinkar PP, Sapinoso LM, Kern SG, Behling CA, Monk BJ, Lockhart DJ, Burger RA, Hampton GM. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc Natl Acad Sci U S A* 2001;98:1176-81.
98. Zhang XY, Li XP, Lai J, Feng J. [Screening for ovarian cancer-associated genes with cDNA microarrays]. *Ai Zheng* 2003;22:943-7.
99. Bignotti E, Tassi RA, Calza S, Ravaggi A, Romani C, Rossi E, Falchetti M, Odicino FE, Pecorelli S, Santin AD. Differential gene expression profiles between tumor biopsies and short-term primary cultures of ovarian serous carcinomas: identification of novel molecular biomarkers for early diagnosis and therapy. *Gynecol Oncol* 2006;103:405-16.
100. Crijns AP, Fehrmann RS, de Jong S, Gerbens F, Meersma GJ, Klip HG, Hollema H, Hofstra RM, te Meerman GJ, de Vries EG, van der Zee AG. Survival-related profile, pathways, and transcription factors in ovarian cancer. *PLoS Med* 2009;6:e24.
101. Donninger H, Bonome T, Radonovich M, Pise-Masison CA, Brady J, Shih JH, Barrett JC, Birrer MJ. Whole genome expression profiling of advance stage papillary serous ovarian cancer reveals activated pathways. *Oncogene* 2004;23:8065-77.
102. Heinzelmann-Schwarz VA, Gardiner-Garden M, Henshall SM, Scurry J, Scolyer RA, Davies MJ, Heinzelmann M, Kalish LH, Bali A, Kench JG, Edwards LS, Vanden Bergh PM, et al. Overexpression of the cell adhesion molecules DDR1, Claudin 3, and Ep-CAM in metaplastic ovarian epithelium and ovarian cancer. *Clin Cancer Res* 2004;10:4427-36.
103. Klinck R, Bramard A, Inkel L, Dufresne-Martin G, Gervais-Bird J, Madden R, Paquet ER, Koh C, Venables JP, Prinos P, Jilaveanu-Pelms M, Wellinger R, et al. Multiple alternative splicing markers for ovarian cancer. *Cancer Research* 2008;68:657-63.
104. Lancaster JM, Dressman HK, Whitaker RS, Havrilesky L, Gray J, Marks JR, Nevins JR, Berchuck A. Gene expression patterns that characterize advanced stage serous ovarian cancers. *J Soc Gynecol Investig* 2004;11:51-9.
105. Le Page C, Ouellet V, Madore J, Ren F, Hudson TJ, Tonin PN, Provencher DM, Mes-Masson AM. Gene expression profiling of primary cultures of ovarian epithelial cells identifies novel molecular classifiers of ovarian cancer. *Br J Cancer* 2006;94:436-45.
106. Li J, Olson LM, Zhang Z, Li L, Bidder M, Nguyen L, Pfeifer J, Rader JS. Differential display identifies overexpression of the USP36 gene, encoding a deubiquitinating enzyme, in ovarian cancer. *Int J Med Sci* 2008;5:133-42.

107. Mougeot JL, Bahrani-Mostafavi Z, Vachris JC, McKinney KQ, Gurlov S, Zhang J, Naumann RW, Higgins RV, Hall JB. Gene expression profiling of ovarian tissues for determination of molecular pathways reflective of tumorigenesis. *J Mol Biol* 2006;358:310-29.
108. Park DC, Yeo SG, Wilson MR, Yerbury JJ, Kwong J, Welch WR, Choi YK, Birrer MJ, Mok SC, Wong KK. Clusterin interacts with paclitaxel and confer paclitaxel resistance in ovarian cancer. *Neoplasia* 2008;10:964-U43.
109. Santin AD, Zhan F, Bellone S, Palmieri M, Cane S, Gokden M, Roman JJ, O'Brien TJ, Tian E, Cannon MJ, Shaughnessy J, Jr., Pecorelli S. Discrimination between uterine serous papillary carcinomas and ovarian serous papillary tumours by gene expression profiling. *Br J Cancer* 2004;90:1814-24.
110. Sunde JS, Donninger H, Wu K, Johnson ME, Pestell RG, Rose GS, Mok SC, Brady J, Bonome T, Birrer MJ. Expression profiling identifies altered expression of genes that contribute to the inhibition of transforming growth factor-beta signaling in ovarian cancer. *Cancer Research* 2006;66:8404-12.
111. Zhang L, Huang J, Yang N, Greshock J, Liang S, Hasegawa K, Giannakakis A, Poulos N, O'Brien-Jenkins A, Katsaros D, Butzow R, Weber BL, et al. Integrative genomic analysis of phosphatidylinositol 3'-kinase family identifies PIK3R3 as a potential therapeutic target in epithelial ovarian cancer. *Clin Cancer Res* 2007;13:5314-21.
112. Zhang L, Huang J, Yang N, Liang S, Barchetti A, Giannakakis A, Cadungog MG, O'Brien-Jenkins A, Massobrio M, Roby KF, Katsaros D, Gimotty P, et al. Integrative genomic analysis of protein kinase C (PKC) family identifies PKC α as a biomarker and potential oncogene in ovarian carcinoma. *Cancer Research* 2006;66:4627-35.
113. Zhang X, Feng J, Cheng Y, Yao Y, Ye X, Fu T, Cheng H. Characterization of differentially expressed genes in ovarian cancer by cDNA microarrays. *Int J Gynecol Cancer* 2005;15:50-7.
114. Moreno-Bueno G, Sanchez-Estevez C, Cassia R, Rodriguez-Perales S, Diaz-Uriarte R, Dominguez O, Hardisson D, Andujar M, Prat J, Matias-Guiu X, Cigudosa JC, Palacios J. Differential gene expression profile in endometrioid and nonendometrioid endometrial carcinoma: STK15 is frequently overexpressed and amplified in nonendometrioid carcinomas. *Cancer Research* 2003;63:5697-702.
115. Zheng M, Simon R, Kononen J, Sauter G, Mihatsch MJ, Moch H. [Analysis of gene expression profiles among 3 epithelial ovarian tumor subtypes using cDNA and tissue microarrays]. *Ai Zheng* 2004;23:771-6.
116. Sugimura M, Sagae S, Ishioka S, Nishioka Y, Tsukada K, Kudo R. Mechanisms of paclitaxel-induced apoptosis in an ovarian cancer cell line and its paclitaxel-resistant clone. *Oncology* 2004;66:53-61.
117. Lamendola DE, Duan Z, Yusuf RZ, Seiden MV. Molecular description of evolving paclitaxel resistance in the SKOV-3 human ovarian carcinoma cell line. *Cancer Research* 2003;63:2200-5.
118. Selvanayagam ZE, Cheung TH, Wei N, Vittal R, Lo KW, Yeo W, Kita T, Ravatn R, Chung TK, Wong YF, Chin KV. Prediction of chemotherapeutic response in ovarian cancer with DNA microarray expression profiling. *Cancer Genet Cytogenet* 2004;154:63-6.
119. Macleod K, Mullen P, Sewell J, Rabiasz G, Lawrie S, Miller E, Smyth JF, Langdon SP. Altered ErbB receptor signaling and gene expression in cisplatin-resistant ovarian cancer. *Cancer Research* 2005;65:6789-800.
120. Samimi G, Manorek G, Castel R, Breaux JK, Cheng TC, Berry CC, Los G, Howell SB. cDNA microarray-based identification of genes and pathways associated with oxaliplatin resistance. *Cancer Chemother Pharmacol* 2005;55:1-11.

121. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, Olson JA, Jr., Marks JR, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439:353-7.
122. Cheng TC, Manorek G, Samimi G, Lin X, Berry CC, Howell SB. Identification of genes whose expression is associated with cisplatin resistance in human ovarian carcinoma cells. *Cancer Chemother Pharmacol* 2006;58:384-95.
123. Xu S, Mou H, Lu G, Zhu C, Yang Z, Gao Y, Lou H, Liu X, Cheng Y, Yang W. Gene expression profile differences in high and low metastatic human ovarian cancer cell lines by gene chip. *Chin Med J (Engl)* 2002;115:36-41.
124. Adib TR, Henderson S, Perrett C, Hewitt D, Bourmpoulia D, Ledermann J, Boshoff C. Predicting biomarkers for ovarian cancer using gene-expression microarrays. *Br J Cancer* 2004;90:686-92.
125. De Cecco L, Marchionni L, Gariboldi M, Reid JF, Lagonigro MS, Caramuta S, Ferrario C, Bussani E, Mezzanzanica D, Turatti F, Delia D, Daidone MG, et al. Gene expression profiling of advanced ovarian cancer: characterization of a molecular signature involving fibroblast growth factor 2. *Oncogene* 2004;23:8171-83.
126. Ouellet V, Provencher DM, Maugard CM, Le Page C, Ren F, Lussier C, Novak J, Ge B, Hudson TJ, Tonin PN, Mes-Masson AM. Discrimination between serous low malignant potential and invasive epithelial ovarian tumors using molecular profiling. *Oncogene* 2005;24:4672-87.
127. Motamed-Khorasani A, Jurisica I, Letarte M, Shaw PA, Parkes RK, Zhang X, Evangelou A, Rosen B, Murphy KJ, Brown TJ. Differentially androgen-modulated genes in ovarian epithelial cells from BRCA mutation carriers and control patients predict ovarian cancer survival and disease progression. *Oncogene* 2007;26:198-214.
128. Denkert C, Budczies J, Darb-Esfahani S, Gyorffy B, Sehouli J, Konsgen D, Zeillinger R, Weichert W, Noske A, Buckendahl AC, Muller BM, Dietel M, et al. A prognostic gene expression index in ovarian cancer - validation across different independent data sets. *J Pathol* 2009;218:273-80.
129. Bali A, O'Brien PM, Edwards LS, Sutherland RL, Hacker NF, Henshall SM. Cyclin D1, p53, and p21Waf1/Cip1 expression is predictive of poor clinical outcome in serous epithelial ovarian cancer. *Clin Cancer Res* 2004;10:5168-77.
130. Ferrandina G, Stoler A, Fagotti A, Fanfani F, Sacco R, De Pasqua A, Mancuso S, Scambia G. p21WAF1/CIP1 protein expression in primary ovarian cancer. *Int J Oncol* 2000;17:1231-5.
131. Plisiecka-Halasa J, Karpinska G, Szymanska T, Ziolkowska I, Madry R, Timorek A, Debniak J, Ulanska M, Jedryka M, Chudecka-Glaz A, Klimek M, Rembiszewska A, et al. P21WAF1, P27KIP1, TP53 and C-MYC analysis in 204 ovarian carcinomas treated with platinum-based regimens. *Ann Oncol* 2003;14:1078-85.
132. Skirnisdottir I, Sorbe B, Seidal T. P53, bcl-2, and bax: their relationship and effect on prognosis in early stage epithelial ovarian carcinoma. *Int J Gynecol Cancer* 2001;11:147-58.
133. Tai YT, Lee S, Niloff E, Weisman C, Strobel T, Cannistra SA. BAX protein expression and clinical outcome in epithelial ovarian cancer. *J Clin Oncol* 1998;16:2583-90.
134. Brustmann H. Immunohistochemical detection of human telomerase reverse transcriptase (hTERT) and c-kit in serous ovarian carcinoma: a clinicopathologic study. *Gynecol Oncol* 2005;98:396-402.
135. Sui L, Dong Y, Ohno M, Watanabe Y, Sugimoto K, Tokuda M. Survivin expression and its correlation with cell proliferation and prognosis in epithelial ovarian tumors. *Int J Oncol* 2002;21:315-20.

136. Hefler LA, Zeillinger R, Grimm C, Sood AK, Cheng WF, Gadducci A, Tempfer CB, Reinthaller A. Preoperative serum vascular endothelial growth factor as a prognostic parameter in ovarian cancer. *Gynecol Oncol* 2006;103:512-7.

137. Buttitta F, Marchetti A, Gadducci A, Pellegrini S, Morganti M, Carnicelli V, Cosio S, Galletti O, Genazzani AR, Bevilacqua G. p53 alterations are predictive of chemoresistance and aggressiveness in ovarian carcinomas: a molecular and immunohistochemical study. *Br J Cancer* 1997;75:230-5.

138. Reles A, Wen WH, Schmider A, Gee C, Runnebaum IB, Kilian U, Jones LA, El-Naggar A, Minguillon C, Schonborn I, Reich O, Kreienberg R, et al. Correlation of p53 mutations with resistance to platinum-based chemotherapy and shortened survival in ovarian cancer. *Clin Cancer Res* 2001;7:2984-97.

139. Diamandis EP, Scorilas A, Fracchioli S, Van Gramberen M, De Bruijn H, Henrik A, Soosaipillai A, Grass L, Yousef GM, Stenman UH, Massobrio M, Van Der Zee AG, et al. Human kallikrein 6 (hK6): a new potential serum biomarker for diagnosis and prognosis of ovarian carcinoma. *J Clin Oncol* 2003;21:1035-43.

140. Luo LY, Katsaros D, Scorilas A, Fracchioli S, Piccinno R, Rigault de la Longrais IA, Howarth DJ, Diamandis EP. Prognostic value of human kallikrein 10 expression in epithelial ovarian carcinoma. *Clin Cancer Res* 2001;7:2372-9.

141. Scambia G, Testa U, Benedetti Panici P, Foti E, Martucci R, Gadducci A, Perillo A, Facchini V, Peschle C, Mancuso S. Prognostic significance of interleukin 6 serum levels in patients with ovarian cancer. *Br J Cancer* 1995;71:354-6.

142. Schmider-Ross A, Pirsig O, Gottschalk E, Denkert C, Lichtenegger W, Reles A. Cyclin-dependent kinase inhibitors CIP1 (p21) and KIP1 (p27) in ovarian cancer. *J Cancer Res Clin Oncol* 2006;132:163-70.

143. Masciullo V, Ferrandina G, Pucci B, Fanfani F, Lovergine S, Palazzo J, Zannoni G, Mancuso S, Scambia G, Giordano A. p27Kip1 expression is associated with clinical outcome in advanced epithelial ovarian cancer: multivariate analysis. *Clin Cancer Res* 2000;6:4816-22.

144. Korkolopoulou P, Vassilopoulos I, Konstantinidou AE, Zorzos H, Patsouris E, Agapitos E, Davaris P. The combined evaluation of p27Kip1 and Ki-67 expression provides independent information on overall survival of ovarian carcinoma patients. *Gynecol Oncol* 2002;85:404-14.

145. Newcomb EW, Sosnow M, Demopoulos RI, Zeleniuch-Jacquotte A, Sorich J, Speyer JL. Expression of the cell cycle inhibitor p27KIP1 is a new prognostic marker associated with survival in epithelial ovarian tumors. *Am J Pathol* 1999;154:119-25.

146. Barbieri F, Lorenzi P, Ragni N, Schettini G, Bruzzo C, Pedulla F, Alama A. Overexpression of cyclin D1 is associated with poor survival in epithelial ovarian cancer. *Oncology* 2004;66:310-5.

147. Levidou G, Korkolopoulou P, Thymara I, Vassilopoulos I, Saetta AA, Gakiopoulou H, Konstantinidou A, Kairi-Vassilatou E, Pavlakis K, Patsouris E. Expression and prognostic significance of cyclin D3 in ovarian adenocarcinomas. *Int J Gynecol Pathol* 2007;26:410-7.

148. Sui L, Dong Y, Ohno M, Sugimoto K, Tai Y, Hando T, Tokuda M. Implication of malignancy and prognosis of p27(kip1), Cyclin E, and Cdk2 expression in epithelial ovarian tumors. *Gynecol Oncol* 2001;83:56-63.

149. Farley J, Smith LM, Darcy KM, Sobel E, O'Connor D, Henderson B, Morrison LE, Birrer MJ. Cyclin E expression is a significant predictor of survival in advanced, suboptimally debulked ovarian epithelial cancers: a Gynecologic Oncology Group study. *Cancer Res* 2003;63:1235-41.

150. Rosen DG, Yang G, Deavers MT, Malpica A, Kavanagh JJ, Mills GB, Liu J. Cyclin E expression is correlated with tumor progression and predicts a poor prognosis in patients with ovarian carcinoma. *Cancer* 2006;106:1925-32.
151. Bedrosian I, Lee C, Tucker SL, Palla SL, Lu K, Keyomarsi K. Cyclin E-associated kinase activity predicts response to platinum-based chemotherapy. *Clin Cancer Res* 2007;13:4800-6.
152. Materna V, Surowiak P, Markwitz E, Spaczynski M, Drag-Zalesinska M, Zabel M, Lage H. Expression of factors involved in regulation of DNA mismatch repair- and apoptosis pathways in ovarian cancer patients. *Oncol Rep* 2007;17:505-16.
153. Psyrri A, Yu Z, Bamias A, Weinberger PM, Markakis S, Kowalski D, Camp RL, Rimm DL, Dimopoulos MA. Evaluation of the prognostic value of cellular inhibitor of apoptosis protein in epithelial ovarian cancer using automated quantitative protein analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1179-83.
154. Skirnisdottir I, Seidal T, Sorbe B. A new prognostic model comprising p53, EGFR, and tumor grade in early stage epithelial ovarian carcinoma and avoiding the problem of inaccurate surgical staging. *Int J Gynecol Cancer* 2004;14:259-70.
155. Psyrri A, Kassar M, Yu Z, Bamias A, Weinberger PM, Markakis S, Kowalski D, Camp RL, Rimm DL, Dimopoulos MA. Effect of epidermal growth factor receptor expression level on survival in patients with epithelial ovarian cancer. *Clin Cancer Res* 2005;11:8637-43.
156. Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, Parker HL. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001;19:980-91.
157. Olivotto IA, Bajdik CD, Ravdin PM, Speers CH, Coldman AJ, Norris BD, Davis GJ, Chia SK, Gelmon KA. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 2005;23:2716-25.
158. Colozza M, Azambuja E, Cardoso F, Sotiriou C, Larsimont D, Piccart MJ. Proliferative markers as prognostic and predictive tools in early breast cancer: where are we now? *Ann Oncol* 2005;16:1723-39.
159. Draghici S, Khatri P, Eklund AC, Szallasi Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet* 2006;22:101-9.
160. Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY, Luo Y, Sun YA, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006;24:1151-61.
161. Darb-Esfahani S, Wirtz RM, Sinn BV, Budczies J, Noske A, Weichert W, Faggad A, Scharff S, Sehouli J, Oskay-Ozcelik G, Zamagni C, De Iaco P, et al. Estrogen receptor 1 mRNA is a prognostic factor in ovarian carcinoma: determination by kinetic PCR in formalin-fixed paraffin-embedded tissue. *Endocr Relat Cancer* 2009;16:1229-39.
162. Expression profiling--best practices for data generation and interpretation in clinical trials. *Nat Rev Genet* 2004;5:229-37.
163. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 2002;18 Suppl 1:S96-104.
164. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31:e15.
165. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;99:6567-72.

166. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116-21.
167. Fogh J. Cultivation, characterization, and identification of human tumor cells with emphasis on kidney, testis, and bladder tumors. *Natl Cancer Inst Monogr* 1978:5-9.
168. Kern MA, Helmbach H, Artuc M, Karmann D, Jurgovsky K, Schadendorf D. Human melanoma cell lines selected in vitro displaying various levels of drug resistance against cisplatin, fotemustine, vindesine or etoposide: modulation of proto-oncogene expression. *Anticancer Res* 1997;17:4359-70.
169. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747-52.
170. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-11.
171. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics* 2002;18:207-8.
172. Materna V, Liedert B, Thomale J, Lage H. Protection of platinum-DNA adduct formation and reversal of cisplatin resistance by anti-MRP2 hammerhead ribozymes in human cancer cells. *Int J Cancer* 2005;115:393-402.
173. Dietel M, Arps H, Lage H, Niendorf A. Membrane vesicle formation due to acquired mitoxantrone resistance in human gastric carcinoma cell line EPG85-257. *Cancer Res* 1990;50:6100-6.
174. Lage H, Jordan A, Scholz R, Dietel M. Thermosensitivity of multidrug-resistant human gastric and pancreatic carcinoma cells. *Int J Hyperthermia* 2000;16:291-303.
175. Lage H, Dietel M. Multiple mechanisms confer different drug-resistant phenotypes in pancreatic carcinoma cells. *J Cancer Res Clin Oncol* 2002;128:349-57.
176. Chen TR, Drabkowski D, Hay RJ, Macy M, Peterson W, Jr. WiDr is a derivative of another colon adenocarcinoma cell line, HT-29. *Cancer Genet Cytogenet* 1987;27:125-34.
177. Soule HD, Vazquez J, Long A, Albert S, Brennan M. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst* 1973;51:1409-16.
178. Cailleau R, Young R, Olive M, Reeves WJ, Jr. Breast tumor cell lines from pleural effusions. *J Natl Cancer Inst* 1974;53:661-74.
179. Karukstis KK, Thompson EH, Whiles JA, Rosenfeld RJ. Deciphering the fluorescence signature of daunomycin and doxorubicin. *Biophys Chem* 1998;73:249-63.
180. Shannon CE. The mathematical theory of communication. 1963. *MD Comput* 1997;14:306-17.
181. Breitling R, Herzyk P. Rank-based methods as a non-parametric alternative of the T-statistic for the analysis of biological microarray data. *J Bioinform Comput Biol* 2005;3:1171-89.
182. Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 2004;573:83-92.
183. Jeffery IB, Higgins DG, Culhane AC. Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. *BMC Bioinformatics* 2006;7:359.
184. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004;2:E108.

185. Galamb O, Gyorffy B, Sipos F, Spisak S, Nemeth AM, Miheller P, Tulassay Z, Dinya E, Molnar B. Inflammation, adenoma and cancer: objective classification of colon biopsy specimens with gene expression signature. *Dis Markers* 2008;25:1-16.
186. Chen Z, McGee M, Liu Q, Scheuermann RH. A distribution free summarization method for Affymetrix GeneChip arrays. *Bioinformatics* 2007;23:321-7.
187. Hochreiter S, Clevert DA, Obermayer K. A new summarization method for Affymetrix probe level data. *Bioinformatics* 2006;22:943-9.
188. Wu Z, Irizarry RA. Stochastic models inspired by hybridization theory for short oligonucleotide arrays. *J Comput Biol* 2005;12:882-93.
189. Hubbell E, Liu WM, Mei R. Robust estimators for expression analysis. *Bioinformatics* 2002;18:1585-92.
190. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 2001;98:31-6.
191. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4:249-64.
192. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;20:307-15.
193. Gyorffy B, Molnar B, Lage H, Szallasi Z, Eklund AC. Evaluation of microarray preprocessing algorithms based on concordance with RT-PCR in clinical samples. *PLoS One* 2009;4:e5645.
194. Sims AH, Smethurst GJ, Hey Y, Okoniewski MJ, Pepper SD, Howell A, Miller CJ, Clarke RB. The removal of multiplicative, systematic bias allows integration of breast cancer gene expression datasets - improving meta-analysis and prediction of prognosis. *BMC Med Genomics* 2008;1:42.
195. Gong Y, Yan K, Lin F, Anderson K, Sotiriou C, Andre F, Holmes FA, Valero V, Booser D, Pippin JE, Jr., Vukelja S, Gomez H, et al. Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. *Lancet Oncol* 2007;8:203-11.
196. Li LF, Xu XJ, Zhao Y, Liu ZB, Shen ZZ, Jin WR, Shao ZM. Integrated gene expression profile predicts prognosis of breast cancer patients. *Breast Cancer Res Treat* 2009;113:231-7.
197. Wittig R, Nessling M, Will RD, Mollenhauer J, Salowsky R, Munstermann E, Schick M, Helmbach H, Gschwendt B, Korn B, Kioschis P, Lichter P, et al. Candidate genes for cross-resistance against DNA-damaging drugs. *Cancer Res* 2002;62:6698-705.
198. Lage H, Christmann M, Kern MA, Dietel M, Pick M, Kaina B, Schadendorf D. Expression of DNA repair proteins hMSH2, hMSH6, hMLH1, O6-methylguanine-DNA methyltransferase and N-methylpurine-DNA glycosylase in melanoma cells with acquired drug resistance. *Int J Cancer* 1999;80:744-50.
199. Grottke C, Mantwill K, Dietel M, Schadendorf D, Lage H. Identification of differentially expressed genes in human melanoma cells with acquired resistance to various antineoplastic drugs. *Int J Cancer* 2000;88:535-46.
200. Ross DD, Yang W, Abruzzo LV, Dalton WS, Schneider E, Lage H, Dietel M, Greenberger L, Cole SP, Doyle LA. Atypical multidrug resistance: breast cancer resistance protein messenger RNA expression in mitoxantrone-selected cell lines. *J Natl Cancer Inst* 1999;91:429-33.
201. Sinha P, Hutter G, Kottgen E, Dietel M, Schadendorf D, Lage H. Search for novel proteins involved in the development of chemoresistance in colorectal cancer and

fibrosarcoma cells in vitro using two-dimensional electrophoresis, mass spectrometry and microsequencing. *Electrophoresis* 1999;20:2961-9.

202. Talantov D, Mazumder A, Yu JX, Briggs T, Jiang Y, Backus J, Atkins D, Wang Y. Novel genes associated with malignant melanoma but not benign melanocytic lesions. *Clin Cancer Res* 2005;11:7234-42.

203. Schaefer KL, Wai DH, Poremba C, Korsching E, van Valen F, Ozaki T, Boecker W, Dockhorn-Dworniczak B. Characterization of the malignant melanoma of soft-parts cell line GG-62 by expression analysis using DNA microarrays. *Virchows Arch* 2002;440:476-84.

204. Hoek K, Rimm DL, Williams KR, Zhao H, Ariyan S, Lin A, Kluger HM, Berger AJ, Cheng E, Trombetta ES, Wu T, Niinobe M, et al. Expression profiling reveals novel pathways in the transformation of melanocytes to melanomas. *Cancer Res* 2004;64:5270-82.

205. Segal NH, Pavlidis P, Noble WS, Antonescu CR, Viale A, Wesley UV, Busam K, Gallardo H, DeSantis D, Brennan MF, Cordon-Cardo C, Wolchok JD, et al. Classification of clear-cell sarcoma as a subtype of melanoma by genomic profiling. *J Clin Oncol* 2003;21:1775-81.

206. Schaefer KL, Brachwitz K, Wai DH, Braun Y, Diallo R, Korsching E, Eisenacher M, Voss R, Van Valen F, Baer C, Selle B, Spahn L, et al. Expression profiling of t(12;22) positive clear cell sarcoma of soft tissue cell lines reveals characteristic up-regulation of potential new marker genes including ERBB3. *Cancer Res* 2004;64:3395-405.

207. Nambiar S, Mirmohammadsadegh A, Doroudi R, Gustrau A, Marini A, Roeder G, Ruzicka T, Hengge UR. Signaling networks in cutaneous melanoma metastasis identified by complementary DNA microarrays. *Arch Dermatol* 2005;141:165-73.

208. de Wit NJ, Rijntjes J, Diepstra JH, van Kuppevelt TH, Weidle UH, Ruiter DJ, van Muijen GN. Analysis of differential gene expression in human melanocytic tumour lesions by custom made oligonucleotide arrays. *Br J Cancer* 2005;92:2249-61.

209. Xu T, Shu CT, Purdom E, Dang D, Ilsley D, Guo Y, Weber J, Holmes SP, Lee PP. Microarray analysis reveals differences in gene expression of circulating CD8(+) T cells in melanoma patients and healthy donors. *Cancer Res* 2004;64:3661-7.

210. Weeraratna AT, Becker D, Carr KM, Duray PH, Rosenblatt KP, Yang S, Chen Y, Bittner M, Strausberg RL, Riggins GJ, Wagner U, Kallioniemi OP, et al. Generation and analysis of melanoma SAGE libraries: SAGE advice on the melanoma transcriptome. *Oncogene* 2004;23:2264-74.

211. Huang E, West M, Nevins JR. Gene expression profiling for prediction of clinical characteristics of breast cancer. *Recent Prog Horm Res* 2003;58:55-73.

212. Pawitan Y, Bjohle J, Amler L, Borg AL, Eghazi S, Hall P, Han X, Holmberg L, Huang F, Klaar S, Liu ET, Miller L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005;7:R953-64.

213. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;365:671-9.

214. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. *Genome Biol* 2007;8:R157.

215. Korkola JE, Blaveri E, DeVries S, Moore DH, 2nd, Hwang ES, Chen YY, Estep AL, Chew KL, Jensen RH, Waldman FM. Identification of a robust gene signature that predicts breast cancer outcome in independent data sets. *BMC Cancer* 2007;7:61.

216. Naderi A, Teschendorff AE, Barbosa-Morais NL, Pinder SE, Green AR, Powe DG, Robertson JF, Aparicio S, Ellis IO, Brenton JD, Caldas C. A gene-expression signature to predict survival in breast cancer across independent data sets. *Oncogene* 2007;26:1507-16.
217. Oh DS, Troester MA, Usary J, Hu Z, He X, Fan C, Wu J, Carey LA, Perou CM. Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol* 2006;24:1656-64.
218. Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, Lindahl T, Pawitan Y, Hall P, Nordgren H, Wong JE, Liu ET, et al. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res* 2006;66:10292-301.
219. Gyorffy B, Dietel M, Fekete T, Lage H. A snapshot of microarray-generated gene expression signatures associated with ovarian carcinoma. *Int J Gynecol Cancer* 2008;18:1215-33.
220. Fedorowicz G, Guerrero S, Wu TD, Modrusan Z. Microarray analysis of RNA extracted from formalin-fixed, paraffin-embedded and matched fresh-frozen ovarian adenocarcinomas. *BMC Med Genomics* 2009;2:23.
221. Grisar D, Hauspy J, Prasad M, Albert M, Murphy KJ, Covens A, Macgregor PF, Rosen B. Microarray expression identification of differentially expressed genes in serous epithelial ovarian cancer compared with bulk normal ovarian tissue and ovarian surface scrapings. *Oncol Rep* 2007;18:1347-56.
222. Heinzelmann-Schwarz VA, Gardiner-Garden M, Henshall SM, Scurry JP, Scolyer RA, Smith AN, Bali A, Vanden Bergh P, Baron-Hay S, Scott C, Fink D, Hacker NF, et al. A distinct molecular profile associated with mucinous epithelial ovarian cancer. *Br J Cancer* 2006;94:904-13.
223. Quinn JE, Carser JE, James CR, Kennedy RD, Harkin DP. BRCA1 and implications for response to chemotherapy in ovarian cancer. *Gynecol Oncol* 2009;113:134-42.
224. Warrenfeltz S, Pavlik S, Datta S, Kraemer ET, Benigno B, McDonald JF. Gene expression profiling of epithelial ovarian tumours correlated with malignant potential. *Mol Cancer* 2004;3:27.
225. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, Szallasi Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* 2010;123:725-31.
226. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC. Jetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics* 2011;12:474.
227. Lacroix M, Querton G, Hennebert P, Larsimont D, Leclercq G. Estrogen receptor analysis in primary breast tumors by ligand-binding assay, immunocytochemical assay, and northern blot: a comparison. *Breast Cancer Res Treat* 2001;67:263-71.
228. Pusztai L, Ayers M, Stec J, Clark E, Hess K, Stivers D, Damokosh A, Sneige N, Buchholz TA, Esteva FJ, Arun B, Cristofanilli M, et al. Gene expression profiles obtained from fine-needle aspirations of breast cancer reliably identify routine prognostic markers and reveal large-scale molecular differences between estrogen-negative and estrogen-positive tumors. *Clin Cancer Res* 2003;9:2406-15.
229. Gyorffy B, Surowiak P, Kiesslich O, Denkert C, Schafer R, Dietel M, Lage H. Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. *Int J Cancer* 2006;118:1699-712.
230. Staunton JE, Slonim DK, Collier HA, Tamayo P, Angelo MJ, Park J, Scherf U, Lee JK, Reinhold WO, Weinstein JN, Mesirov JP, Lander ES, et al. Chemosensitivity prediction by transcriptional profiling. *Proc Natl Acad Sci U S A* 2001;98:10787-92.

231. Scherf U, Ross DT, Waltham M, Smith LH, Lee JK, Tanabe L, Kohn KW, Reinhold WC, Myers TG, Andrews DT, Scudiero DA, Eisen MB, et al. A gene expression database for the molecular pharmacology of cancer. *Nat Genet* 2000;24:236-44.

232. Dietel M, Bals U, Schaefer B, Herzig I, Arps H, Zabel M. In vitro prediction of cytostatic drug resistance in primary cell cultures of solid malignant tumours. *Eur J Cancer* 1993;29A:416-20.

233. Gyorffy B, Serra V, Materna V, Schafer R, Dietel M, Schadendorf D, Lage H. Analysis of gene expression profiles in melanoma cells with acquired resistance against antineoplastic drugs. *Melanoma Res* 2006;16:147-55.

234. Timar J, Gyorffy B, Raso E. Gene signature of the metastatic potential of cutaneous melanoma: too much for too little? *Clin Exp Metastasis* 2010;27:371-87.

235. Gyorffy B, Serra V, Jurchott K, Abdul-Ghani R, Garber M, Stein U, Petersen I, Lage H, Dietel M, Schafer R. Prediction of doxorubicin sensitivity in breast tumors based on gene expression profiles of drug-resistant cell lines correlates with patient survival. *Oncogene* 2005;24:7542-51.

236. Suganuma K, Kubota T, Saikawa Y, Abe S, Otani Y, Furukawa T, Kumai K, Hasegawa H, Watanabe M, Kitajima M, Nakayama H, Okabe H. Possible chemoresistance-related genes for gastric cancer detected by cDNA microarray. *Cancer Sci* 2003;94:355-9.

237. Kang HC, Kim IJ, Park JH, Shin Y, Ku JL, Jung MS, Yoo BC, Kim HK, Park JG. Identification of genes with differential expression in acquired drug-resistant gastric cancer cells using high-density oligonucleotide microarrays. *Clin Cancer Res* 2004;10:272-84.

238. Ein-Dor L, Kela I, Getz G, Givol D, Domany E. Outcome signature genes in breast cancer: is there a unique set? *Bioinformatics* 2005;21:171-8.

239. Abdul-Ghani R, Serra V, Gyorffy B, Jurchott K, Solf A, Dietel M, Schafer R. The PI3K inhibitor LY294002 blocks drug export from resistant colon carcinoma cells overexpressing MRP1. *Oncogene* 2006;25:1743-52.

240. Murata S, Sasaki K, Kishimoto T, Niwa S, Hayashi H, Takahama Y, Tanaka K. Regulation of CD8+ T cell development by thymus-specific proteasomes. *Science* 2007;316:1349-53.

241. Adams J, Palombella VJ, Sausville EA, Johnson J, Destree A, Lazarus DD, Maas J, Pien CS, Prakash S, Elliott PJ. Proteasome inhibitors: a novel class of potent and effective antitumor agents. *Cancer Res* 1999;59:2615-22.

242. Schewe DM, Aguirre-Ghiso JA. Inhibition of eIF2alpha dephosphorylation maximizes bortezomib efficiency and eliminates quiescent multiple myeloma cells surviving proteasome inhibitor therapy. *Cancer Res* 2009;69:1545-52.

243. Kane RC, Farrell AT, Sridhara R, Pazdur R. United States Food and Drug Administration approval summary: bortezomib for the treatment of progressive multiple myeloma after one prior therapy. *Clin Cancer Res* 2006;12:2955-60.

244. Ciolli S, Leoni F, Casini C, Breschi C, Santini V, Bosi A. The addition of liposomal doxorubicin to bortezomib, thalidomide and dexamethasone significantly improves clinical outcome of advanced multiple myeloma. *Br J Haematol* 2008;141:814-9.

245. Munkacsy G, Abdul-Ghani R, Mihaly Z, Tegze B, Tchernitsa O, Surowiak P, Schafer R, Gyorffy B. PSMB7 is associated with anthracycline resistance and is a prognostic biomarker in breast cancer. *Br J Cancer* 2010;102:361-8.

246. Kamata S, Kishimoto T, Kobayashi S, Miyazaki M. Expression and localization of ATP binding cassette (ABC) family of drug transporters in gastric hepatoid adenocarcinomas. *Histopathology* 2008;52:747-54.

247. Larbcharoensub N, Leopairat J, Sirachainan E, Narkwong L, Bhongmakapat T, Rasmeepaisarn K, Janvilisri T. Association between multidrug resistance-associated protein 1

and poor prognosis in patients with nasopharyngeal carcinoma treated with radiotherapy and concurrent chemotherapy. *Hum Pathol* 2008;39:837-45.

248. Molina R, Barak V, van Dalen A, Duffy MJ, Einarsson R, Gion M, Goike H, Lamerz R, Nap M, Soletormos G, Stieber P. Tumor markers in breast cancer- European Group on Tumor Markers recommendations. *Tumour Biol* 2005;26:281-93.

249. Mano MS, Rosa DD, De Azambuja E, Ismael GF, Durbecq V. The 17q12-q21 amplicon: Her2 and topoisomerase-IIalpha and their importance to the biology of solid tumours. *Cancer Treat Rev* 2007;33:64-77.

250. Nguyen T, Menocal EM, Harborth J, Fruehauf JH. RNAi therapeutics: an update on delivery. *Curr Opin Mol Ther* 2008;10:158-67.

251. Rho JH, Qin S, Wang JY, Roehrl MH. Proteomic expression analysis of surgical human colorectal cancer tissues: up-regulation of PSB7, PRDX1, and SRP9 and hypoxic adaptation in cancer. *J Proteome Res* 2008;7:2959-72.

252. Kwak MK, Huang B, Chang H, Kim JA, Kensler TW. Tissue specific increase of the catalytic subunits of the 26S proteasome by indirect antioxidant dithiolethione in mice: enhanced activity for degradation of abnormal protein. *Life Sci* 2007;80:2411-20.

253. Tegze B, Szallasi Z, Haltrich I, Penzvalto Z, Toth Z, Liko I, Gyorffy B. Parallel evolution under chemotherapy pressure in 29 breast cancer cell lines results in dissimilar mechanisms of resistance. *PLoS One* 2012;7:e30804.

254. Saad ED, Katz A, Buyse M. Overall survival and post-progression survival in advanced breast cancer: a review of recent randomized clinical trials. *J Clin Oncol* 2010;28:1958-62.

255. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta* 2010;1805:105-17.

256. Gerlinger M, Swanton C. How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br J Cancer* 2010;103:1139-43.

257. Penzvalto Z, Tegze B, Szasz AM, Sztupinszki Z, Liko I, Szendroi A, Schafer R, Gyorffy B. Identifying resistance mechanisms against five tyrosine kinase inhibitors targeting the ERBB/RAS pathway in 45 cancer cell lines. *PLoS One* 2013;8:e59503.

258. Jung EJ, Moon HG, Park ST, Cho BI, Lee SM, Jeong CY, Ju YT, Jeong SH, Lee YJ, Choi SK, Ha WS, Lee JS, et al. Decreased annexin A3 expression correlates with tumor progression in papillary thyroid cancer. *Proteomics Clin Appl* 2010;4:528-37.

259. Yan X, Yin J, Yao H, Mao N, Yang Y, Pan L. Increased expression of annexin A3 is a mechanism of platinum resistance in ovarian cancer. *Cancer Res* 2010;70:1616-24.

260. Peraldo-Neia C, Migliardi G, Mello-Grand M, Montemurro F, Segir R, Pignochino Y, Cavalloni G, Torchio B, Mosso L, Chiorino G, Aglietta M. Epidermal Growth Factor Receptor (EGFR) mutation analysis, gene expression profiling and EGFR protein expression in primary prostate cancer. *BMC Cancer* 2011;11:31.

261. Nam KT, Lee HJ, Smith JJ, Lapierre LA, Kamath VP, Chen X, Aronow BJ, Yeatman TJ, Bhartur SG, Calhoun BC, Condie B, Manley NR, et al. Loss of Rab25 promotes the development of intestinal neoplasia in mice and is associated with human colorectal adenocarcinomas. *J Clin Invest* 2010;120:840-9.

262. Cheng JM, Volk L, Janaki DK, Vyakaranam S, Ran S, Rao KA. Tumor suppressor function of Rab25 in triple-negative breast cancer. *Int J Cancer* 2010;126:2799-812.

263. Gyorffy B, Lage H. A Web-based data warehouse on gene expression in human malignant melanoma. *J Invest Dermatol* 2007;127:394-9.

264. Zerial M, McBride H. Rab proteins as membrane organizers. *Nat Rev Mol Cell Biol* 2001;2:107-17.

265. Gyorffy B, Schafer R. Biomarkers downstream of RAS: a search for robust transcriptional targets. *Curr Cancer Drug Targets* 2010;10:858-68.
266. Magnani L, Stoeck A, Zhang X, Lanczky A, Mirabella AC, Wang TL, Gyorffy B, Lupien M. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proc Natl Acad Sci U S A* 2013;110:E1490-9.
267. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET, Bergh J. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* 2005;102:13550-5.
268. Loi S, Haibe-Kains B, Desmedt C, Lallemant F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, Klijn JG, Larsimont D, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* 2007;25:1239-46.
269. Ioannidis JP. Microarrays and molecular research: noise discovery? *Lancet* 2005;365:454-5.
270. Fekete T, Raso E, Pete I, Tegze B, Liko I, Munkacsy G, Sipos N, Rigo J, Jr., Gyorffy B. Meta-analysis of gene expression profiles associated with histological classification and survival in 829 ovarian cancer samples. *Int J Cancer* 2012;131:95-105.
271. Abu-Jawdeh GM, Jacobs TW, Niloff J, Cannistra SA. Estrogen receptor expression is a common feature of ovarian borderline tumors. *Gynecol Oncol* 1996;60:301-7.
272. Choe SE, Boutros M, Michelson AM, Church GM, Halfon MS. Preferred analysis methods for Affymetrix GeneChips revealed by a wholly defined control dataset. *Genome Biol* 2005;6:R16.
273. Barash Y, Dehan E, Krupsky M, Franklin W, Geraci M, Friedman N, Kaminski N. Comparative analysis of algorithms for signal quantitation from oligonucleotide microarrays. *Bioinformatics* 2004;20:839-46.
274. Pepper SD, Saunders EK, Edwards LE, Wilson CL, Miller CJ. The utility of MAS5 expression summary and detection call algorithms. *BMC Bioinformatics* 2007;8:273.
275. Gyorffy B, Lanczky A, Szallasi Z. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocr Relat Cancer* 2012;19:197-208.
276. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;474:609-15.

AZ ÉRTEKEZÉS ALAPJÁUL SZOLGÁLÓ KÖZLEMÉNYEK

Közlemény	IF
1. Pénczvártó Zs, Tegze B, Szász AM, Sztupinszki Zs, Likó I, Szendrői A, Schäfer R, Györffy B. Identifying Resistance Mechanisms against five Tyrosine Kinase Inhibitors Targeting the ERBB/RAS Pathway in 45 Cancer Cell Lines. PLoS One , 2013;8(3):e59503. doi: 10.1371/journal.pone.0059503. Epub 2013 Mar 29.	3.730*
2. Magnani L, Stoeck A, Zhang X, Lánckzy A, Mirabella AC, Wang TL, Györffy B, Lupien M. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. Proc Natl Acad Sci U S A . 2013 Apr 16;110(16):E1490-9. doi: 10.1073/pnas.1219992110. Epub 2013 Apr 1. PubMed PMID: 23576735.	9.737*
3. Györffy B, Lánckzy A, Szállási Z. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data of 1287 patients, Endocrine-Related Cancer . 2012 Apr 10;19(2):197-208.	5.261
4. Györffy B, Benke Z, Lánckzy A, Balázs B, Szállási Z, Timár J, Schäfer R. RecurrenceOnline: an online analysis tool to determine breast cancer recurrence and hormone receptor status using microarray data, Breast Cancer Res Treat , 2012;132:1025–1034.	4.469
5. Fekete T, Rásó E, Pete I, Tegze B, Liko I, Munkácsy G, Sipos N, Rigó J, Györffy B. Meta-analysis of gene expression profiles associated with histological classification and survival in 829 ovarian cancer samples. Int J Cancer , 2012 Jul 1;131(1):95-105. doi: 10.1002/ijc.26364.	6.198
6. Tegze B, Szállási Z, Haltrich I, Pénczvártó Z, Tóth Z, Likó I, Györffy B. Parallel Evolution under Chemotherapy Pressure in 29 Breast Cancer Cell Lines Results in Dissimilar Mechanisms of Resistance. PLoS One . 2012;7(2):e30804. Epub 2012 Feb 2.	3.730
7. Li Q, Birkbak NJ, Györffy B, Szállási Z, Eklund AC. Jetset: selecting the optimal microarray probe set to represent a gene, BMC Bioinformatics , 2011 Dec 15;12:474.	2.751
8. Györffy B, Lánckzy A, Eklund AC, Denkert C, Budczies J, Li Q, Szállási Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1809 patients, Breast Cancer Res Treatment , 2010 Oct;123(3):725-31.	4.859
9. Györffy B, Schäfer R. Biomarkers Downstream of RAS: A Search for Robust Transcriptional Targets. Curr Cancer Drug Targets . 2010 Dec 1;10(8):858-68.	4.771
10. Munkácsy G, Abdul-Ghani R, Mihály Z, Tegze B, Tchernitsa O, Surowiak P, Schäfer R, Györffy B. PSMB7 is associated with anthracycline resistance and is a prognostic biomarker in breast cancer, Br J Cancer . 2010 Jan 19;102(2):361-8.	4.831
11. Timár J, Györffy B, Rásó E. Gene signature of the metastatic potential of cutaneous melanoma: too much for too little? Clin Exp Metastasis . 2010 Aug;27(6):371-87.	4.113
12. Györffy B, Schafer R. Meta-analysis of gene expression profiles related to	4.696

relapse-free survival in 1079 breast cancer patients, Breast Cancer Res Treatment , 2009 Dec;118(3):433-41.	
13. Denkert C, Budczies J, Darb-Esfahani S, Györffy B, Sehouli J, Könsgen D, Zeillinger R, Weichert W, Noske A, Buckendahl AC, Müller BM, Dietel M, Lage H. A prognostic gene expression index in ovarian cancer-validation across different independent data sets. J Pathol . 2009 Jun;218(2):273-80.	6.466
14. Györffy B, Molnár B, Lage H, Szállási Z, Eklund A. Evaluation of Microarray Preprocessing Algorithms Based on Concordance with RT-PCR in Clinical Samples, PLoS One , 2009 May 21;4(5):e5645.	4.351
15. Györffy B, Dietel M, Fekete T, Lage H. A snap shot of microarray-generated gene expression signatures associated with ovarian carcinoma, Int J Gynecol Cancer . 2008 Nov-Dec;18(6):1215-33.	1.932
16. Györffy B, Galamb O, Sipos F, Spisák S, Németh AM, Miheller P, Tulassay Z, Dinya E, Molnár B. Inflammation, adenoma and cancer: objective classification of colon biopsy specimens with gene expression signature Dis Markers . 2008;25(1):1-16.	2.303
17. Györffy B, Lage H. A Web-Based Data Warehouse on Gene Expression in Human Malignant Melanoma. J Invest Dermatol . 2007 Feb;127(2):394-9.	4.829
18. Györffy B, Serra V, Schäfer R, Dietel M, Schadendorf D, Lage H. Analysis of gene expression profiles in melanoma cells with acquired resistance against antineoplastic drugs. Melanoma Res . 2006 Apr;16(2):147-155.	1.704
19. Györffy B, Surowiak P, Kiesslich O, Denkert C, Schäfer R, Dietel M, Lage H. Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. Int J Cancer . 2006 Apr 1;118(7):1699-712.	4.693
20. Abdul-Ghani R, Serra V, Györffy B, Jurchott K, Solf A, Dietel M, Schäfer R. The PI3K inhibitor LY294002 blocks drug export from resistant colon carcinoma cells overexpressing MRP1. Oncogene . 2006 Mar 16;25(12):1743-52.	6.582
21. Györffy B, Serra V, Jurchott K, Abdul-Ghani R, Garber M, Stein U, Petersen I, Lage H, Dietel M, Schäfer R: Prediction of Doxorubicin Sensitivity in Breast Tumors Based on Gene Expression Profiles of Drug Resistant Cell Lines Correlates with Patient Survival, Oncogene . 2005;24(51):7542-51.	6.872
Az értekezés alapjául szolgáló közlemények összesen	98,878

AZ ÉRTEKEZÉSTŐL FÜGGETLEN IMPAKT FAKTOROS SAJÁT KÖZELEMÉNYEK A PHD FOKOZAT MEGSZERZÉSE UTÁN

Közlemény	IF
22. Mihály Z, Kormos M, Lánckzy A, Dank M, Budczies J, Szász AM, Györffy B. A meta-analysis of gene expression based biomarkers predicting outcome after tamoxifen treatment in breast cancer. Breast Cancer Res Treat , 2013. <i>In press</i> .	4.469*
23. Malek A, Györffy B, Catapano CV, Schäfer R. Selection of optimal combinations of target genes for therapeutic multi-gene silencing based on	2.945*

miRNA co-regulation. Cancer Gene Therapy , 2013 May;20(5):326-9. doi: 10.1038/cgt.2013.20.	
24. Maciejczyk A, Szelachowska J, Czapiga B, Matkowski R, Halon A, <u>Györfly B</u> , Surowiak P. Elevated BUBR1 Expression Is Associated with Poor Survival in Early Breast Cancer Patients--16 Years Follow-up Analysis. J Histochem Cytochem . 2013 May;61(5):330-9. doi: 10.1369/0022155413480148. Epub 2013 Feb 7.	2.255*
25. Szász AM, Eklund AC, Li Q, Sztupinszki Z, Tóké AM, Rowan A, Székely B, Kiss A, Szendrői M, <u>Györfly B</u> , Szállási Z, Swanton C, Kulka J. The CIN4 chromosomal instability qPCR classifier defines tumour aneuploidy and stratifies outcome in grade 2 breast cancer. PLoS One , 2013;8(2):e56707. doi: 10.1371/journal.pone.0056707. Epub 2013 Feb 26.	3.730*
26. Nagy GR, <u>Györfly B</u> , Nagy B, Rigó J Jr. Lower risk for Down syndrome associated with longer oral contraceptive use: a case-control study of women of advanced maternal age presenting for prenatal diagnosis. Contraception . 2013 Apr;87(4):455-8. doi: 10.1016/j.contraception.2012.08.040.	3.090*
27. Maciejczyk A, Lacko A, Ekiert M, Jagoda E, Wysocka T, Matkowski R, Halon A, <u>Györfly B</u> , Lage H, Surowiak P. Elevated nuclear S100P expression is associated with poor survival in early breast cancer patients. Histology and Histopathology , 2013 Apr;28(4):513-24. Epub 2013 Jan 31.	2.281*
28. Mihály Z, Sztupinszki Z, Surowiak P, <u>Györfly B</u> . A comprehensive overview of targeted therapy in metastatic renal cell carcinoma. Curr Cancer Drug Targets . 2012 Sep 1;12(7):857-72.	4.000
29. Porter DC, Farmaki E, Altília S, Schools GP, West DK, Chen M, Chang DB, Puzyrev AT, Lim C, Rokow-Kittell R, Friedhoff RT, Papavassiliou AG, Kalurupalle S, Hurteau G, Shi J, Baran PS, <u>Györfly B</u> , Wentland MP, Broude EV, Kiaris H, Roninson IB. CDK8 mediates chemotherapy-induced tumor-promoting paracrine activities. Proc Natl Acad Sci U S A . 2012 Aug 21;109(34):13799-804.	9.737
30. Budczies J, Klauschen F, Sinn BV, <u>Györfly B</u> , Schmitt WD, Darb-Esfahani S, Denkert C. Cutoff finder: a comprehensive and straightforward web application enabling rapid biomarker cutoff optimization. PLoS One . 2012;7(12):e51862. doi: 10.1371/journal.pone.0051862.	3.730
31. Budczies J, Denkert C, Müller BM, Brockmöller SF, Klauschen F, <u>Györfly B</u> , Dietel M, Richter-Ehrenstein C, Marten U, Salek RM, Griffin JL, Hilvo M, Orešič M, Wohlgemuth G, Fiehn O. Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue - a GC-TOFMS based metabolomics study. BMC Genomics . 2012 Jul 23;13:334.	4.397
32. Halon A, Nowak-Markwitz E, Donizy P, Matkowski R, Maciejczyk A, Gansukh T, <u>Györfly B</u> , Spaczynski M, Zabel M, Lage H, Surowiak P. Enhanced Immunoreactivity of TIMP-2 in the Stromal Compartment of Tumor as a Marker of Favorable Prognosis in Ovarian Cancer Patients. J Histochem Cytochem . 2012 Jul;60(7):491-501.	2.255
33. Maciejczyk A, Jagoda E, Wysocka T, Matkowski R, <u>Györfly B</u> , Lage H, Surowiak P. ABCC2 (MRP2, cMOAT) Localized in the Nuclear Envelope of Breast Carcinoma Cells Correlates with Poor Clinical Outcome. Pathol Oncol Res . 2012 Apr;18(2):331-42.	1.555
34. Szász AM, Németh Z, <u>Györfly B</u> , Micsinai M, KrenácsT, Baranyai Z,	3.325

Harsányi L, Kiss A, Schaff Z, Tőkés AM, Kulka J. Identification of a claudin-4/E-cadherin score (CURIO) to predict prognosis in breast cancer. Cancer Sci. 2011 Dec;102(12):2248-2254.	
35. Halon A, Nowak-Markwitz E, Maciejczyk A, Pudelko M, Gansukh T, <u>Györfly B</u> , Donizy P, Murawa D, Matkowski R, Spaczynski M, Lage H, Surowiak P. Loss of estrogen receptor beta expression correlates with shorter overall survival and lack of clinical response to chemotherapy in ovarian cancer patients. Anticancer Res. 2011 Feb;31(2):711-8.	1.725
36. Tchernitsa O, Kasajima A, Schäfer R, Kuban RJ, Ungethüm U, <u>Györfly B</u> , Neumann U, Simon E, Weichert W, Ebert MP, Röcken C. Systematic evaluation of the miRNA-ome and its downstream effects on mRNA expression identifies gastric cancer progression. J Pathol. 2010 Nov;222(3):310-9.	7.274
37. Darb-Esfahani S, Sinn BV, Weichert W, Budczies J, Lehmann A, Noske A, Buckendahl AC, Müller BM, Sehoul J, Koensgen D, <u>Györfly B</u> , Dietel M, Denkert C. Expression of classical NF-kappaB pathway effectors in human ovarian carcinoma. Histopathology. 2010 May;56(6):727-39.	3.569
38. Gosepath EM, Eckstein N, Hamacher A, Servan K, Jonquieres G, Royer HD, Lage H, <u>Györfly B</u> , Kassack MU. Acquired cisplatin resistance in the head-neck cancer cell line Cal27 is associated with decreased DKK1 expression and can partially be reversed by overexpression of DKK1. Int J Cancer. 2008 Nov 1;123(9):2013-9.	4.734
39. Galamb O, <u>Györfly B</u> , Sipos F, Dinya E, Krenács T, Berczi L, Szöke D, Spisák S, Solymosi N, Németh AM, Juhász M, Molnár B, Tulassay Z. Helicobacter pylori and antrum erosion-specific gene expression patterns: the discriminative role of CXCL13 and VCAM1 transcripts. Helicobacter. 2008 Apr;13(2):112-26.	2.470
40. Szöke D, Györfly A, Surowiak P, Tulassay Z, Dietel M, <u>Györfly B</u> . Identification of consensus genes and key regulatory elements in 5-fluorouracil resistance in gastric and colon cancer. Onkologie. 2007 Sep;30(8-9):421-6.	1.227
41. Györfly A, Surowiak P, Tulassay Z, <u>Györfly B</u> . Highly expressed genes are associated with inverse antisense transcription in mouse. J Genet. 2007 Aug;86(2):103-9.	0.567
42. Schäfer R, Tchernitsa OI, <u>Györfly B</u> , Serra V, Abdul-Ghani R, Lund P, Sers C. Functional transcriptomics: an experimental basis for understanding the systems biology for cancer cells. Adv Enzyme Regul. 2007;47:41-62.	1.587
43. Györfly A, Tulassay Z, Surowiak P, <u>Györfly B</u> . Computational analysis reveals 43% antisense transcription in 1182 transcripts in mouse muscle, DNA Sequence 2006 Dec;17(6):422-30.	0.569
44. Nagy GR, <u>Györfly B</u> , Galamb O, Molnár B, Nagy B, Papp Z. Use of routinely collected amniotic fluid for whole-genome expression analysis of polygenic disorders. Clin Chem. 2006 Nov;52(11):2013-20.	5.454
45. Surowiak P, Materna V, <u>Györfly B</u> , Matkowski R, Wojnar A, Maciejczyk A, Paluchowski P, Dzięgiel P, Pudelko M, Kornafel J, Dietel M, Kristiansen G, Zabel M, Lage H. Multivariate analysis of estrogen receptor alpha, pS2, metallothionein and CD24 expression in invasive breast cancers. Brit J Cancer. 2006 Aug 7;95(3):339-46.	4.459

46. Györffy A, Vászrhelyi B, Szőke D, Dietel M, Tulassay T, <u>Györffy B</u> . Comparative promoter analysis of doxorubicin resistance associated genes suggests e47 as a key regulatory element. Anticancer Res , 2006 Jul-Aug;26(4B):2971-6.	1.479
47. Surowiak P, Suchocki S, <u>Györffy B</u> , Gansukh T, Wojnar A, Maciejczyk A, Pudelko M, Zabel M. Stromal myofibroblasts in breast cancer: relations between their occurrence, tumor grade and expression of some tumor markers. Folia Histochem Cytobiol , 2006;44(2):111-6.	0.897
48. Vannay A, Vászrhelyi B, Környei M, Treszl A, Kozma G, <u>Györffy B</u> , Tulassay T, Sulyok E. Single-nucleotide polymorphisms of VEGF gene are associated with risk of congenital valvuloseptal heart defects. Am Heart J . 2006 Apr;151(4):878-81.	3.514
49. Vászrhelyi B, Cseh A, Kocsis I, Treszl A, <u>Györffy B</u> , Rigó J Jr. Three mechanisms in the pathogenesis of pre-eclampsia suggested by over-represented transcription factor-binding sites detected with comparative promoter analysis. Mol Hum Reprod . 2006 Jan;12(1):31-4.	2.760
50. Surowiak P, Matkowski R, Materna V, <u>Györffy B</u> , Wojnar A, Pudelko M, Dziegiel P, Kornafel J, Zabel M. Elevated metallothionein (MT) expression in invasive ductal breast cancers predicts tamoxifen resistance. Histol Histopathol . 2005 Oct;20(4):1037-44.	2.023
Az értekezéstől független közlemények összesen	96,077

*: pontos impakt faktor 1 év múlva érhető el

AZ ÉRTEKEZÉSTŐL FÜGGETLEN IMPAKT FAKTOROS SAJÁT KÖZELEMÉNYEK A PHD FOKOZAT MEGSZERZÉSÉIG

Közlemény	IF
51. Kocsis I, Vászrhelyi B, Györffy A, <u>Györffy B</u> : Reanalysis of genotype distributions published in Neurology between 1999 and 2002, Neurology , 2004 Jul 27;63(2):357-8.	5.973
52. Kocsis I, <u>Györffy B</u> , Vászrhelyi B: Examination of Hardy-Weinberg equilibrium in papers of Kidney International: an underused tool, Kidney Int , 2004 May;65(5):1956-1958.	4.790
53. <u>Györffy B</u> , Vászrhelyi B, Krikovszky D, Madácsy L, Tordai A, Tulassay T, Szabó A: Gender-specific association of vitamin D receptor polymorphism combinations with type 1 diabetes mellitus. Eur J Endocrinol 2002 Dec;147(6):803-8.	2.560
Közlemények a PhD fokozat megszerzéséig	13,323

MTMT-ben ellenőrzött összesített impakt faktor = 208,3

KÖSZÖNETNYILVÁNÍTÁS

A leírt kutatások során sokmindenkitől kaptam támogatást. Mivel ezeket nem lehet rangsorolni, ezért időrendben szeretném áttekinteni azokat a legfontosabb állomásokat, amelyek nélkül ez a munka nem születhetett volna meg.

Negyedéves hallgató koromban kezdetem TDK munkát végezni a Semmelweis Egyetem I. Gyermekklinikáján **Vásárhelyi Barna** vezetésével. Neki köszönöm, hogy a tudományos kutatás rendszerével megismertetett, megtanított tudományos közlemények és előadások készítésére és elindított kutatói pályámon. A TDK munka során együtt dolgoztam több hallgatótársammal, **Kocsis Istvánnal, Vannay Ádámmal, Tóth Judittal és Fekete Andreával**, akiknek köszönöm az akkori és az azóta eltelt idő alatt végzett kutatások során nyújtott segítségét.

Az egyetem elvégzése után PhD hallgató lettem **Szabó András** vezetése alatt. Neki köszönöm, hogy az experimentális kutatás módszertanával megismertetett, hogy a külföldi együttműködések fontosságára és a kutatás rendszerezettségének szükségességére felhívta figyelmemet. Köszönöm PhD hallgató társaimnak, elsősorban **Treszl Andrásnak** és **Krikovszky Dórának** a PhD során végzett kutatásokban nyújtott együttműködést.

A PhD képzés után a Berliini Charité Patológiai Intézetében kezdtem dolgozni ösztöndíjasként. Köszönöm mentoromnak, **Reinhold Schäfer**-nek, hogy a türelem és az empátia fontosságára felhívta figyelmemet és megismertette velem a német kutatási rendszer felépítését. Köszönöm ottani posztdoktor kollegáimnak, **Violetta Serrá**-nak, **Pawel Surowiak**-nak és **Jan Budczies**-nek az azóta is tartó együttműködést.

Visszatérve Magyarországra újra a Semmelweis Egyetemen kezdtem dolgozni, és azóta is **Tulassay Tivadar** kutatócsoportjában vagyok. Neki köszönöm, hogy lehetővé tette a kutatásokat, hogy önzetlen támaszt nyújtott kritikus pillanatokban, valamint a megoldhatatlannak tűnő kérdésekhez adott tanácsait.

A Semmelweis Egyetem-i kollegák közül köszönöm az elmúlt évek alatt nyújtott szakmai együttműködést **Szász Marcellnek, Tímár Józsefnek, Rásó Erzsébetnek, Dinya Eleknek, Haltrich Irénnek, Szőke Dominikának, Tulassay Zsoltnak, Kulka Janinának, Nagy Gyula Richárdnak, Rigó Jánosnak, Molnár Bélának és Galamb Orsolyának.**

Köszönöm az együttműködés során nyújtott munkát a Semmelweis Egyetemen kívüli együttműködő hazai kutatási partnereimnek, **Pongor Sándornak, Szállási Zoltánnak, Pete Imrének, és Likó Istvánnak.**

Külön köszönöm PhD hallgatóimnak, **Munkácsy Gyöngyinek, Győrffy Andrásnak, Tegze Bálintnak, Pénzváltó Zsófiának és Mihály Zsuzsannának** a közös munkát. Meg kell említeni azokat a graduális hallgatókat is, akik egyetemi diplomamunkájukat készítették témavezetésemmel: **Lánczky Andrást, Benke Zsombort, Kokas Mártont, Sztupinszki Zsófiát, Weltz Boglárkát, Tilesch Esztert, Vikor Attilát és Balázs Bálintot.**

Köszönöm a nemzetközi együttműködésben végzett munkát **Hermann Lage-nak, Luca Magnani-nak, Igor Roninson-nak, Carsten Denkert-nek, Aron Eklund-nak, Oleg Tchernitsa-nak, Manfred Dietel-nek és Rula Abdul Ghani-nak.**

Végezetül köszönöm az anyagi támogatását azoknak a szervezeteknek, amelyek pályázatokon keresztül lehetővé tették a kutatások elvégzését: **OTKA, Európai Unió, Kutatási és Technológiai Innovációs Alap, ETT, Magyar Posztdoktori Ösztöndíj, Semmelweis Egyetem, Breast Cancer Research Foundation, Fuss a rákkutatásért Alapítvány és a Nemzeti Fejlesztési Ügynökség.**