

Bírálat

Győrffy Balázs: „Transzkriptom szintű adatok alkalmazása a rosszindulatú daganatos betegek várható terápiás válaszána és túlélésének előrejelzésére” című akadémiai doktori értekezéséről.

Győrffy Balázs akadémiai doktori értekezésében az elmúlt 10 év kutatásainak eredményét foglalta részben össze. Mivel még 2004-ben én bíráltam PhD értekezését is, jól nyomon tudtam követni, hogy ez alatt az idő alatt milyen nagyot fejlődött és milyen óriási munkát végzett Balázs. Jellemző erre, hogy az értekezésben csak az azóta írt cikkeinek kevesebb, mint fele szerepel és ezekkel is bőven teljesíti a doktori követelményeket.

Balázs kutatási témáinak és értekezésének fontosságát és jelentőségét mutatja a számos magas kitüntetés, amit az elmúlt években az elért eredményekért kapott.

Az értekezés fő témája annak vizsgálata, hogy a gének expressziója használható-e biomarkerként rosszindulatú daganatos betegek terápiás válaszána előrejelzésére? A dolgozat, bár egy fő téma köré épül, távolról sem egysíkú, hiszen tartalmaz in vitro vizsgálatokat sejtvonalakon, in silico adatok elemzését, saját mintagyűjtést és annak vizsgálatát, rezisztencia mechanizmus vizsgálatát siRNS-sel és széles körben használható bioinformatikai eszközök fejlesztését.

Az értekezés elején található a gondosan összeállított rövidítések jegyzéke, melyet az irodalmi háttérrel is tartalmazó bevezető rész követ 24 oldalon. Az egész dolgozatra jellemző a világos fogalmazás, jól megszerkesztett, logikus felépítés, a kerek és értelmes mondatok. Helyesírási hibák, hibásan megfogalmazott mondatokat csak nagyon keveset találunk, ami nagyban hozzájárul a jó olvashatósághoz. Itt említenék meg egy, az egész dolgozatban gyakran előforduló hibát. A humán gének nevét hivatalosan csupa dőlt nagybetűvel kell írni, ez a szabály itt nagyon sokszor sérül. Sőt nagyon sokszor a génnek nem a hivatalos neve szerepel, hanem egy korábbi, igaz általában széles körben használt név. Pl. a p21-ként említett gén hivatalos neve *CDKN1A*, a survivin *BIRC5*, a p27-é *CDKN1B*, stb. Az rendben van, hogy a jobb érthetőség kedvéért a gén ismertebb, korábbi nevét is használjuk, de az egységes, egymással akár bioinformatikai módszerekkel is összehasonlítható tudományos szakirodalom miatt mindenképpen tanácsos a gének standardizált, szabályos nevét is alkalmazni. Ennek fontosságát maga az értekezés is bizonyítja, hiszen pl. a petefészekrákkal kapcsolatos irodalomkutatás során itt is a HUGO nevezéktan szerint azonosítják a géneket.

A bevezető után, viszonylag hosszan, 5 oldalon a kutatási célok kerülnek megfogalmazásra. A hosszúság időnként egy kicsit zavaró, igaz a probléma inkább csak technikai jellegű, hiszen a célok mellett néha az is leírásra kerül, hogy mit és hogyan vizsgáltak a cél elérésének érdekében. Ennek persze az az oka, hogy a célok retrospektív lettek az értekezésben megfogalmazva.

A következő rész 24 oldalon található az alkalmazott módszerek ismertetése, amely megfelelő színvonalon, tömören, mégis érthetően ismerteti ezeket. Számomra különösen meggyőző a leírások interdiszciplináris jellege, és hogy a jelölt milyen egymástól teljesen más szakértelmet igénylő (pl. orvosi, genetikai, informatikai) szakterületen mutat kiemelkedő hozzáértést.

A következő közel 40 oldalon található az értekezés talán legfontosabb része, az eredmények ismertetése. Az értekezés továbbra is magas színvonalú, azonban itt szeretnék megjegyezni egy formai hibát, ami számomra igen megnehezítette a dolgozat megfelelő értékelését. Az értekezés a végén található irodalomjegyzék alapján 21 közlemény alapján készült. Ezekre az értekezésen belül közvetlenül nem található hivatkozás, csak a 276 hivatkozás közé elbújtatva, így sokszor nekem kellett kitalálnom, hogy melyik eredmény melyik cikkből származott. Ráadásul a hivatkozás sokszor csak megbeszélés során kerül megemlítésre. Nekem nem sikerült mind a 21 cikket beazonosítanom. Jó lenne, ha az egyes eredményekhez közvetlenül hozzá lennének rendelve a felsorolt az értekezés alapjául szolgáló közlemények.

Természetesen ez nem befolyásolja azt, hogy a jelöltnek meggyőző irodalmi háttere van és a leírt eredmények bőven elérik az MTA doktora cím megítéléséhez szükséges szintet.

A dolgozatban az eredmények logikusan, egymásra épülve kerültek leírásra. Az első pontban 30 sejtvonalon 12 kemoterápiás szerrel szembeni rezisztenciáját vizsgálták, majd az egyes gyógyszerekre rezisztens és érzékeny sejtvonalak génexpresszióját hasonlították össze microarray-jel. Úgy is lehet tekinteni, hogy ezzel a vizsgálattal alapozták meg a továbbiakat, hiszen a rezisztens és érzékeny sejtek között expressziós különbséget mutató gének a jelölt elmélete szerint in vivo is felhasználhatók predikciós célokra. Ezt az elméletet később a gyakorlatban is igazolták. Az egyes gének előrejelzésének hatékonyságát aztán ugyanezekben a sejtvonalakon tesztelték, illetve 46 kiválasztott gén TaqMan RT-PCR-rel validáltak. Sajnos ez utóbbi validációról csak nagyon homályos adatok szerepelnek az értekezésben. A hozzá tartozó cikkben van egy táblázat, amelyben viszont több mint 20 esetben a különbséget jelző P érték nagyobb, mint 0,05. Ez azt jelenti, hogy nem minden esetben sikerült igazolni a microarray eredményeket? Mekkora az egyezés a kétféle módszer között?

Az értekezésben ezt követően számos érdekes és értékes eredményt találunk, amelynek részletezésétől itt most eltekintek, és csak azokkal foglalkozom, amelyekkel kapcsolatban kérdés merült fel bennem.

Sokszor az értekezésben hiányzó adat, megtalálható a hozzá tartozó cikkben és általában jellemző, hogy a cikkek egy minőségi fokozattal jobbak, mint az értekezésben leírtak, illetve ritkán a kettő között ténybeli különbség is van. Például a 4.3-as pontban leírt eredményeknél a doxorubicin és mitoxantron rezisztenciával kapcsolatba hozott gének validációját 44, irodalomban azonosított betegen végzik, azonban az értekezésben nem szerepel, hogy milyen betegekről van szó, pedig az is fontos adat.

Nem tartom szerencsésnek, hogy a szövegben nem szerepelnek ábra sorszámkra való hivatkozások, legfeljebb csak annyi, hogy lásd a következő ábrát, pedig az ábráknak vannak számaik.

A 63. oldalon található 12. ábrán nekem a „censored” fordításaként nem tetszett a „kizárt”. Szerintem pl. a cenzorált szó jobb lett volna.

Az értekezésre, illetve a jelölt kutatómunkájára általában jellemző, hogy megfigyeléseket ír le, nem nagyon próbál a miértre, a megfigyelt dolgok mögött mechanizmusokra magyarázatot találni. Ez egyébként egyáltalán nem baj, hiszen így is rendkívül hasznos, a gyakorlati életben is hasznosítható eredményekről olvashatunk. Mégis számomra nagyot emelt a dolgozat értékén és érdekességén az a munka, amelyben siRNS felhasználásával igazolták, hogy a *PSMB7* gén driver génként szerepel. Ennek az eredménynek is komoly klinikai következményei lehetnek.

A 65. oldalon a *PSMB7* gén csendesítésének ellenőrzésénél PCR szerepel, mint módszer.

Mivel itt a génexpresszió csökkenésének ellenőrzése zajlik, valószínűleg RT-PCR volt a módszer.

A 15. ábra ismét egy példa arra, hogy az értekezés kicsit kevésbé alaposan lett összeállítva, mint a cikkek. Itt a rajzon a kipróbált siRNS-ek különbözőképpen vannak bekeretezve. Az csak a cikkből derül ki, hogy ennek jelentősége van, a leghatásosabb siRNS-t máshogy van jelölve, mint a kevésbé hatásosak.

A 16. ábra az eredményekhez tartozó cikkben (ahol 3B néven szerepel) lényegesen jobb és informatívabb. Nem értem miért nem ez került be az értekezésbe.

A 4.5. pontban a szövegben az szerepel, hogy 40 sejtvonalból 27 rezisztens leányvonalat fejlesztettek ki. Ezzel szemben a 18. ábrán, illetve a hozzá tartozó cikkben is az szerepel, hogy 2 sejtvonalból 29 leányvonalat fejlesztettek ki. Melyik az igaz?

A 68. oldalon számomra egy kicsit homályos megfogalmazással az olvasható, hogy azt találták, hogy egy-egy sejtvonalban a rezisztencia kialakulásánál mindössze egy-két mechanizmus volt aktív. Ez az addig közltekől számomra nem következik. Miből állapították meg? Ráadásul az eredményekhez kapcsolható cikkben ezt a megállapítást nem találtam.

A 4. táblázatban láthatjuk a paclitaxel és a doxorubicin rezisztenciával kapcsolatba hozott géneket. Itt szerepelnek például a *TUBB2* és *TUBB3* gének, mint a paclitaxel rezisztencia génjei. Bár az értekezésben nem szerepel, a cikkben erről ezt olvashatjuk:

„It is also notable that two genes selected for their previously described association with taxane resistance (*TUBB2C* and *TUBB3*) were associated with doxorubicin resistance in our cell line panel, but with the opposite trend,”

Ezek szerint a cikkben a doxorubicinnel mutat rezisztenciát, míg az értekezésben a paclitaxellel. Mi az igazság? Illetve mi lehet a magyarázata, hogy bár asszociációt mutatott a cikkben doxorubicin rezisztenciával, de ellentétes trenddel?

Még egy ellentmondást találtam a cikk és az értekezés között. Az értekezésben ez szerepel: Több, korábban már leírt különféle mechanizmus aktiválódását (pld *TOP2A*, *TUBB3*, *ABCB1*) tudtuk igazolni a rezisztencia létrejötte során. Ezzel szemben a cikkben:

ABCB1 reached a p value of 0.03 and 0.07 in the doxorubicin- and paclitaxel-treated cell lines, respectively but these were not significant after multiple testing corrections.

A kérdés az, hogy az *ABCB1* igazoltan aktiválódik, vagy egyelőre a statisztikai adatok nem meggyőzőek a szerepét illetően?

A következőkben az első kísérlethez hasonlóan sejtvonalakat vizsgáltak, összesen 45-t és nézték rezisztenciájukat 5 célzott kemoterápiás szerre. Ebben a kísérletben nem saját maguk határozták meg a kiválasztott sejtek génexpresszióját, hanem a caArray project microarray méréseit hasznosították. Itt az eredmények leírásánál, illetve a megbeszéléseknél ez a lépés nincs megemlítve, így a megértéshez vissza kell emlékezni, vagy lapozni a 33 oldallal korábbi leírásra. Viszont, hasonlóan az előzőekhez itt is akár a gyakorlatban is hasznosítható eredményeket kaptak, aminek egy részét 39 veserákostól gyűjtött mintán

immunhisztokémiával is megvizsgáltak és 3 génterméknél a festődés intenzitása összefüggött a jobb túléléssel.

Szintén nagyon hasznosnak és értékesnek tartom a melanoma, emlőrák és petefészekrák kapcsán elvégzett metaanalíziseket, ahol olyan leközölt vizsgálatokat elemeztek, amelyekben a génexpressziós mintázatok és a klinikai kép között kerestek összefüggéseket. Mindig is nagyon fontosnak tartottam, hogy az egyes cikkekben leközölt eredmények ne vesszenek el, hanem ilyen, minőségellenőrzéssel egybekötött együttes elemzéssel az eredeti cikkekben benne levő esetleges rejtett értékek, eredmények is hasznosuljanak. Én hiszek abban, hogy a PubMed-ben található több mint 24 millió cikkben választ lehetne megfelelő elemzéssel találni rengeteg fontos, és jelenleg nem megoldott problémára, kérdésre. A jelölt ebben, szerintem a világszínvonalon is jelentős teljesítményt nyújt, és talán ez az egyik legfontosabb hozzájárulása a tudományos ismeretekhez.

A következő fejezetek ezekben a bioinformatikai projekteknél elért eredményeit ismerteti a jelölt olyan témákkal, mint a microarray adatok előfeldolgozása és más különböző bioinformatikai fejlesztések. Nagyon értékesnek tartom a világhálón keresztül mindenki számára elérhető különböző web-es felületek kifejlesztését, melyet kutatók, sőt klinikusok is használhatnak, és mint láthatjuk, széles körben használnak is.

Az értekezés megbeszélés részében az eredmények magas színvonalú interpretálása található 20 oldalon. Ahogy a korábban is említettem, itt történik kicsit elbújtatva sok, saját cikke az első hivatkozás. Jó ötletnek tartom viszont, hogy az adott témakör lényeges elemei, a cikk mondanivalója pontokba van szedve, megkönnyítve ezzel az érthetőséget. A jelölt ebben a fejezetben is meggyőz minket felkészültségéről, a különböző tudományágakban való jártasságáról és arról a képességéről, hogy a saját eredményeit, a szakirodalmat és az egyes tudásbázisokat kritikusan értékelje.

Az értekezés 6. fejezetében Györffy Balázs 3 fő pontban és 12 alpontban foglalja össze saját kutatási eredményeit. A felsorolás reális és meggyőző.

A hivatkozások jegyzékében 276 cikk található megfelelő formátumban. Kivétel a 162. hivatkozás, ahol a szerzők hiányoznak.

Az értékelés közben feltett kérdéseim mellett az értekezés kapcsán a következőket kérdezném:

1. Különböző eredetű sejtvonalak génexpresszióit hasonlította össze az alapján, hogy rezisztensek, vagy érzékenyek egy adott kemoterápiás szerre. Az eredetüktől függően

ezek a sejtvonalak a rezisztenciától függetlenül is nyilván különböztek a génexpresszióban. Nem okoz-e ez torzítást az eredményekben? Tehát pl., ha összehasonlít egy prosztatata és egy vastagbél eredetű sejtvonalat, azok nyilván eredendően más expressziós profillal rendelkeznek.

2. A sejttenyészeteket hosszú ideig különböző kemoterápiás szerek jelenlétében tartották fent. Nem változtak-e (differenciálódtak) meg teljesen ezek a sejtvonalak ilyen hosszú idő alatt?
3. A legújabb egy sejt szinten történő NGS (pl. nuc-seq) eredmények alapján egy tumorban minden sejt genetikailag különböző. Egyrészt mi a véleménye erről, másrészt, ha ez így van, nem befolyásolja-e az eredményt, hogy génexpressziós vizsgálathoz honnan vesszük a mintát? Van-e erre valami szabály?
4. Hogyan látja a jövőjét a génexpresszió alapú diagnosztikának?
5. A genomika fejlődésével milyen ígéretes módszereket prognosztizál, amelyek jelentősen javíthatják a tumorok diagnózisát és a személyreszabott terápiát?

Összességében megállapítom, hogy Györffy Balázs szcientrometriai adatai kimagaslóak, cikkei és a benyújtott értekezés magas színvonalúak. Az értekezéssel párhuzamosan benyújtott tézisfüzet jól foglalja össze az értekezés fő mondanivalóit. A Györffy Balázs PhD fokozatának elérése óta egyrészt óriásit fejlődött, másrészt jelentős mennyiségű eredeti cikkeivel és kiemelkedő minőségű kutatásaival, fejlesztéseivel jelentős mértékben járult hozzá nemcsak a magyar, hanem a nemzetközi tudományos ismeretekhez jelentősen gyarapítva a tudományszakot és hozzájárult a tudomány fejlődéséhez. Ezek **alapján a nyilvános vita kitűzését és a mű elfogadását javaslom.**

Budapest, 2014. szeptember 8.

Dr. Szalai Csaba MTA doktora
SE Genetikai Sejt és Immunbiológiai Intézet
Igazgató helyettes
Heim Pál Kórház