

**SEMMELWEIS EGYETEM**  
**DOKTORI ISKOLA**

**Ph.D. értekezések**

**2478.**

**NAGY ÁDÁM**

**Onkológia**  
című program

Programvezető: Dr. Kopper László, professor emeritus

Témavezető: Dr. Gyórfy Balázs, egyetemi tanár

# **Prognosztikus biomarkerek vizsgálata különböző daganattípusokban**

Doktori értekezés

**Nagy Ádám**

Semmelweis Egyetem

Patológiai tudományok Doktori Iskola



Témavezető:

Dr. Győrffy Balázs D.Sc., egyetemi tanár

Hivatalos bírálók:

Dr. Tantos Ágnes Ph.D., tudományos főmunkatárs

Dr. Hritz István Ph.D., egyetemi docens

Szigorlati bizottság elnöke:

Dr. Alpár Alán D.Sc., egyetemi tanár

Szigorlati bizottság tagjai:

Dr. Vannay Ádám Ph.D., tudományos főmunkatárs

Dr. Szüts Dávid Ph.D., tudományos főmunkatárs

Dr. Than Nándor Gábor Ph.D., tudományos főmunkatárs

Budapest

2020

**TARTALOMJEGYZÉK**

<b>1</b>	<b>RÖVIDÍTÉSEK JEGYZÉKE .....</b>	<b>6</b>
<b>2</b>	<b>BEVEZETÉS.....</b>	<b>8</b>
2.1	A daganatok előfordulása világszerte és Magyarországon.....	8
2.2	A daganatok általános jellemzői, daganat hallmarkok .....	10
2.2.1	A sejtosztódás fenntartása .....	11
2.2.2	A tumorszuppresszorok visszaszorítása .....	11
2.2.3	A sejthalál kikerülése.....	12
2.2.4	Korlátlan osztódási képesség.....	12
2.2.5	Érképzés serkentése .....	12
2.2.6	Invázió és metasztázis indukálása .....	13
2.2.7	A sejtek energiametabolizmusának az átprogramozása .....	14
2.2.8	A genom stabilitás fenntartása.....	15
2.2.9	A gyulladás serkentése .....	15
2.2.10	Az immunrendszer visszaszorítása.....	15
2.3	Egyes daganattípusok jellemzése, molekuláris terápiás célpontok.....	16
2.3.1	A tüdődaganatok.....	16
2.3.1.1	Terápiás célpontok a nem-kissejtes tüdőrákban .....	17
2.3.1.1.1	Az epidermális növekedési faktor receptor (EGFR) mutációk .....	18
2.3.1.1.2	Az anaplasztikus lymphoma kináz (ALK) mutációk .....	18
2.3.1.1.3	A BRAF mutációk.....	19
2.3.1.1.4	A KRAS mutációk .....	19
2.3.2	A májdaganatok.....	20
2.3.2.1	Biomarkerek a hepatocelluláris karcinómában.....	21
2.3.2.1.1	Génexpresszió alapú biomarkerek .....	21
2.3.2.1.2	Génmutáció alapú biomarkerek .....	22
2.3.2.1.3	mikro-RNS alapú biomarkerek .....	23
2.3.2.2	A hepatocelluláris karcinóma célzott terápiás lehetőségei .....	23
2.3.2.3	A vaszkuláris invázió szerepe a daganat kiújulásában .....	25
2.3.3	Az akut mieloid leukémia.....	25

2.3.3.1	Biomarkerek az akut mieloid leukémiában .....	26
2.3.3.2	Az NPM1 mutációk .....	26
2.4	A nukleinsav szekvenálás .....	27
2.4.1	Sanger szekvenálás .....	28
2.4.2	Az újgenerációs szekvenálás .....	29
2.5	A génextpresszió mérésére alkalmas módszerek .....	33
2.5.1	A kvantitatív polimeráz láncreakció .....	33
2.5.2	A microarray .....	34
2.5.3	Az RNS szekvenálás .....	35
<b>3</b>	<b>CÉLKITŰZÉSEK.....</b>	<b>38</b>
<b>4</b>	<b>MÓDSZEREK .....</b>	<b>40</b>
4.1	Mutációs és génextpressziós adatok feldolgozása .....	40
4.1.1	mRNS expressziós adatok .....	40
4.1.1.1	RNS szekvenálás alapú génextpressziós adatok feldolgozása .....	40
4.1.1.2	Microarray alapú génextpressziós adatok feldolgozása .....	41
4.1.2	Mikro-RNS expressziós adatok .....	42
4.1.2.1	Microarray alapú mikro-RNS expressziós adatok feldolgozása .....	42
4.1.2.2	RNS szekvenálás alapú mikro-RNS expressziós adatok .....	43
4.1.3	Szomatikus mutációk azonosítása .....	43
4.1.4	Kópiaszám eltérések azonosítása .....	44
4.1.5	TCGA programból elérhető mutációs adatok .....	44
4.2	Statisztikai számítások .....	44
4.2.1	A szomatikus mutációk és a génkifejeződés kapcsolata .....	44
4.2.1.1	A KRAS mutációk és kópiaszám eltérések hatása a génkifejeződésre nem-kissejtes tüdőrákban .....	45
4.2.1.2	Az NPM1 mutációk hatása a génkifejeződésre akut mieloid leukémiában .....	46
4.2.2	Mikro-RNS kifejeződés normál és tumoros májszövetben .....	46
4.2.3	Túlélés elemzés .....	46
4.2.3.1	Univariáns túlélés elemzés .....	47
4.2.3.2	Multivariáns túlélés elemzés .....	47
4.2.4	Daganat jellemző specifikus génextpressziós mintázat .....	48

4.3	Klinikai minták feldolgozása .....	48
4.3.1	Az NPM1 gén Sanger szekvenálása .....	48
4.3.2	RNS izolálás .....	49
4.3.3	Kvantitatív polimeráz láncreakció .....	49
<b>5</b>	<b>EREDMÉNYEK .....</b>	<b>51</b>
5.1	KRAS mutáció függő génexpressziós változások vizsgálata nem-kissejtes tüdőrákban .....	51
5.1.1	A nem-kissejtes tüdőrák adatbázis felépítése .....	51
5.1.2	A KRAS mutációk .....	52
5.1.3	A KRAS mutációkkal összefüggő transzkriptomikus változások .....	56
5.1.4	A KRAS amplifikációval és delécióval összefüggő génexpressziós mintázat hatása a túlélésre .....	59
5.2	Az NPM1 mutációk transzkriptomikus lenyomata és validálása akut mieloid leukémiában .....	62
5.2.1	Akut mieloid leukémia adatbázis .....	62
5.2.2	Az NPM1 mutációkkal összefüggő gének .....	68
5.2.3	Az NPM1 mutációk transzkriptomikus lenyomatának igazolása klinikai mintákon .....	77
5.2.4	A HOX gének és a kofaktoraik összefüggése .....	80
5.3	Prognosztikus biomarkerek azonosítása máj hepatocelluláris karcinómában .	81
5.3.1	Adatbázis felépítése .....	81
5.3.1.1	mikro-RNS expressziós adatbázis .....	81
5.3.1.2	mRNS expressziós adatok .....	84
5.3.2	Prognosztikus mikro-RNS-ek vizsgálata máj hepatocelluláris karcinómában .....	84
5.3.2.1	Korábban publikált prognosztikus mikro-RNS-ek .....	84
5.3.2.2	Prognosztikus mikro-RNS-ek validálása .....	85
5.3.2.3	Eltérő kifejeződést mutató mikro-RNS-ek a normál és a tumoros májszövet között .....	90
5.3.2.4	Prognosztikus mikro-RNS-ek korrelációja .....	96
5.3.3	Génexpresszió alapú biomarkerek vizsgálata máj hepatocelluláris karcinómában .....	96

5.3.3.1	Prognosztikus biomarkerek irodalmi keresése a máj hepatocelluláris karcinómában .....	96
5.3.3.2	Prognosztikus biomarkerek vizsgálata az ázsiai és a kaukázusi betegpopulációkban.....	97
5.3.3.3	Multivariáns túlélés elemzés.....	102
5.3.3.4	A vaszkuláris invázióval összefüggő gének .....	102
5.4	Daganat asszociált gének prognosztikus hatásának vizsgálata különböző tumortípusokban .....	104
5.4.1	RNS szekvenálási adatbázis felépítése .....	104
5.4.2	Daganat asszociált gének hatása a teljes túlélésre.....	106
5.4.3	Transzkriptomikus mintázatok prognosztikus hatása.....	106
5.4.4	Prognosztikus gének eloszlása különböző daganattípusokban.....	107
<b>6</b>	<b>MEGBESZÉLÉS .....</b>	<b>109</b>
6.1	KRAS mutáció függő génexpressziós változások vizsgálata nem-kissejtes tüdőrákban .....	109
6.2	Az NPM1 mutációk transzkriptomikus lenyomata és validálása akut mieloid leukémiában .....	111
6.3	Prognosztikus biomarkerek azonosítása máj hepatocelluláris karcinómában	113
6.4	Daganat asszociált gének prognosztikus hatásának vizsgálata különböző tumortípusokban .....	114
<b>7</b>	<b>KÖVETKEZTETÉSEK.....</b>	<b>118</b>
<b>8</b>	<b>ÖSSZEFOGLALÁS.....</b>	<b>120</b>
<b>9</b>	<b>SUMMARY .....</b>	<b>121</b>
<b>10</b>	<b>IRODALOMJEGYZÉK .....</b>	<b>122</b>
<b>11</b>	<b>SAJÁT PUBLIKÁCIÓK JEGYZÉKE .....</b>	<b>149</b>
11.1	A disszertációhoz kapcsolódó publikációk jegyzéke .....	149
11.2	A disszertációtól független publikációk jegyzéke.....	149
<b>12</b>	<b>KÖSZÖNETNYILVÁNÍTÁS .....</b>	<b>150</b>

## 1 RÖVIDÍTÉSEK JEGYZÉKE

5-FU - 5-fluorouracil

AAV2 - Adeno-asszociált vírus 2

ALK - Anaplasztikus lymphoma kináz

BECN1 - Beclin-1

cDNS - Komplementer DNS

CEA - Carcinoembrionic antigén

CI - Konfidencia intervallumok

CNV - Kópiaszám eltérései

DAVID - Database for Annotation, Visualization and Integrated Discovery

EGFR - Epidermális növekedési faktor receptor

ER – Ösztrogén receptor

FAB - French-American-British

FDR - False discovery rate

FLT3 - FMS-like tyrosine kinase 3

GDC - Genomic Data Commons

GTP - Guanozin-trifoszfát

HBV - Hepatitis B vírus

HCV - Hepatitis C vírus

HR - Hazard ratio

IDH1/2 - Izocitrát dehidrogenáz

IHC - Immunhisztokémia

ITD - Internal tandem duplikációk

MAF - Mutation annotated format

MDSC - Mieloid szuppresszor sejtek

NCBI GEO - National Center for Biotechnology Information Gene Expression Omnibus

NPM1 - Nucleophosmin gén

NSCLC – Non-small cell lung cancer

PANTHER - Protein ANalysis THrough Evolutionary Relationships

qPCR - Kvantitatív polimeráz láncreakció

SNP - Egy nukleotidos polimorfizmus

TCGA - The Cancer Genome Atlas

Tregs - Regulátor T sejtek

TSP1 - Thrombospondin-1

VEGFA - Vaszkuláris endotheliális növekedési faktor-A



## 2 BEVEZETÉS

### 2.1 A daganatok előfordulása világszerte és Magyarországon

A rák előfordulása és mortalitása világszerte rohamosan növekszik. Ezt a növekedést számos tényező befolyásolja, amelyek közül fontos kiemelni a népesség elöregedését és a népesség növekedését. A GLOBOCAN 2018 becslései alapján, a világon évente megközelítőleg 18 millió új rosszindulatú daganatos megbetegedést diagnosztizálnak és ebből mintegy 9,5 millió esetben a betegség végkimenetele halállal végződik (1). A daganatos megbetegedések a vezető halálokok közé tartoznak az egész világon, amelyek nemtől és kortól függetlenek. Az új esetek gyakorisága alapján elmondható, hogy a világon minden második férfi, valamint háromból egy nő az élete során rosszindulatú tumorban szenved (1). Továbbá, számos afrikai régióban, a nők körében, a daganattal kapcsolt halálozási esetek száma megközelíti az észak-amerikai és egyes nyugat-európai ország halálozási rátáját. A daganatos betegségek és az ezekkel kapcsolt halálokok döntő többségét az ázsiai populáció adja, Európa az összes új eset 23%-át, míg az amerikai kontinens 21%-át teszi ki. A daganattal összefüggő halálozási gyakoriság tekintetében Európa az összes eset 20%-át, míg Amerika 14,4% adja (1).

Az egyik leggyakoribb daganattípus a tüdőrák, mintegy 12%-át adja az összes diagnosztizált daganatos megbetegedésnek. A tüdőrákot számos további magas incidenciájú daganat követi, úgymint az emlő, a kolorektális daganat és a prosztatatarák. A tüdőrák, a halálozás gyakoriság alapján is a ranglétra tetején áll (18,4%), de a kolorektális daganatok (9,2%), a gyomor (8,2%) és a májdaganat (8,2%) esetén is magas a halálozás (1). A nemek tekintetében, a férfiak körében a tüdőrák a leggyakoribb daganattípus és a halálozási ráta is ezen tumortípus esetén a legnagyobb. Továbbá szintén magas gyakorisággal bír a prosztatata és a kolorektális daganat, valamint a halálokok körében a tüdőrák mellett a máj és a gyomorrák vezet. A nőknél az emlőrákot a kolorektális daganat és a tüdőrák követi a leggyakrabban diagnosztizált daganattípusok körében, a vezető halálokokhoz pedig a tüdőrák, a kolorektális daganat és az emlőrák tartozik (1).

A daganatok incidenciája és a mortalitása területi eloszlásuk alapján elmondható, hogy a férfiak körében a prosztatatarák a legtöbb országban – ide tartozik Észak-, Közép- és Dél-Amerika, Ausztrália, Afrika és Európa nagy része – a leggyakrabban diagnosztizált daganatos megbetegedés (1). A prosztatatarákot a tüdőrák követi a

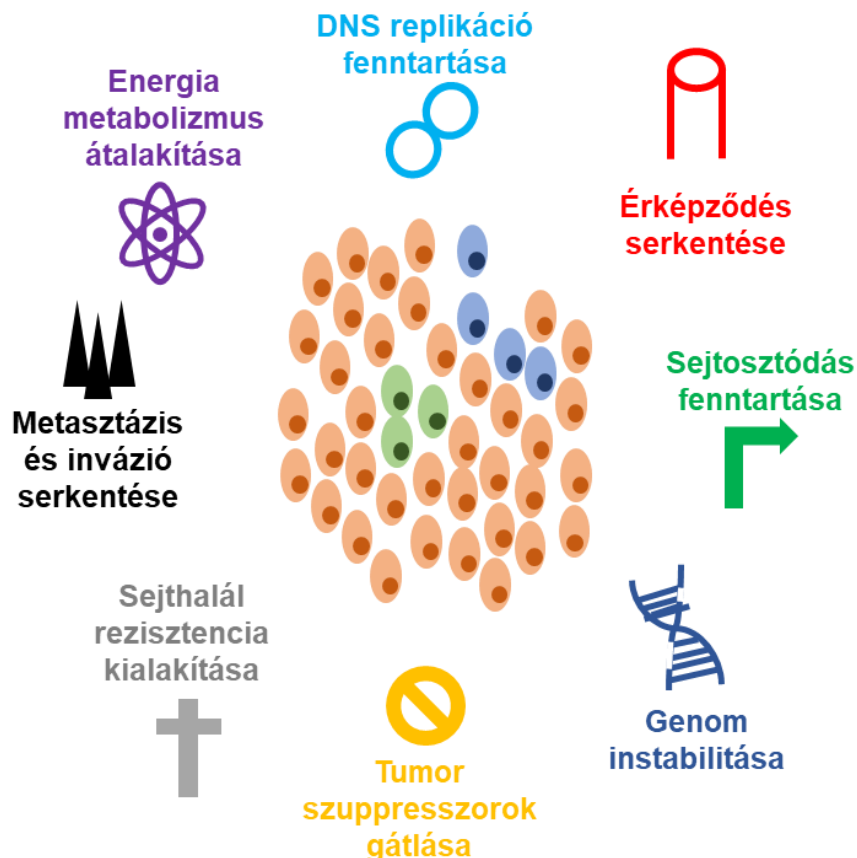
megbetegedések szempontjából, amely esetében legérzékenyebb részek Kelet-Európa, Oroszország, Közel-Kelet, Kína és Észak-Afrika egyes országai. Érdemes kiemelni azokat a daganatokat, amelyek egy adott földrajzi régióban vagy országban mutatnak halmozódást, ilyen például a májdaganat, amely Délkelet-Ázsia egyes országaiban és Mongóliában mutatja a legmagasabb esetszámot. A halálozási ráta szempontjából a legtöbb országban a tüdőrákkal kapcsolt halálozás a leggyakoribb. Ide sorolható Észak- és Dél-Amerika, Európa, Ázsia nagy része és Ausztrália, továbbá a prosztaták mortalitása Közép-Amerikában, Közép-Afrikában a leggyakoribb, valamint a májdaganatnak pedig Észak-Afrika egyes országaiban, Mongóliában és Délkelet-Ázsiában a leggyakoribb vezető daganatos halálok (1). A nők körében az emlőrák a legtöbb országban diagnosztizált leggyakoribb daganatos megbetegedés, amely esetében fontos kiemelni, hogy néhány ország kivételével az incidencia szempontjából ezen tumortípus a vezető. A daganat mortalitási ráta földrajzi eloszlása, azonban az incidenciához képest viszonylag heterogén, legmagasabb Közép- és Dél-Amrikában, Észak-Afrikában, Közel-Keleten, Oroszországban, Indiában és Délkelet-Ázsiában. A mortalitás körében azonban az emlőrák mellett méhnyakrákot is fontos kiemelni, amely Közép- és Dél-Afrika országaiban vezeti a daganattal összefüggő halálozási ranglistát (1).

A daganatstatisztikát hazánkra szűkítve láthatjuk, hogy a tüdőrák, kolorektális daganatok, valamint a szájüregi és a gégedaganatok a leggyakrabban diagnosztizált daganattípusok közé tartoznak (2). A környező országokban ezen daganattípusok szintén magas gyakorisági rátával rendelkeznek (3). Kutatócsoportunk egy korábbi munkája során, a Magyarországon rendelkezésre álló epidemiológiai adatok felhasználásával helyspecifikus előrejelzést készítettek az összes daganatos megbetegedésről (4). Magyarországon a férfiak esetén a teljes élethosszra vonatkoztatott daganat kialakulási valószínűség 56,9%, a daganattal kacsolt halálozási ráta pedig 27,6% (4). Ezek az értékek valamivel alacsonyabbak a nőknél, a daganat kialakulási valószínűsége az életük során 51,9%, a daganattal összefüggő halálozási gyakoriságuk pedig 21,7% (4). Ezen tanulmány során előrejelezték, hogy a teljes népesség 2016 és 2030 között várhatóan 6%-kal fog csökkenni, miközben a 60 éves és az annál idősebbek száma 18%-kal fog növekedni. Ez a daganat kialakulásának 35%-os növekedését és a rosszindulatú daganattal kapcsolatos halálozás 30%-os növekedését eredményezi a 65–85 éves korosztály körében. A nők esetében az tüdőrák incidenciája 2016 és 2030 között évente

előre láthatólag 1,9%-kal, az emlőráké pedig 1,7%-kal fog növekedni. A férfiak körében a prosztatatarák és a vastagbélrák gyakorisága sokkal intenzívebben évi 3,6%-kal és 2,1%-kal fog emelkedni (4).

## 2.2 A daganatok általános jellemzői, daganat hallmarkok

A daganatok kialakulása egy meglehetősen összetett, soklépcsős folyamat, amelyért számos sejten belüli és sejten kívüli folyamat rendellenes működése tehető felelőssé. Hanahan és Weinberg korábbi munkájuk során leírtak nyolc jellemzőt, amelyek összefüggenek a daganatos elváltozásokkal (5). Ide tartozik a folyamatos sejtosztódás stimulálása, a növekedést gátlók megfékezése, sejthalál kikerülése, a folyamatos DNS szintézis fenntartása, érékpeződés, valamint az sejtek inváziós és áttétképző képességének indukálása (**1. ábra**). Ezekhez később, a genomi instabilitás, a gyulladás, az immunmikrokörnyezet és a sejtek energia metabolizmusának átalakulása társult (6).



**1. ábra.** A rosszindulatú daganatok kialakulásában szerepet játszó jellemzők.

### 2.2.1 A sejtosztódás fenntartása

A sejtosztódás fenntartása, az egyik olyan tényező, amely elengedhetetlen a normál sejtek daganatos sejté történő átalakulásához. A sejtosztódási jeleket többnyire növekedési faktorok indukálják, amelyek a sejtek felszínén lévő receptorokhoz kötődnek. A kölcsönhatás során a receptorok a sejt belső részén található tirozin kináz egység aktiválódik és így továbbítódik a jel a sejtmag felé, amelynek a következtében a sejtosztódás előrehaladása és a sejtnövekedés szabályozódik. Azonban más folyamatokat is képes befolyásolni, mint például a sejtek túlélését és energia metabolizmusát (5). A folyamatos osztódási potenciál kialakulásáért elsőként a növekedési faktor receptorokban vagy a receptorhoz kapcsolt jelátviteli folyamat egyéb más elemeit érintő szomatikus mutációk tehetőek felelőssé. Erre példa a BRAF, mint szerin-threonin kináz gént érintő mutációk, amelyek következtében a MAP kináz kaszkádon keresztül folyamatosan aktív sejtosztódás indukálódik (7). Hasonló hatása van a foszfoinozitol 3 kináz gén izoformákat érintő szomatikus mutációknak is (8). Továbbá, fontos kiemelni még azon géneket – PTEN és mTOR –, amelyek a sejtciklus negatív szabályozásában vesznek részt és az őket érintő mutációk szintén fokozott sejtosztódáshoz vezetnek (8), (9).

### 2.2.2 A tumorszuppresszorok visszaszorítása

A sejtosztódás fenntartása mellett a daganatok kialakulásában jelentős szerep jut a tumorszuppresszorok visszaszorításának is. Ezekre példaként a TP53 és a RB géneket érdemes megemlíteni, melyeknek a feladatuk az, hogy szabályozzák a sejtciklust, azaz ők döntenek arról, hogy a sejt beléphet-e a sejtciklusba vagy pedig sejtöregedés és a sejthalál útjára lépjen (10). A sejtek egymás közötti kontakt gátlásának az elkerülése szintén ehhez a csoporthoz tartozik. A kontakt gátlás folyamatának az elkerülése kóros sejtnövekedéshez, karcinogenezishez vezet. Az NF2 génnek kiemelt szerepe van a sejtfelszíni adhézións molekulák és a tirozin kináz doménnel rendelkező receptorok közötti kapcsolódás szabályozásában, amely során kontrollált sejtosztódás válik lehetővé (11). A kontakt gátlás folyamatánál az LKB1 tumorszuppresszor gént is fontos megemlíteni, mivel jelentős szerep játszik a sejtfelszíni kapcsolóstruktúrák és a sejntegritás szabályozásában, hiányában jelentős mitotikus aktiválódás történik (12). Utolsóként az antiproliferatív TGF- $\beta$  receptorhoz kapcsolt jelátviteli útvonalat említeném meg,

amelynek aberráns működése számos késői stádiumú daganat esetén az epitheliális-mezenchimális átalakulással hozható összefüggésbe (13).

### **2.2.3 A sejthalál kikerülése**

A daganat patogenezise során az intrinszik, azaz a mitokondriumhoz kapcsolt sejthalál útvonal jelent számottevőbb akadályt (14). A sejthalál szabályozásában kiemelt szerepet kap a már korábban említett TP53 tumorszuppresszor (10). Ezen fehérje érzékeli a DNS száltöréseket, valamint egyéb kromoszómális elváltozásokat és apoptózist indukál a Noxa és a Puma fehérjék felülregulálásán keresztül. A daganatok esetében a leggyakoribb a TP53 gén tumorszuppresszor funkciójának az elvesztése, amely során nem történik meg a DNS károsodásának az érzékelése és a sejt kikerüli a sejthalált (10). Továbbá, hasonló hatás érhető el az antiapoptotikus fehérjék kifejeződésének a növelésével – ilyen például a Bcl-2 fehérje –, valamint a proapoptotikus fehérjék – Bax, Bim, Puma – szintjének a csökkentésével (15). Az apoptózis elkerülése mellett az autofágia is jelentős szereppel bír a tumorigenezisben. A BECN1 (Beclin-1) gén minkét alléljának vagy az autofágiában szerepet játszó egyéb komponensek kiesése a daganat kialakulásának kedvez (16).

### **2.2.4 Korlátlan osztódási képesség**

A daganatos sejteknek korlátlan replikációs potenciálra van szükséges ahhoz, hogy a makroszkopikus méretű tumor létrejöhön. A számtalan osztódás során a telomerek védik a kromoszómák végeit a rövidüléstől. A telomerázok felelősek a telomerek meghosszabbításáért, amelyek a daganatos sejtekben fokozott génkifejeződést mutatnak (17). Azonban a telomerázoknak nemcsak a telomerák meghosszabbításában van szerepük, hanem a sejtosztódásban is szerepet játszanak, függetlenül telomerek fenntartásától (18). Továbbá, a sejthalál elkerülésének a kialakításában, DNS hibajavítási folyamatokban és RNS függő RNS polimeráz működésének a szabályozásában is részt vesz (19), (20).

### **2.2.5 Érképzés serkentése**

A daganatok fennmaradásához és nevekedeése növekedéséhez megfelelő érellátottságra van szükségük, amely biztosítja számukra az oxigén és az oldott

tápanyagok felvételét, valamint az anyagcseretermékek leadását (21). Fontos kiemelni, hogy a daganatok érképző képessége különböző lehet. Egyes típusok, mint például a hasnyálmirigy duktális adenokarcinóma kifejezetten avaszkuláris, azaz a daganatszövet érellátása kevés (22). Azonban a hasnyálmirigy neuroendokrin karcinóma, valamint a veserákok esetén az érképző képesség igen jelentős (23), (24). A vaszkuláris endotheliális növekedési faktor-A (VEGFA) és a thrombospondin-1 (TSP1) géneknek kiemelkedő jelentőségük van az érképződés szabályozásában (25). A VEGFA serkenti, míg a TSP1 pedig gátolja az érképződést (25). A VEGF-hez kapcsolódó jelátviteli útvonal tirozin kináz aktivitással rendelkező receptorokon (VEGFR1-3) keresztül szabályozódik (26). Számos korábbi vizsgálat igazolta, hogy a VEGF génkifejeződését számos tényező befolyásolhatja (26). Így például a hypoxia és az onkogén stimulus (RAS és MYC) jelentősen növelte a VEGF kifejeződését daganatban (26). Továbbá, a daganatban megfigyelték, hogy a VEGF-en kívül számos más érképződéssel kapcsolatos komponens szabadul fel eredetileg inaktív állapotban (MMP9), amelyeket az extracelluláris mátrix proteázai aktiválnak (27). A VEGF-hez hasonlóan a fibroblaszt növekedési faktor (FGF) szintén felülregulálódik a daganatban hozzájárulva az érképződés folyamatához (25). A TSP1, szemben a VEGF és az FGF érújdonképzésben szerepével, az érendothel sejtek felületén található receptorokhoz kötődik és gátolja az érképzés folyamatát, így létrehozva egy stabil egyensúlyt (28).

### **2.2.6 Invázió és metasztázis indukálása**

A rosszindulatú daganatoknak szintén egy másik fontos tulajdonsága az áttétképzés, vagyis az invázió és a metasztázis. Ez egy több fázisú folyamat, amely során a daganatsejt az ér vagy a nyirok pályába belépve eljut a test egy távolabbi pontjára, ahol makroszkopikus tumor képes kialakulni (29). A folyamat első pontja a lokális invázió, ahol a daganatsejtek áttörnek a bazális membránt és ezt követően jutnak be a vér- vagy a nyirokkeringésbe. Itt a keringés alakos elemeihez tapadva elszállítódnak távolabbi helyekre, ahol úgynevezett extravazációval kilépnek a keringésből, majd a célszerveben mikrometasztázisokat hoznak létre. Végül a mikrometasztázisok a daganatsejtek folyamatos osztódása révén makroszkopikus méretű daganatokká fejlődhetnek (29). A normál sejtek esetén a sejt-sejt adhézió kialakításában az E-cadherin, az egyik kulcsfontosságú elem. Az E-cadherin fokozott génkifejeződése hozzájárul a sejtek

nyugvó állapotának a fenntartásához, csökkentve a sejtek inváziós és metasztatizáló képességét. A daganatok esetében, azonban az E-cadherin alacsony génkifejeződést mutat, valamint gyakori a gén mutációk által történő inaktiválódása is (30). Továbbá, a sejt-sejt és a sejt-extra celluláris mátrix kapcsolatában részt vevő gének közül fontos kiemelni az N-cadherint is. Ezen gén normál kifejeződése szükséges a neuronok és a mezenchimális sejtek migrációjához, azonban a magas kifejeződése daganatos fenotípussal társul (30). Az epitheliális sejtek nyugvó állapotú sejtek, azaz az inváziós és metasztatikus képességük nagyon alacsony, ellenben a mezenchimális sejtekével. A daganatok fokozott inváziós és a metasztatikus képességéhez nagymértékben hozzájárulnak, az ezen két állapot közötti átalakulásban, vagyis az epitheliális-mezenchimális átalakulásban részt vevő gének fokozott kifejeződése (31). Közöttük szerepel számos transzkripciós faktor, mint például a Snail, Slug, Twist, és Zeb1/2, amelyek a daganatsejtekben különböző kombinációkban fejeződnek ki és járulnak hozzá a daganatsejtek migrációs képességéhez (31).

### **2.2.7 A sejtek energiametabolizmusának az átprogramozása**

A daganatos sejtek energia metabolizmusának az átprogramozása nélkülözhetetlen a korlátlan sejtosztódás támogatásához. Normál sejtek aerob (oxigén dús) körülmények között a glikolízis során, a glükózt piruváttá alakítják a sejtplazmában, amely majd a mitokondriumban energiává alakul a citomsavciklus és a terminális oxidáció révén. Anaerob (oxigén hiányos) körülmények között a glikolízisben keletkező piruvát nagy része tejsavvá alakul és csak kis része hasznosul a mitokondriumban. A daganatok esetében a glikolízis az anaerob glikolízis mintáján alapszik, a különbség az, hogy aerob körülmények között megy végbe (32). Ezt a folyamatot 1920-ban Otto Heinrich Warburg írta le, és ezért Warburg effektusnak vagy más néven aerob glikolízisnek hívják (32). Ezt a daganatsejtek számos gén működésén átalakítása révén érik el. Az egyik ilyen gén a GLUT1, amely a glükóz sejtplazmába történő szállításáért felel. A daganatsejtek ezen transzporter génkifejeződésének a fokozása révén érik el a glükóz felhalmozódást a sejtplazmában (33). A daganatok megváltozott glikolízise összefügg az onkogének aktiválódásával (RAS, MYC) és a mutáns TP53 gén működésével, amely fokozott sejtosztódást és a sejthalál elkerülését eredményezi (33). Továbbá, kiemelendő még az izocitrát dehidrogenáz  $\frac{1}{2}$  (IDH1/2) gének funkció nyerő

mutációi, amelyet számos tumortípusban, többek között a gliómában írtak le (34). A mutáns IDH1/2 gének működése összefügg a HIF1 transzkripciós faktorok aktivitásával, amelyek befolyásolhatják a daganatok genom stabilitását és az érképződését (35).

### **2.2.8 A genom stabilitás fenntartása**

A daganatok heterogenitását számos modell alapján lehet értelmezni, amely során a szerzett mutációk az egyes daganatsejteknek különféle szelekciós előnyt biztosítanak a környező sejtekhez képest (például gyógyszer szenzitivitás). Így a daganatsejteknek egy adott szövetkörnyezetben lehetőségük van kiválni és megszerezni a dominanciát a többi sejt felett (36). A korábbi évek során számos olyan gént azonosítottak, amelyek a sejt DNS tartalmának a védelmére vannak szakosodva (37). Vannak olyan gének, amelyek azonosítják a DNS károsodást és DNS hibajavítási folyamatot indukálnak, vannak akik közvetlenül a DNS hibajavításban vesznek részt, valamint egyesek pedig megpróbálják hatástalanítani a DNS károsodást okozó molekulákat (38). Ezen gének nagy többsége mind tumorsuppresszor, a bennük létrejövő mutáció vagy epigenetikai módosulás daganatprogressziót indukál (38).

### **2.2.9 A gyulladás serkentése**

A daganatos szövetben megfigyelhetőek a természetes és az adaptív immunitás sejtjei, amelyek a nem tumoros szövet gyulladásos állapotát reprezentálják (39). A gyulladás összefügg számos más daganatot kiváltó tényezővel. A gyulladás során különböző bioaktív molekulák juthatnak a tumor mikrokörnyezetébe, például növekedési faktorok, amelyek indukálják a sejtosztódást, valamint olyan faktorok, amelyek a daganatsejtek sejthalál kikerülését eredményezhetik, extra celluláris mátrixot módosító enzimeket aktiválhatnak, amelyek serkentik az érképződést, és az áttétképzést (40). Továbbá, a gyulladásos sejtekből reaktív oxigén gyökök is felszabadulhatnak, amelyek a környező rosszindulatú daganatos sejtek számára rendkívül mutagén, így felgyorsítva a daganatsejtek genetikai evolúcióját az agresszív daganatos fenotípus felé (40).

### **2.2.10 Az immunrendszer visszaszorítása**

Az immunrendszer és a daganat progressziója közötti összefüggés máig egy viszonylag kérdéses terület. Korábbi vizsgálatok mutatják, hogy az immunhiányos



egerekben sokkal gyorsabb a daganat növekedése, szemben a megfelelő immunrendszerrel ellátott kontroll egércsoporttal (41). Továbbá, vastagbél és petefészek daganatok esetén kimutatták a nagyfokú citotoxikus T sejtes, valamint természetes ölősejtes beszűrődést, amely jobb betegség túléléssel társult (42). A folyamatos immunfelügyelet során az immunrendszer elemei megpróbálják az újonnan kialakuló daganatsejteket felismerni és eltávolítani, azonban egyes daganatsejteknek, vagy sejtklónoknak bizonyos mértékig sikerül távol tartaniuk magukat az immunfelügyeletről (43). Megfigyelték, hogy az egyes daganatklónok TGF- $\beta$  és egyéb immunoszuppresszív faktorok szekréciója során képesek megbénítani az immunsejteket, így biztosítva a daganatok saját fennmaradásukat (44). Valamint még további immunrendszer ellenes mechanizmusuk közé tartozik az, hogy képesek immunszuppressziót kiváltani a regulátor T sejtek (Tregs) és a mieloid szuppresszor sejtek (MDSC) toborzásával, amelyek gátolják a citotoxikus T sejtek működését (45), (46).

### **2.3 Egyes daganattípusok jellemzése, molekuláris terápiai célpontok**

Az értekezésem alapját főként a tüdő-, a máj-, a gyomorrák és az akut mieloid leukémia daganattípusokon alapuló tanulmányok képezik, ezért a következő alfejezetekben az egyes daganattípusokat és azok prognosztikus és prediktív biomarkereit ismertetem.

#### **2.3.1 A tüdődaganatok**

A tüdődaganat az egyik leggyakoribb és a legnagyobb a halálozási rátával rendelkező daganattípus mind a férfiak és a nők körében (1), (2).

A tüdődaganatok szövettanilag két csoportra oszthatóak: kissejtes tüdőrák és a nem kissejtes tüdőrák. Az összes tüdőrákos eset 15%-át a kissejtes, 85%-át pedig a nem-kissejtes tüdőrák teszi ki (47).

A kissejtes tüdőrák kialakulása teljes mértékben a dohányzáshoz köthető és az előfordulási gyakorisága a nők és a férfiak körében egyenlő arányban van (48). Az osztódási gyakorisága sokkal gyorsabb a nem-kissejtes tüdőrákhoz képest és már a daganat korai stádiumában is képes az áttétképzésre. A daganat kezelése során a kezdeti kemoterápia nagyon hatásos, ennek ellenére magas a halálozási arány a daganat kiújulása kapcsán (49). Bár a célzott terápia a kissejtes tüdőrák esetén jelenleg még nem elérhető,

azonban számos korábbi vizsgálatban olyan molekuláris eltéréseket azonosítottak, amelyek lehetséges célpontok lehetnek a daganat kezelése során. Az egyik ilyen célpont, a sokszorozódott SOX2 gén, amely hozzájárul a daganatsejtek osztódási képességének a növekedéséhez (50). A SOX2 sokszorozódása mellett, a RLF-MYCL1 génfúzió, valamint a CREBBP, EP300, és a MLL gén mutációk szintén a sejtosztódási programot gyorsítják a kissejtes tüdőrák esetén (51). Továbbá, ezen daganattípusra szintén jellemző a TP53, a RB1 gének, valamint az egyes NOTCH géncsaládba tartozó gének funkcióvesztő mutációinak jelenléte (52), (53). A kissejtes tüdőrák esetén a génmutációk mellett fontos kiemelni a génkifejeződési változásokat is. Egy korábbi génexpressziós vizsgálatban olyan géncsoportot azonosítottak, amelyek alapján a kissejtes tüdőrák két fő csoportra osztható: az első csoport a CHGA, a GRP, az ASCL1 és a DLK1 géneket alacsonyan, a második csoport ezen géneket magasan kifejező minták csoportja (52). További génexpressziós vizsgálatok során sikerült olyan géneket azonosítani, amelyek hozzájárulnak a daganat növekedéséhez, valamint szerepet játszanak a kemoterápiás válasz kialakításában. Ilyenek a TAF5L, TFCP2L4, PHF20 és az LMO4 gének fokozott kifejeződése, amely a kissejtes tüdőrák kemorezisztenciájával függ össze (54).

A tüdőrákos esetek másik, egyben a legnagyobb csoportját a nem-kissejtes tüdőrák teszi ki, amely számos további altípusba osztható: tüdő adenokarcinóma, laphámsejtes karcinóma és nagysejtes karcinóma (47). A tüdő adenokarcinóma a leggyakoribb tüdőrák altípus, az esetek mintegy 40%-ában fordul elő függetlenül a kortól, a nemtől és a dohányzástól. A daganat progressziója a többi altípushoz képest sokkal lassabb, így a korai diagnosztizálása jelentősen hozzájárul a várható túlélés növekedéséhez (55). A laphámsejtes karcinóma, az esetek 30%-át teszik ki, a betegség kialakulása szorosan korrelál a dohányzással (56). A nagysejtes karcinóma 5-10%-os gyakorisággal bír a tüdőrákos betegek esetén és hasonlóan a laphámsejtes karcinómához, előfordulása szorosan összefügg a dohányzással. Fontos megemlíteni, hogy ezen daganat altípusra nagyfokú differenciálatlanság jellemző, amely jelentősen megnehezíti a pontos diagnózis megállapítását (55), (57).

### **2.3.1.1 Terápiás célpontok a nem-kissejtes tüdőrákban**

Szemben a kissejtes tüdőrákkal, számos molekuláris célpont áll rendelkezésre a nem-kissejtes tüdőrák célzott terápiája esetén. Jelenleg már a forgalomban vannak olyan

gyógyszerek, amelyek az epidermális növekedési faktor receptor (EGFR) mutációkra, vagy az anaplasztikus lymphoma kináz (ALK) gént érintő molekuláris változásokra specifikusak. Jelenleg számos olyan gyógyszer van, ami még csak pre-klinikai vagy klinikai kísérleti fázisban van, ezek például a KRAS, BRAF, ERBB2, PIK3C génmutációkra, a RET és ROS gének transzlokációjára és a c-MET amplifikációra specifikus gátlók (58).

#### **2.3.1.1.1 Az epidermális növekedési faktor receptor (EGFR) mutációk**

Az epidermális növekedési faktor receptor (EGFR) mutációk voltak az első molekuláris célpontok, amelyeket jóváhagytak a nem-kissejtes tüdőrák terápiája során (59). Az EGFR egy sejtfelszíni receptor, amely a sejtek környezetéből érkező jelek hatására aktiválódik és a sejt belseje felé sejtosztódási és sejtnövekedési jelet továbbít (60).

Az EGFR receptort érintő mutációk legtöbb esetben a receptor sejt belüli részén található tirozin kináz domént érintik. Ezen mutációk a tüdődaganatos betegekben 20%-os gyakorisággal fordulnak elő, főként az európai és ázsiai betegpopulációkban, továbbá a nem dohányzók és a nők körében (61). Ilyen a mutációk például a 19-es exont érintő deléción, vagy a 18-as és a 21-es exont érintő pontmutációk, amelyek esetén nő a daganat érzékenysége az – erlotinib (62), gefitinib (61) és afatinib (63) – EGFR tirozin kináz gátlókra.

#### **2.3.1.1.2 Az anaplasztikus lymphoma kináz (ALK) mutációk**

Az anaplasztikus lymphoma kináz (ALK) gént érintő mutációk a tüdőrákos betegek mintegy 2-7%-ban találhatóak meg és főként a tüdő adenokarcinóma altípusra jellemzőek.

Az egyik leggyakoribb ALK gént érintő molekuláris változás az EML4-ALK génfüzió, amely a nem vagy ritkán dohányzók körében gyakoribb és nem fordul elő más onkogenikus driver mutációval együtt, úgymint az EGFR, a KRAS vagy a HER2 géneket érintő mutációkkal (64), (65).

Egy korábbi vizsgálat során megfigyelték, hogy az ALK gént érintő átrendeződés befolyásolta a tüdő adenokarcinómás betegekben a EGFR tirozin kináz gátlókkal szembeni rezisztenciát, valamint az ALK gátlókra – crizotinib és ceritinib – való

érzékenységet (66).

### **2.3.1.1.3 A BRAF mutációk**

A BRAF úgynevezett proto-onkogént érintő mutációk, a nem-kissejtes tüdőrákos betegek 1-4%-ában fordulnak elő és főként az adenokarcinóma altípusra jellemzőek (67). A BRAF mutációk, ellentétben az EGFR és az ALK gént érintő mutációkkal, az aktív dohányosok körében a gyakoribbak (67).

A legismertebb BRAF mutációk a nem-kissejtes tüdőrákban a V600E, G469A és a D594G eltérések, amelyek közül a leggyakoribb V600E mutáció, amely az összes BRAF mutáció 50%-át teszi ki (67). A BRAF mutációk jelenléte kizárja a többi onkogén mutációk jelenlétét, úgymint az EGFR, ALK és a KRAS mutációkat (67).

Jelenleg a BRAF mutációk esetén a tüdőrákban nincsen elfogadott célzott terápiás gyógyszer.

### **2.3.1.1.4 A KRAS mutációk**

A RAS géncsalád három a plazmamembránhoz kötött fehérjét kódol, amelyek saját guanozin-trifoszfát (GTP) bontó aktivitással rendelkeznek, úgymint a KRAS, NRAS és a HRAS. A RAS fehérjék egy molekuláris kapcsolóként működnek, amely során a sejtfelszíni receptoroktól a jelet továbbítják a sejtmag felé, ahol sejtosztódási és sejtnövekedési programokat aktiválhatnak (68). A KRAS, NRAS és HRAS gének számos daganattípusban hordoznak mutációkat, többek között a nem-kissejtes tüdőrákban is (69).

A KRAS gén, a tüdő adenocarcinómás betegek mintegy 25%-ban hordoz eltérést és a mutációk gyakorisága a dohányzók és a kaukázusi betegpopuláció körében a legnagyobb (70). A KRAS gén gyakori mutációi közé tartoznak a G12C, G12D és a G12V mutációk, amelyek az enzimet folyamatosan aktív állapotban tartják (71).

Egyelőre nem tisztázott, hogy a KRAS génben található szomatikus mutációk közvetlen oka-e a rossz túléléshez és a terápiás rezisztenciához, vagy valamilyen más közvetett úton. A KRAS mutációk prognosztikus hatását először vastagbél-daganatban írták le (72), azonban ezeket az eredményeket későbbi tanulmányok nem tudták alátámasztani (73), (74). Számos meta-analízis mutatja, hogy a KRAS mutációk hatása jelentősen különbözött az ázsiai és a nem ázsiai populációk között, és az ázsiai betegek körében a KRAS mutációk rossz prognosztikai faktoroknak bizonyultak (75).

Korábbi vizsgálatok azt mutatják, hogy azon nem-kissejtes tüdőrákos betegek esetében, amelyek KRAS mutációkat is hordoznak, EGFR tirozin kináz gátlókkal szembeni rezisztencia lép fel (76). Ez a KRAS a jelátviteli útvonalban való elhelyezkedésével magyarázható, mivel a KRAS fehérje az EGF receptortól kapja a jelet és továbbítja azt a sejtmag felé. Így a mutáns KRAS fehérje, az EGF receptortól függetlenül képes a jelátviteli kaszkád aktiválására. A KRAS mutációk negatív előrejelzői az EGFR tirozin kináz gátlókkal történő terápiának, mivel ezen gyógyszerekre a KRAS mutáns betegek nem fognak reagálni (77).

Korábbi meta-analízisek révén igazolták az összefüggést a KRAS gén mutációs státusza és az EGFR tirozin kináz gátlókra adott terápiás válasz között (78), amelyet mára már az NCCN leírás is tartalmaz (<https://www.nccn.org/>). Így, a KRAS mutációk magasabb, mint 99%-os prediktív értékkel, ideális prediktív biomarkerként szolgálnak a nem-kissejtes tüdőrák terápiája során (78).

A KRAS mutáns tüdődaganatok célzott terápiáját képezheti, a KRAS-tól a sejtmag felé irányuló jelátviteli útvonalak – RAF-MEK-ERK és a PI3K-AKT-mTOR – gátlása. Egy korábbi tanulmány mutatja, hogy a KRAS mutáns nem-kissejtes tüdőrák egér xenograft modellekben, a selumetinib – amely egy MEK1/2 gátlószer –, szignifikánsan csökkentette a daganat növekedését (79). Egy másik tanulmányban a selumetinibet docetaxel kemoterápiás kezeléssel kombinálták, és igaz, sokkal több mellélhatás jelentkezett, de a medián progresszió mentes túlélés 5,3 hónapra nőtt az eredeti 2,1 hónaphoz képest (80). A jelen álláspont szerint a KRAS mutáns nem-kissejtes tüdőrákos betegek esetén nincs hivatalosan elfogadott terápiás protokoll.

### **2.3.2 A májdaganatok**

A májdaganatok, a hatodik leggyakrabban diagnosztizált daganattípusok világszerte, az esetek legnagyobb részét többnyire Észak- és Kelet-Afrika, valamint Kelet- és Délkelet-Ázsia egyes országainak betegei adják (1).

A májdaganatok szövettani besorolásuk alapján több altípusra oszthatók, amelyek eltérő prognosztikus tulajdonsággal rendelkeznek. Az egyik leggyakoribb májdaganat típus a hepatocelluláris karcinóma, amely az összes májrákos eset mintegy 90%-át képezi és a legrosszabb prognózissal rendelkezik (81). A WHO osztályozás szerinti további májdaganat altípusok az intrahepatikus cholangiocarcinoma, a hepatocelluláris

cholangiocarcinoma, a fibrolamelláris hepatocelluláris karcinoma és a gyermekkori hepatoblasztóma (82). Ezen tumortípusok esetén a betegség prognózisa az intrahepatikus cholangiocarcinomától a gyermekkori hepatoblasztóma felé haladva egyre javuló tendenciát mutat (82).

A hepatocelluláris karcinóma incidenciája és mortalitása évenként nő, amelyhez számos tényező kapcsolódik, mint például a hepatitis B vírus (HBV) és a hepatitis C vírus (HCV) fertőzések, az élelmiszerek aflatoxin tartalma, az elhízás, a diabétesz, a dohányzás, illetve a túlzott alkoholfogyasztás (1). Fontos kiemelni, hogy ezen kockázati tényezők, az egyes etnikai csoportok között eltérő megoszlást mutatnak. Ennek alapján például az ázsiai országokban főként a HBV és a HCV fertőzések kapcsolhatók a legtöbb hepatocelluláris karcinómás megbetegedés kialakulásához (1). A fejlett országokban a májdaganatok kialakulása krónikus májbetegségekre vezethető vissza. Ilyen betegségek lehetnek a túlzott alkoholfogyasztásból adódó májcirrózis, a diabétesz, valamint a zsírmáj (83).

### **2.3.2.1 Biomarkerek a hepatocelluláris karcinómában**

A hepatocelluláris karcinómában számos olyan biomarkert azonosítottak, amelyek lehetővé teszik a daganat pontosabb molekuláris besorolását, előrejelzik a várható túlélést és a terápiára adott választ (84).

A hepatocelluláris karcinóma egy meglehetősen heterogén tumor, amelynek a molekuláris osztályozása jelenleg nem ismert (85). A mutációk száma tumoronként körülbelül 40-50 darab, ezek között 5-8 driver mutáció található (85). A citotoxikus kemoterápia hatástalan vagy annyira toxikus, hogy már a cirrhotikus májat is roncsolja. A prognózis számos klinikai-patológia paraméteren alapul, úgymint a tumorsejtek aránya, proliferációs markerek, vaszkuláris invázió, a májfunkció és az általános egészségi állapot. Jelenleg számos májspecifikus biomarker van klinikai tanulmány alatt, amely során vizsgálják, hogy melyik az a terápia, amelyik beteg számára a legjobb és a legkevesebb mellékhatással jár (86).

#### **2.3.2.1.1 Génexpresszió alapú biomarkerek**

Microarray génexpressziós vizsgálatok során a hepatocelluláris karcinóma két fő csoportra bontható (87). Az egyik csoportot nagyfokú sejtsztódási potenciállal

rendelkező sejtek alkotják, míg a másik csoport sejtjeinek osztódási potenciálja meglehetősen alacsony. A gyorsan osztódó típus esetén magas a daganat kiújulási kockázata és jelentősen alacsonyabb a túlélési idő (87). Fokozott génkifejeződés figyelhető meg az IGF1, az MTOR és a NOTCH géncsalád egyes génjei esetén, továbbá ezen daganat altípusra jellemző még az eltérő mikro-RNS kifejeződési és DNS metilációs mintázat, valamint a CCND1, FGF19 és a VEGFA gének sokszorozódása is (87). A másik csoportot a kismértékben osztódó altípus képviseli, amelyre jellemző az EGF receptorhoz kapcsolt jelátviteli útvonal egyes elemeit érintő elváltozások, a gyulladási folyamatok aktiválódása, valamint a daganat nagyfokú differenciáltsága és az alacsony alfa-fetoprotein szintje (87).

### **2.3.2.1.2 Génmutáció alapú biomarkerek**

A génexpresszió alapú biomarkerek mellett fontos kiemelni génmutációkat is. Teljes genom és teljes exom szekvenálással mutációs mintázatokat azonosítottak hepatocelluláris karcinómában, amelyek szoros összefüggést mutattak például a dohányzással és az alkoholfogyasztással (85). Teljes exom szekvenálással olyan gén mutációkat azonosítottak –TERT, CTNNB1, TP53, AXIN1 és ARID1A –, amelyek hozzájárulnak a hepatocelluláris karcinóma patogeneziséhez (88), (89). Jelentős összefüggést írtak le hepatocelluláris karcinóma sejtvonalakban a TSC2 gént érintő kópiaszám változások és az MTOR jelátvitel között. Megfigyelték, hogy a TSC2 gént érintő törlődések növelték a daganatsejtek érzékenységét az everolimusra, amely az MTOR fehérje kinázra specifikus gátlószer (90).

A hepatocelluláris karcinóma kialakulásában és patogenezisében nagyon fontos szerepet játszanak a HBV és HCV vírusok. Teljes genom szekvenálás során lehetővé vált a vírusgenomok daganatgenomba történő integrálódásának a vizsgálata, amely során megfigyelték, hogy a HBV vírus daganatgenomba történő integrálódása esetén a TERT, MLL4 és a CCNE1 gének működése megváltozik, amely hatással van a betegek túlélésére (91). Hasonlóan a HBV vírushoz, az adeno-asszociált vírus 2 (AAV2) genom beépülést is kimutatták a hepatocelluláris karcinómában, amely az összes betegnek mintegy öt százalékában volt jelen és a TERT, a KMTB2, a CCNA2 és a CCNE1 gének működésével mutatott összefüggést (92).

### 2.3.2.1.3 mikro-RNS alapú biomarkerek

A mikro-RNS-ek kisméretű nem kódoló RNS molekulák, amelyek a humán gének mintegy 10%-ának a transzkripció és poszt-transzkripció szinten történő szabályozásában vesznek részt (93). A humán mikro-RNS-ek („hsa-miR”) jelentős szerepet játszanak a daganatok kialakulásában, fejlődésében és a terápiás válasz kialakításában. Különböző tumorokban egyes mikro-RNS-ek onkogén vagy akár tumorszuppresszor tulajdonsággal is rendelkezhetnek (94). Ennek alapján a nem-kissejtes tüdőrákban a hsa-miR-374a mikro-RNS alacsony kifejeződése a daganat fokozott progressziójával társul, amely rossz túléléshez vezet (95). A kolorektális daganatokban a hsa-miR-185 magas, míg a hsa-miR-133b mikro-RNS-ek alacsony kifejeződése áttétképzéssel és rosszabb túléléssel korrelál (96). Számos további mikro-RNS van, amelyek egy adott tumortípusra specifikus. Ilyen a hsa-miR-34a, amely az emlőrákra, a hsa-miR-10b pedig a hasnyálmirigy rák kialakulásával és növekedésével hozható összefüggésbe (97), (98).

Az eltérő kifejeződést mutató mikroRNS-ek a hepatocelluláris karcinómában potenciális biomarkerként szolgálhatnak a daganat diagnózisára, az áttétképző és inváziós képességére, a terápiás válasz előrejelzésére, a daganat kiújulására és a túlélésre (94), (99), (100). A hepatocelluláris karcinómában alulregulált mikro-RNS-ek a hsa-miR-200c, hsa-miR-200, hsa-miR-21, hsa-miR-224, hsa-miR-10b és a hsa-miR-222, valamint a hsa-miR-224 esetén a magas kifejeződés jellemző (101). Vannak olyan mikro-RNS-ek, amelyek a hepatocelluláris karcinóma valamelyik rizikó tényezőjével függenek össze, azaz például a hsa-miR-96 fokozott kifejeződést mutatnak a HBV fertőzött daganatos betegek esetén (101). Továbbá, a hsa-miR-126 alulszabályozódik azon májdaganatok esetén, amelyek kialakulása a jelentős alkoholfogyasztás eredménye (101). Továbbá, fontos kiemelni a keringő mikro-RNS-eket – hsa-miR-939, hsa-miR-595, hsa-miR-519d és a hsa-miR-494 –, amelyeknek fontos szerepük van a daganat diagnosztizálásában májcirrhosisban érintett betegek esetén (102).

### 2.3.2.2 A hepatocelluláris karcinóma célzott terápiás lehetőségei

A hepatocelluláris karcinóma gyors növekedése, nagy áttétképző képessége és a gyógyszerrezisztenciára való fokozott hajlama révén rossz prognózissal bír. A legtöbb



esetet közepes vagy már jelentősen előrehaladott stádiumában diagnosztizálják, és a betegek mindössze 30%-a részesül valamilyen terápiás kezelésben (103).

A hepatocelluláris karcinóma első vonalbeli célzott terápiája esetén jóváhagyott a sorafenib, amely egy receptor tirozin kináz típusú gátlószer (104). A sorafenib képes gátolni a RAF1, BRAF, VEGFR és a PDGFR-B géneken keresztül a sejtosztódást és a daganat érképzését, valamint serkenti az apoptózist. A gyógyszer terápiás alkalmazását a jelentősen előrehaladott stádiumú – vaszkuláris inváziót és májszöveten kívüli terjedést is mutató – hepatocelluláris karcinómás betegek körében hagyták jóvá (104). A sorafenib alkalmazása során megfigyelték, hogy a kelet-ázsiai, csendes-óceáni és az ausztráliai betegpopulációkban, a kezelés során a betegek medián teljes túlélése a 4,2 hónapról 6,5 hónapra nőtt (105). Azonban fontos kiemelni, hogy a kezelés során a mellékhatások fokozódnak és ez jelentősen hozzájárul az életminőség romlásához. A sorafenib mellett, a hepatocelluláris karcinóma első vonalbeli kezelése során jóváhagyott a lenvatinib, amely egy szintén tirozin kináz típusú gátlószer és a VEGFR, FGFR, PDGF receptor, valamint a RET és a KIT onkogének gátolhatók (106).

A hepatocelluláris karcinóma célzott terápiája során másod vonalban jóváhagyott a regorafenib, receptor tirozin kináz típusú gátlószer, amely a VEGF receptorhoz kötődve gátolja a daganat érképzését (107). Korábbi vizsgálatok mutatják, hogy a regorafenib kezelés hatására szignifikánsan nőtt mind a teljes, mind pedig a progresszió mentes túlélés is (107).

A mikro-RNS-ek terápiás célpontként is szolgálhatnak a hepatocelluláris karcinóma terápiája során. Sejtkultúras vizsgálatokban igazolták, hogy egyes mikro-RNS-ek képesek a sorafenib kezelés hatékonyságát befolyásolni. A hsa-miR-122 magasabb kifejeződése a sejtek sorafenib érzékenységét befolyásolta, míg a hsa-miR-34a alacsonyabb kifejeződése a sorafenib rezisztenciával korrelált (108), (109).

A mikro-RNS-ek közvetlenül is gátolhatóak úgynevezett mikro-RNS antagonistákkal, amelyek kémiaiilag módosított zárt láncú RNS molekulák és az egyik fontos tulajdonságuk az, hogy védettek az RNS lebontó enzimekkel (RNázok) szemben (110). Számos olyan sejtvonalakon valamint egérmodellben végzett vizsgálat létezik, ahol a mikro-RNS antagonisták terápiás hatását vizsgálják. Az első tanulmányban hepatocelluláris karcinóma sejtvonalakon végzett vizsgálat során a sejtek kemoterápiára – IFN- $\alpha$ /5-fluorouracil (5-FU) kombináció – adott válasza javult, az antiszensz hsa-miR-

21 mikro-RNS transzfekciójával (111). Egy másik vizsgálatban pedig transzgenikus egérmodellben vizsgálták, hogy a hsa-miR-221 gátlásával szignifikánsan csökkent a tumor mérete és a száma, valamint nőtt a teljes túlélési idő is (112). A mikro-RNS kifejeződésnek fontos szerepe van hepatocelluláris karcinóma interferonokkal történő terápiájában is. Egy szintén sejtkultúrán végzett vizsgálat bizonyítja, hogy a magasabb hsa-miR-146 kifejeződés interferon rezisztenciát, míg az alacsony hsa-miR-26 kifejeződés interferon terápiával szembeni fokozott érzékenységet okozott hepatocelluláris karcinóma sejtvonalban (113), (114).

### **2.3.2.3 A vaszkuláris invázió szerepe a daganat kiújulásában**

A hepatocelluláris karcinóma kiújulásával erősen korrelál a vaszkuláris invázió, amely a portális vagy a májvénán keresztül történik. Egy korábbi tanulmány során megfigyelték, hogy májtranszplantációval érintett betegek esetén, a hepatocelluláris karcinóma kiújulása több mint 4-szer nagyobb (115). Ugyanakkor, azonban a pre-operatív képalkotó eljárások során nem figyelhetők meg a mikrovaszkuláris invázió jelei (116). A boncolási adatok alapján kiderült, hogy a betegek 40%-a esetén, az 5cm-nél kisebb daganatoknál már megfigyelhetők voltak az invázió jelei (117). Így, a vaszkuláris invázióval összefüggő molekuláris változásoknak kiemelkedő szerepük van a túlélés előrejelzésében, amelyek nagymértékben hozzájárulnak a hepatocelluláris karcinóma kiújulásához.

### **2.3.3 Az akut mieloid leukémia**

Az akut mieloid leukémia, a mieloid daganatok csoportjába tartozik, amely során a sejtek, a mieloid sejtdifferenciáció valamelyik korai fázisában rekednek meg. A sorozatos sejtosztódások során a felszaporodott úgynevezett mieloid sejtek vagy blasztok, kiszorítják a normális fejlődési út során létrejövő vörsejteket.

Az akut mieloid leukémia mintegy 1,1%-át képezi az összes új rosszindulatú daganatos megbetegedéseknek, és az idősebb korosztályban, azon belül is a férfiak körében a leggyakoribb. A halálozási gyakoriság szintén az idősebb korosztály (65 év feletti) körében a nagyobb és sajnos ez a tendencia évről évre nő (1).

### 2.3.3.1 Biomarkerek az akut mieloid leukémiában

Az akut mieloid leukémia patogenezisében fontos szerepet játszanak a kromoszomális strukturális eltérések és a molekuláris genetikai eltérések. A „The Cancer Genome Atlas” tanulmány alapján az akut mieloid leukémiában az öt leggyakrabban mutált gén az NPM1, IDH1, IDH2, DNMT3A és az FLT3 gének (118).

Az izocitrát-dehidrogenáz 1/2 (IDH1/2) gént érintő mutációk, a betegek 15%-ban fordulnak elő és az életkor emelkedésével egyre jobban nő a gyakoriságuk. Ezekben a génekben található mutációk a DNS és hiszton hipermetilációjával, a megváltozott génextpresszióval és haematopoeticus progenitor sejtek differenciációjának megállításával vannak kapcsolatban (119).

Az FMS-like tyrosine kinase 3 (FLT3) gén egy 3-as típusú tirozin kináz aktivitással rendelkező receptort kódol, amelynek szerepe van a hematopoézis szabályozásában beleértve az őssejtek differenciációját és az osztódásukat. Az FLT3 gén mutációi rossz túléléssel társulnak a fiatal páciensek körében. Az FLT3 gén tirozin kináz doménjét érintő aktiváló mutációk, az akut mieloid leukémiás betegek 15%-ban találhatóak meg (120).

### 2.3.3.2 Az NPM1 mutációk

A nucleophosmin gén (NPM1) az egyik leggyakrabban mutált gén az akut mieloid leukémiában (121). Az NPM1 génnek szerepe van a riboszómák képződésében és a transzportjában, a nukleoluszban lévő p14Arf onkoszuppresszor fehérje stabilizálásában, továbbá a centroszóma megkettőződés szabályozásában. Az NPM1 mutációk az akut mieloid leukémiás betegek 20-30%-ban találhatóak meg, hatásukra megváltozik a sejt citoplazmájában a fehérjék kihorgonyozása (rögzítése), amely a leukemogenezis egyik legfontosabb történése (122).

Korábbi tanulmányok igazolták, hogy az NPM1 mutációk főként a mieloid sejtekre jellemzőek, mivel a fehérjék aberráns citoplazmatikus lokalizációja, a csontvelőből vagy a perifériális vérből származó limfoid sejtek esetében nem figyelhető meg (123). Az NPM1 szomatikus mutációk gyakran együtt fordulnak elő az FLT3 génben található „internal tandem duplikációkkal” (ITD) és a DNMT3A mutációkkal. Az NPM1 mutációk az FLT3-ITD és a DNMT3A mutációk mellett gyakran az IDH1/2 és a TET2 mutációkkal is mutatnak asszociációt (124). Az NPM1 mutációkkal ritkán előforduló

mutációk a RUNX1, CEBPA és a TP53 génekben található szomatikus mutációk, valamint az MLL génben található duplikáció. Az FLT3 tirozin kináz domént érintő mutációkkal az NPM1 mutációk ritkán fordulnak elő (124). Egy korábbi tanulmány mutatja, hogy NPM1 mutációval rendelkező normális kariotípusú akut mieloid leukémiás betegeknek jobb a túlélése (125). Egy másik vizsgálat pedig azt mutatja, hogy a kariotípus, életkor, NPM1 mutációs státusz, fehérvérsejt szám, laktát dehidrogenáz és a CD34 kifejeződés egymástól független prognosztikus markerei a teljes túlélésnek (126). Szintén korábbi vizsgálatok mutatják, hogy az IDH1 mutációk jobb túléléssel társulnak NPM1 mutáns és FLT3-ITD vad típusú betegek esetén (127).

Jelenleg a kemoterápia az elsődleges terápiás lehetőség a fiatal és megfelelő egészségi állapottal rendelkező akut mieloid leukémiás páciensek esetén, amely általában egy antraciklin (mint például idarubicin) és cytarabin kombinációjából áll (128), (129). Az akut mieloid leukémia célzott terápiája során számos lehetséges terápiás készítmény áll rendelkezésre, azonban ezek nagy része még klinikai vizsgálat alatt áll. Ilyenek például a tirozin kináz gátlók – sorafenib (130), midostaurin (131), quizartinib (132) és crenolanib (133) –, amelyek az FLT3 receptor tirozin kináz egységét gátolják. Az akut mieloid leukémiás betegek 50%-ban a STAT3 fehérje felülszabályozódik és ez rossz túléléssel társul. Így a betegek terápiája lehetséges a STAT3 inhibitorokkal is (C188-9, OPB-31121), amelyek specifikusan gátolják a STAT3 fehérje foszforilációját (134), (135). Számos további lehetséges célzott terápiás készítmények, úgymint IDH1 (136) és IDH2 (137) gátlószerek, sejtmag transzport gátlók és CD33 (138), valamint CD123 antigén specifikus gátlószerek (139) állnak rendelkezésre az akut mieloid leukémia kezelése során.

## **2.4 A nukleinsav szekvenálás**

A biológiai minták nukleinsav szekvenálása számos tudományterületen elterjedt, mára már szinte alapvető molekuláris biológiai módszer. A DNS szekvenálás az elmúlt évtizedekben rendkívül nagy változásokon ment keresztül, kiindulva a rövid oligonukleotid szakaszok szekvenálásától egészen a teljes genom szekvenálásáig.

A kezdetekben a Maxam és Gilbert által kifejlesztett a DNS radioaktív jelölésén és különböző kémiai módosításán alapuló szekvenálási módszer nagy áttörést jelentett a DNS bázissorrendjének megállapításában (140). Azonban sajnos ez a módszer rendkívül

idő-, és munkaigényes volt, így egy ettől sokkal hatékonyabb módszer kifejlesztésére volt szükség (140). Az 1970-es évek közepére Fredrick Sanger, egy sokkal gyorsabb és hatékonyabb egy úgynevezett „lánc-termináció” alapuló DNS szekvenálási technikát fejlesztett ki (141). A következő néhány évtizedben ez a módszer tovább fejlődött, automatizálódott és így drámai módon felgyorsult a nukleinsav szekvenálásának a folyamata.

Az első generációs szekvenálási módszereket az úgynevezett újgenerációs szekvenálási módszerek követték. Ezek közé tartozik az egyik legkorábbi újgenerációs módszer, a piroszekvenálás. Ez egy szintézis alapú módszer, ahol a DNS szintézisét végző DNS polimeráz enzim aktivitását a luciferáz enzim fluoreszcens fénykibocsátásával detektálnak (142). A piroszekvenálás egyik legfontosabb előnye a Sanger szekvenálással szemben, hogy egyetlen szekvenálási reakcióban akár 20 millió nukleotid is leolvasható, szemben a Sanger szekvenálás kevesebb, mint 1000 nukleotid hosszúságú hatékonyságával (142). A piroszekvenálást számos újgenerációs szekvenálási módszer kifejlesztése követett, amely során lehetővé vált szekvenálási reakciók párhuzamosítása és így akár egyetlen szekvenálási reakcióban a teljes genom bázissorrendje is megismerhető.

A következő alfejezetekben részletesen ismertetem a Sanger szekvenálás és az újgenerációs szekvenálás elméleti háttérét, valamint ezek klinikumban betöltött szerepét.

#### **2.4.1 Sanger szekvenálás**

Az egyik legjobban elterjedt első generációs DNS szekvenálási eljárás a Sanger féle DNS láncterminációs szekvenálás, amely során a reakcióban a DNS polimerázok és kémiai módon módosított úgynevezett didezoxinukleotidok segítségével különböző hosszúságú DNS fragmensek keletkeznek (141). A didezoxinukleotidok egyik fontos tulajdonsága az, hogy a dezoxiribóz harmadik szénatomján nem található hidroxil csoport, ezért nem képes hozzákapcsolódni a következő nukleotid ötödik szénatomjának foszfátjához. Tehát a folyamat során a didezoxinukleotidok megjelenése a szintetizálódó DNS láncban terminálja a további polimerizációt, az új lánc nem nő tovább és különböző hosszúságú DNS szakaszok keletkeznek (141).

A szekvenálási folyamat során négy párhuzamos szekvenálási reakció zajlik. Minden reakció tartalmazza az egyszálú templát vagy minta DNS-t, a szekvenciára

specifikus, kisméretű nukleotid (úgynevezett primer) szekvenciákat, a DNS polimeráz enzimet, az adenin, guanin, citozin és timin tartalmú dezoxinukleotidokat (dATP, dGTP, dCTP és dTTP), valamint reakcióelegyenként a radioaktív izotóppal jelölt didezoxinukleotidokat (ddATP, ddGTP, ddCTP, és ddTTP). A didezoxinukleotidok koncentrációja a normál dezoxinukleotidok koncentrációjához képest sokkal kisebb, így a láncterminálódás gyakorisága nagyon kicsi és véletlenszerű. A szekvenálás során keletkeznek olyan nagyon rövid DNS szálak, amelyekbe már korán beépült a didezoxinukleotid és keletkeznek olyanok is, amelyekbe sokáig egy darab didezoxinukleotid sem épült be, ezért ezek hosszúak (141).

A szekvenálási reakciók végén a különböző hosszúságú DNS szakaszokat poliakrilamid gélelektroforézissel választják szét, így mind a négy reakció esetén különböző hosszúságú, az adott nukleotidra specifikusan végződő DNS szakaszok keletkeznek. A DNS szekvencia a gél aljától, azaz a rövid szakaszoktól visszafelé azaz a hosszabb szakaszok felé haladva olvasható le (141).

A Sanger féle szekvenálási módszer számos fejlesztésen esett át, így például a didezoxinukleotidok radioaktív jelölése helyett áttértek a fluoreszcens jelölésre, amely során a DNS szakaszok méret szerinti elválasztása egyetlen kapillárisban végezhető el (143). A módszer az egyszerűsödése nagymértékben hozzájárult az automatizált DNS szekvenálás létrejöttéhez, amely során lehetővé vált az egyre összetettebb fajok genomjának a vizsgálata.

A Sanger szekvenálás egy széles körben alkalmazott módszer, amely során lehetőség van a pontmutációk vagy kisméretű deléciók, duplikációk azonosítására (144). Jelentős szerepe van a daganatok mutációs spektrumának meghatározásában, és jelenleg a daganatdiagnosztika oszlopos tagja (145). A módszer egyik legfontosabb limitációja, hogy legfeljebb 1000 bázispár hosszúságú szakaszt olvashatunk le, és amennyiben több tíz kilobázis hosszúságú régiót szeretnénk meghatározni, akkor érdekesebb valamilyen költséghatékonyabb, nagyobb áteresztő képességű szekvenálási eljárást választani.

#### **2.4.2 Az újgenerációs szekvenálás**

Mint ahogy már korábban is említettem, a legelső újgenerációs szekvenálási módszer a piroszekvenálás volt, amelyet a Roche cég fejlesztett ki és a 454 szekvenálási technológia néven terjedt el (142). A piroszekvenálást számos újgenerációs szekvenálási

eljárás követett, ilyenek például a Solexa módszer, amelyet az Illumina, vagy a SOLiD, amelyet az Applied Biosystems fejlesztett ki (146).

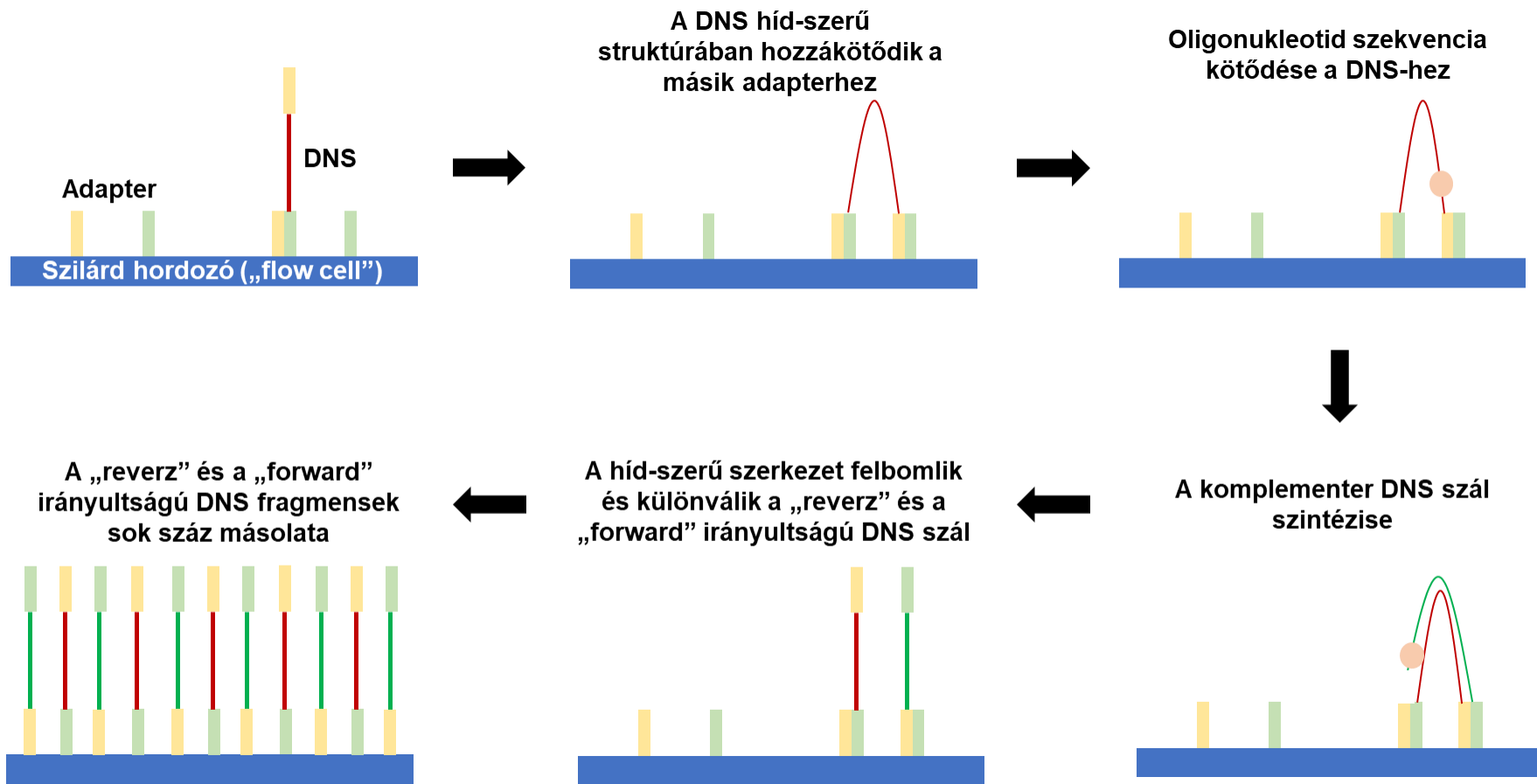
Az újgenerációs szekvenálási eljárások módszertanilag hasonlóak. Mindegyik esetben a szekvenálási könyvtár készítése a DNS véletlenszerű fragmentálásával kezdődik, majd ezt követően a fragmenseket úgynevezett adapter szekvenciák közé illesztik (146). Majd a fragmenseket különböző PCR alapú eljárások során felszorzósítanak, ez a Solexa eljárás esetén a híd PCR, valamint a SOLiD eljárás esetén az emulziós PCR (147), (148). Maga a szekvenálás folyamata során enzimek – többek között DNS polimeráz – által, fluoreszcens molekulával kapcsolt dezoxinukleotidok épülnek be az újonnan keletkező DNS szálba, ahol a keletkező fluoreszcens jelet detektálják (146).

Az újgenerációs szekvenálásnak a hagyományos Sanger szekvenálással szemben számos előnye van. Elsősorban az, hogy lehetőség van arra, hogy egyszerre nagyszámú szekvenálási reakció is elvégezhető, így akár a teljes genom egyetlen szekvenálási folyamat során leolvasható. Egy másik fontos előnye Sanger módszerrel szemben a költséghatékonyság, mivel egyetlen szekvenálás során csupán néhány mikroliter mennyiségű reagens szükséges, így akár egy szervezet teljes DNS vagy RNS tartalma is leolvasható (146).

Az egyik legelterjedtebb újgenerációs szekvenálási eljárás az Illumina által kifejlesztett Solexa. A szekvenálás, egy szilárd hordozón az úgynevezett „flow cell” felületén zajlik (149). Első lépésében, a már *in vitro* körülmények között létrehozott szekvenálási könyvtár felszorzósítása történik híd PCR reakcióval (**2. ábra**). A PCR reakció során minden egyes, a DNS könyvtárból származó fragmensből nagyjából 1000 darab másolatot kapunk (147). Miután megtörtént a fragmensek felszorzósítása, a dupla szálú fragmensek denaturálását követően a primerek bekötődése következik a fragmensek végeihez. A szekvenálási reakció számos cikluson keresztül zajlik le és minden szekvenálási ciklusban a DNS polimeráz, a primer 3' hidroxil csoportjától kezdve beépít egy nukleotidot az újonnan szintetizálódó DNS szálba. Az új szálba beépülő nukleotidokhoz négy fajta fluoreszcens módosítást tartalmaznak, amely során minden egyes beépülő nukleotid esetén fluoreszcens jel keletkezik, amelyet a készülékben lévő kamerák érzékelnek (150).

Az újgenerációs szekvenálás számos területen alkalmazható, úgymint a teljes genom szekvenálás, a *de novo* szekvenálás, a teljes exom, valamint a teljes RNS szekvenálás. A teljes RNS szekvenálás során lehetőségünk van eltérő kifejeződést mutató gének azonosítására különböző minták között. Továbbá, alkalmazható még kromatin immunprecipitációs szekvenálására, valamint eltérő kifejeződést mutató mikro-RNS-ek vagy egyéb nem kódoló kis RNS-ek azonosítására is (146).





2. ábra. A híd-PCR reakció folyamata az Illumina Solexa szekvenálási módszer során.

## 2.5 A génextpresszió mérésére alkalmas módszerek

A funkcionális genomikai kutatásokhoz, többek között egyes gének funkciójának, működésének és a szabályozódásának a megértéséhez, elengedhetetlen génkifejeződés vizsgálata.

A génkifejeződés mérését számos területen alkalmazzák. Alkalmazható különböző genotípusok, valamint szövettípusok közötti eltérő kifejeződést mutató gének azonosítására, valamilyen kezelés hatásának vizsgálatára, vagy az egyes fejlődési szakaszban bekövetkező génextpressziós változások detektálására is (151). A génkifejeződés kvantitálására mára már számos módszer áll rendelkezésre, ezek közül a legelterjedtebbek a valós idejű polimeráz láncreakció (qPCR), a microarray és a teljes RNS szekvenálás (152), (153), (154).

### 2.5.1 A kvantitatív polimeráz láncreakció

A polimeráz láncreakció (PCR) módszer kifejlesztése Kary Mullis nevéhez fűződik, amely során az egyes DNS szakaszok, a DNS polimeráz enzim segítségével felszorzósíthatóak (155). A kvantitatív polimeráz láncreakció (qPCR) egy olyan PCR alapú eljárás, ahol valós időben detektálható egy adott génről átírt transzkript mennyisége. Mára a qPCR, a számos génextpresszió mérésére alkalmas módszerek közül az egyik „standard” módszernek számít, főként a nagy vizsgálati specificitása és a nagy detektálási érzékenysége miatt (152).

A qPCR reakció során az átírt mRNS mennyisége többféle módon mérhető. Mindegyik módszer esetén a kiindulási minta a komplementer DNS (cDNS). A cDNS, a sejtből vagy a szöveti mintából izolált mRNS, reverz transzkriptáz enzimmel, dupla szálú DNS-sé visszaírt változata (152).

Az egyik legelterjedtebb génextpressziós kvantitálásra alkalmas qPCR alapú módszer a SYBR Green I eljárás (156). Ezen eljárás népszerűsége elsősorban a módszer egyszerűségének és az alacsony költségének volt köszönhető. A SYBR Green I egy fluoreszcens festék, amely a kettős szálú DNS kis árkában kötődik be, és a kötődést követően fluoreszcens jelet bocsájt ki. Ennek során, a PCR reakció ciklusszámának a növekedésével, egyre több SYBR Green I festékmolekulával jelölt, dupla szálú PCR

termék keletkezik és a kibocsátott fluorszcens jel erőssége arányos a PCR termék mennyiségének növekedésével (156).

A valós idejű qPCR eljárás egy másik módszere esetén, fluoreszcens molekulával jelölt, hidrolízis próbákat alkalmaznak. Ezeket, a vizsgálni kívánt génszakaszra specifikus oligonukleotid szakaszokat, TaqMan próbáknak hívják (157). A próba két fajta fluorofórral van megjelölve, az egyik az úgynevezett kioltó, a másik pedig a jeladó molekula. A próba 5' vége felé a kioltó, a 3' vége irányába pedig a jeladó molekula helyezkedik el. Az amplifikáció kezdetén a kioltó molekula folyamatosan elnyeli a jeladó molekula által kibocsátott fluoreszcens jelet. A DNS polimeráz a PCR reakció során az új DNS szál szintézisét végzi és amikor eléri a próba molekulát, az 5' exonukleáz aktivitása révén elbontja a próbát, így a jeladó molekula által kibocsátott fluoreszcens jel detektálhatóvá válik (157). Fontos kiemelni, hogy a hidrolízis próbák alkalmazása a génexpresszió mérése során sokkal nagyobb specificitást eredményeznek a SYBR Green I eljáráshoz képest, mivel a próbák szekvenciája specifikus, a vizsgálni kívánt gén szekvenciájára.

A qPCR eljárás egyik legfontosabb előnye az, hogy a reakció a mintaelőkészítéssel együttvéve meglehetősen gyorsan kivitelezhető, olcsó, és magas szenzitivitással és specificitással működik. Azonban az egyik, és egyben a legnagyobb hátránya ennek a módszernek az, hogy ha több gén kifejeződésének a mérése a cél, az már sok időt és energiát igényel, valamint a költségvetés is jelentősen megemelkedhet.

### **2.5.2 A microarray**

A génexpressziós microarray segítségével lehetőség van az összes gén, egy adott időpillanatban történő transzkripció aktivitásának a vizsgálatára (153). A microarray számos területen alkalmazható, úgymint különböző betegség altípusok azonosításában, gyógyszerhatás vizsgálatában és különböző diagnosztikában használt eszközöknek a fejlesztése során (158). Azonban, fontos megemlíteni, hogy a microarray vizsgálatok feldolgozása és kiértékelése, szemben a qPCR adatokkal, komoly technikai kihívást jelent. Egy részről a kihívást a nagy mennyiségű adat és a nagy számításigény jelenti, másik részről pedig nagy körültekintést igényel az adatok feldolgozására használt programok, szoftverek helyes kiválasztása (158).

Számos génexpressziós microarray módszert fejlesztettek ki az utóbbi időben, ilyenek például az Affymetrix, Agilent és a NimbleGen cégek által kidolgozott módszerek (153). A microarray, egy DNS hibridizációs technika, amely egy üvegfelületen zajlik. Ezen az üvegfelületen több ezer pont helyezkedik el, és minden egyes pont egy adott génre specifikus oligonukleotid vagy genomiális DNS szekvenciát tartalmaz, sok millió kópiában. A microarray egyik leggyakoribb alkalmazása a két minta csoport (például kontroll és kezelt) közötti génkifejeződés eltéréseinek vizsgálata (153).

A módszer első lépésében a kontroll és a kezelt mintákból RNS-t izolálnak, amelyet reverz transzkripció során cDNS-é fordítanak át. Ezt követi a minta csoportok megjelölése, amely két színű (piros és zöld) fluoreszcens festékekkel történik. A következő lépés a hibridizáció, amely során a különböző fluoreszcens festékekkel jelölt minták cDNS molekulái hozzákötődnek az üvegfelületen kihorgonyozott megfelelő komplementer szekvenciához (153). A microarray-en az egyes pontokhoz kötődött cDNS mennyisége arányos, az adott mintában lévő, adott génhez tartozó RNS molekulák számával. A minták DNS-éhez kötődött piros és a zöld festékmolekulák, lézer sugárral történő gerjesztését követően történik az emittált jel detektálása. A fluoreszcens jel intenzitása arányos a próbához kötött DNS molekulák számával. A detektálás során egy adott pont esetén négy féle szín lehetséges: a piros jel esetén az egyik mintában, míg a zöld jel esetén a másik mintában nagyobb az adott gén kifejeződése. Amennyiben mind a két mintában egyenlő mértékben fejeződik ki az adott gén, akkor sárga jelet kapunk, valamint ha egyik mintában sem történik kifejeződés, akkor pedig a fekete szín jelenik meg (153).

Manapság a microarray technikát egyre inkább kezdi felváltani az újgenerációs RNS szekvenálás, amely segítségével az egyes RNS átiratok szekvenciája is megismerhető, azonban ennek költsége jelenleg meglehetősen magas (158). Továbbá fontos megemlíteni, hogy korábbi vizsgálatok bizonyítják, hogy a két génexpresszió mérésére alkalmas módszer között nagy a hasonlóság (159). Így mindaddig, amíg a szekvenálás alapú módszerek költséghatékonysága és könnyen alkalmazhatósága nem javul, a mikroarray sok orvos számára kívánatos alternatíva marad.

### **2.5.3 Az RNS szekvenálás**

Az újgenerációs DNS szekvenálás elterjedése teret adott a teljes RNS szekvenálás módszerének létrejöttéhez. Az RNS szekvenálás kezdeti szakaszában cDNS könyvtárakat

készítettek és ezek nukleotid szekvencia sorrendjét a hagyományos Sanger szekvenálással olvasták le (160). Ezzel a technikai megközelítéssel csak néhány gén volt vizsgálható, és a génexpresszió mértékéről nem adott információt. További probléma volt az, hogy csak az egyes transzkriptek voltak vizsgálhatóak ezzel a módszerrel, és az egyes transzkript izoformákat nem lehetett megkülönböztetni egymástól (160).

Az RNS szekvenálás, az újgenerációs szekvenálási módszerek közé tartozik, amely során lehetővé válik egy adott szervezet teljes mRNS, vagy nem kódoló RNS (például mikro-RNS) tartalmának a vizsgálata. Továbbá, ezzel a módszerrel lehetőség van az alternatív splicing folyamatának vizsgálatára, egy nukleotidos polimorfizmusok (SNP) és fúziós gének vizsgálatára, valamint az úgynevezett egysejtes („single-cell”) RNS szekvenálásra is (161).

Az RNS szekvenálás hasonlóan a többi újgenerációs szekvenálási eljáráshoz, a könyvtárkészítésen alapszik, amely során a cDNS-sé átírt teljes RNS mennyiséget fragmentálják és adapter szekvenciák közé illesztik. Majd ezt követi a fragmens szekvenciák szilárd felszínen történő felsokszorosítása és a szintézis alapú leolvasása (154). A szekvenálás részletei a 2.4.2 alfejezetben olvasható.

Az RNS szekvenálásnak számtalan előnye és természetesen korlátai is vannak. A módszer előnyeik közül fontos kiemelni, hogy nemcsak szekvenciainformációt kapunk a minta teljes RNS tartalmáról, hanem a transzkriptek mennyiségét is megtudjuk egy adott időpillanatban, így ezzel a módszerrel akár különböző tényezők génexpresszióra gyakorolt hatása is vizsgálhatóvá válik (161). Továbbá, az RNS szekvenálással sokkal pontosabban határozható meg egy gén kifejeződési szintje, mint egy microarray eljárással (161). A génkifejeződés mértéke, a referencia genomhoz történő illesztés során kapható meg, ahol egy adott gén kifejeződési szintje, az exonokhoz tartozó leolvasások számával egyenlő. Ez az érték az úgynevezett „HTSeq-counts” érték, amely kiszámításhoz a leggyakrabban használt program a HTSeq program (162).

A módszer egyik limitációja az, hogy az adatok feldolgozása rendkívül számítás igényes és megfelelő bioinformatikai tudást igényel. Továbbá nélkülözhetetlen, hogy a vizsgált minta esetén legyen ismert és feltérképezett referencia genom szekvencia (163). Szintén a módszer korlátai közé sorolható az, hogy a génexpressziós értékeket nagymértékben befolyásolhatja a könyvtárkészítés, valamint a szekvencia illesztés során bekövetkező hibák (163). Végül, fontos kiemelni a reprodukálhatóságot, mivel az egyes

szekvenálási platformok nincsenek standardizálva, így például ugyanazon mintát különböző RNS szekvenálási módszerrel lemérve, eltérő szekvenálási eredményeket kaphatunk, amelyek főképpen a szekvenálás mélységében nyilvánulhatnak meg (163).

Az RNS szekvenálás mára az egyik legelterjedtebb globális génexpresszió mérésére alkalmas eljárás, mivel magas a szenzitivitása, széles tartományban képes detektálni a génekifejeződést és a transzkriptvariánsok vizsgálatára is alkalmas. Egyre inkább nő az alkalmazása például az onkológiai biomarkerek azonosításban, amely lehetővé teszi a daganatok pontosabb altípusba sorolását, valamint segítenek a terápiás válasz előrejelzésében (163).

### 3 CÉLKITŰZÉSEK

- 3.1. Irodalmi adatok alapján szomatikus KRAS mutációk és a KRAS génexpresszió nem mutatnak szignifikáns összefüggést a túléléssel a nem-kissejtes tüdőrákban. A munkám során célom volt a szomatikus KRAS mutációkkal összefüggő transzkriptomikus mintázat prognosztikus hatásának vizsgálata nem-kissejtes tüdőrákban.
  - 3.1.1. Ehhez az elsődleges célom volt egy mutációs és génkifejeződési adatokkal kombinált adatbázis létrehozása nem-kissejtes tüdőrákban..
  - 3.1.2. További célom volt a szomatikus KRAS mutációkkal és a KRAS génben lévő kópiaszám eltérésekkel összefüggő génkifejeződési változások azonosítása,
  - 3.1.3. és ezeknek a génexpressziós változásoknak a túléléssel való összefüggésének a vizsgálata.
- 3.2. Az NPM1 mutációk kiemelkedő prognosztikus markerek az akut mieloid leukémiában, valamint az optimális kezelési stratégia kiválasztásában is jelentős szerepük van. Célom volt megvizsgálni, az NPM1 mutációk génexpresszióra és a betegek túlélésre gyakorolt hatását akut mieloid leukémiában.
  - 3.2.1. Az első lépésben célom az adatbázisépítés volt, amely során különböző akut mieloid leukémiás betegcsoportok génkifejeződési, mutációs és klinikai adatát gyűjtöttem össze.
  - 3.2.2. A második célom az volt, hogy az NPM1 szomatikus mutációkkal összefüggő génkifejeződés változásokat azonosítsak, és vizsgáljam a hatásukat a túlélésre különböző betegcsoportokon.
  - 3.2.3. Végül, klinikai minták bevonásával, *in vitro* vizsgálat során, szeretném igazolni a kiválasztott gének kifejeződésbeli változásait, az NPM1 mutáns és vad típusú betegek között.
- 3.3. Számos korábban publikált mRNS és mikro-RNS alapú prognosztikus biomarker létezik máj hepatocelluláris karcinóma esetén, azonban ezek prognosztikus hatása sok esetben nem egyértelmű. A kutatásom során további célom volt, hogy ezeknek az mRNS és mikro-RNS alapú biomarkereknek a túléléssel való összefüggését vizsgáljam meg a máj hepatocelluláris karcinómában.

- 3.3.1. Elsődleges célunk egy irodalmi keresés, ahol összegyűjtjük a korábban azonosított mRNS és mikro-RNS biomarkereket hepatocelluláris karcinómában.
  - 3.3.2. Ezt követően célunk nyilvánosan elérhető mRNS és mikro-RNS szekvenálási és microarray génkifejeződési adatok felhasználásával egy integrált adatbázist készíteni.
  - 3.3.3. Végül Cox regresszió során igazolni a korábban publikált biomarkerek génexpressziójának a hatását a hepatocelluláris karcinóma túlélésére.
- 3.4. A daganatok kialakulásában számos gén játszik szerepet, ezeket korábbi tanulmányok funkcionális csoportokba, úgynevezett daganat jellemzőkbe sorolták. A célom az volt, hogy megvizsgáljam a daganat asszociált gének prognosztikus hatását, különböző tumortípusokban.
- 3.4.1. Az első lépésben célom volt a TCGA (The Cancer Genome Atlas) adatbázisban elérhető össze tumortípus RNS szekvenálásai adatának összegyűjtése, adatbázisba integrálása.
  - 3.4.2. A második lépés során célom volt a daganattényezőkkel összefüggő gének kifejeződése és a túlélés közötti kapcsolat vizsgálata, valamint meghatározni a túléléssel összefüggő gének megoszlását a különböző daganattípusok között.
  - 3.4.3. A harmadik lépésben célom volt a fő daganatjellemzők által meghatározott transzkriptomikus mintázatok prognosztikus hatásának vizsgálata, különböző daganattípusok között.



## 4 MÓDSZEREK

### 4.1 Mutációs és génexpressziós adatok feldolgozása

Az adatok feldolgozását és a statisztikai számításokat R v.3.2.3 statisztikai programkörnyezetben végeztem (<https://www.r-project.org/>). A statisztikai elemzések során a szignifikancia küszöb  $P < 0,05$  volt.

#### 4.1.1 mRNS expressziós adatok

Az mRNS expressziós adatokat két ingyenesen elérhető adatbázisból töltöttem le: a teljes RNS szekvenálási adatok a Genomic Data Commons (GDC, <https://portal.gdc.cancer.gov/>) adatbázisból, és azon belül is a The Cancer Genome Atlas (TCGA) genomikai együttműködési programból származtak. Az mRNS kifejeződési adatoknak egy másik típusa, a microarray génexpressziós adatok, a National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO, <https://www.ncbi.nlm.nih.gov/geo/>) adatbázisból voltak elérhetőek.

##### 4.1.1.1 RNS szekvenálás alapú génexpressziós adatok feldolgozása

A TCGA programból származó, Illumina HiSeq 2000 RNS szekvenálási módszeren alapuló nyers RNS szekvenálási (htseq counts) adatokat, a betegekhez tartozó klinikai adatokkal együtt töltöttem le. A TCGA program 33 különböző tumortípus RNS szekvenálási adatát tartalmazza, azonban csak azon 28 tumortípus adata került letöltésre és feldolgozásra, amelyek esetében legalább 100 minta volt elérhető.

A nyers adatok normalizálását DESeq algoritmussal végeztem, amely a negatív binomiális eloszláson alapul (164). Ezen program, a különböző RNS szekvenálási adatok normalizálására alkalmas algoritmussal szemben, képes a megfelelő hamis pozitív arány megfelelő szinten tartására. Az Ensembl transzkript azonosítók és a HUGO gén azonosítók (<https://www.genenames.org/>) összekapcsolását az „AnnotationDbi” R programcsomag (<http://bioconductor.org/packages/AnnotationDbi/>) segítségével végeztem. Egy további, úgynevezett skálázó normalizálást is alkalmaztam az adatok zajmentesítésére, amely során a gének átlagos expressziós értékét minden mintában 1000-re állítottam (165).

#### 4.1.1.2 Microarray alapú génexpressziós adatok feldolgozása

A microarray génexpressziós adatokat a nem-kissejtes tüdőrák, akut mieloid leukémia és a máj hepatocelluláris karcinóma esetén használtam fel.

A microarray adatok feldolgozása esetén első lépésben elvégeztük a minták (génchipek) minőség ellenőrzését a „yaqcaffy” R Bioconductor programcsomag (<http://bioconductor.org/packages/yaqcaffy/>) segítségével. Ezt követően, csak azon minták kerültek kiválasztásra, amelyek megfelelnek az egyes minőségi paramétereknek: a génchip nyers minőségi értéke, oligonukleotid próba hibridizációk százalékos aránya, a BioB-/C-/D-csúcsok jelenléte, a GAPDH és béta-aktin 3'-5' arányok (166). A génexpressziós adatok normalizálását MAS5 algoritmussal végeztük, amely az „affy” R Bioconductor (<http://bioconductor.org/packages/affy/>) programcsomag része. A háttérzaj csökkentésére továbbá egy második, már az RNS szekvenálási adatok feldolgozásánál ismertetett úgynevezett skálázó normalizálást is alkalmaztunk (165). Egy génhez több próba is tartozhat, és a legmegbízhatóbb próba kiválasztására a JetSet adatbázist alkalmaztuk (167).

A következőkben, az egyes daganattípusok esetén ismertetem, hogy konkrétan mely microarray tanulmányokat választottuk ki, és a fentebb említett általános adatfeldolgozási lépéseken kívül milyen egyéb speciális módszereket alkalmaztunk. Fontos kiemelni, hogy mindegyik daganattípus esetén azok a génexpressziós vizsgálatok kerültek kiválasztásra, amelyek mind nyers génexpressziós és mind klinikai adatokat is tartalmaztak.

Az NCBI GEO adatbázisban elérhető nem-kissejtes tüdőrák génexpressziós adatok az Affymetrix HG-U133A 2.0 (GPL571), HG-U133 Plus 2.0 (GPL570) és HG-U133A (GPL96) microarray módszereken alapulnak. A teljes, klinikai adatokkal is ellátott adatbázis létrehozása és a minőségellenőrzése, a kutatócsoport egy korábbi munkája során készült el (166). A nem-kissejtes tüdőrákkal kapcsolatos elemzések során fontos volt, hogy a microarray és az RNS szekvenálási adatok egymással összehasonlíthatók legyenek. Így, a normalizálási lépések előtt, azon gének kerültek kiválasztásra ( $n = 11500$  gén), amelyek mind a két módszer során mértek. Továbbá, a háttérzaj csökkentése érdekében azon géneket kiszűrtük, amelyek esetében a betegekben mért átlagos génexpresszió (a normalizálást követően) 100 alatti volt.

Az akut mieloid leukémiás betegek esetén a GSE6891 (168) és a GSE1159 (169) vizsgálatok microarray génexpressziós adatait használtam fel. Az első tanulmány mintái az Affymetrix HG-U133 Plus 2.0 (GPL570), míg a második tanulmány mintái az Affymetrix HG-U133A (GPL96) típusú microarray eljárásból származtak. A vizsgált génexpressziós tanulmányok keresése és megtalálása az NCBI GEO adatbázisban a következő keresőszavakkal történt: „AML,” „GPL570” és „GPL96”.

A máj hepatocelluláris karcinóma microarray adatokat szintén az NCBI GEO adatbázisból származtak és a „hepatocellular”, „carcinoma”, „vascular” és „invasion” keresőszavak használatával azonosítottuk és töltöttük le a vizsgálatokat. A statisztikai tesztek elvégzéséhez szükséges megfelelő mintaszám figyelembevételével csak azon vizsgálatok génexpressziós adataival foglalkoztunk a továbbiakban, amelyek legalább 50 betegmintát tartalmaztak. Ezek alapján az első vizsgálat a GSE20017 (103), amely esetében a génexpressziót az Illumina HumanRef-8 WG-DASL v.3.0 micorarray módszerrel határozták meg. Az azonosított második vizsgálat a GSE9843 (170), ahol a minták génexpresszióját Affymetrix HG-U133 Plus 2.0 (GPL570) eljárással mérték.

#### **4.1.2 Mikro-RNS expressziós adatok**

Az mRNS expressziós adatokhoz hasonlóan, a mikro-RNS expressziós adatok is a GDC-TCGA és az NCBI GEO adatbázisokból származtak. Klinikai adatokkal összekapcsolt mikro-RNS expressziós adatbázist a máj hepatocelluláris karcinóma esetén hoztunk létre.

##### **4.1.2.1 Microarray alapú mikro-RNS expressziós adatok feldolgoása**

A máj hepatocelluláris karcinóma esetén a „hepatocellular”, „carcinoma”, „miRNA” és „survival” kulcsszavakat alkalmaztam a mikro-RNS alapú expressziós vizsgálatok keresésére. Azon vizsgálatokat választottam ki – GSE31384 (100), GSE10694 (171) és GSE6857 (172) –, amelyek legalább 60 mintát tartalmaztak mikro-RNS expressziós és túlélési adattal. Egy másik nagyon fontos szempont volt az adatok kiválasztásánál, hogy az ugyanazon elsődleges tumor mintához tartozzon normál (nem tumoros) minta expressziós adat, a minták mikro-RNS expresszióinak páronkénti összehasonlítás céljából. Az expressziós adatokat újrnormalizálás nélkül használtuk fel további elemzések céljából.

A GSE31384 és a GSE10694-es microarray vizsgálatok során a mikro-RNS kifejeződési adatok normalizálására kvantilis, míg a GSE6857-es expressziós adatok esetén medián normalizációt alkalmaztak. Fontos kiemelni, hogy a GSE31384-es mikro-RNS expressziós vizsgálat esetén a tumoros és a hozzá tartozó normál mintapárból származó jelölt mikro-RNS-eket, ugyanazon microarrayen mérték. Ezen adatok esetén a mikro-RNS expressziós értékek, a tumoros és a normál szövetből származó normalizált  $\log_2$  arányt jelentik.

Azonban, nem csak az egyes vizsgálatok és a minták szintjén szűrtünk, hanem a mikro-RNS-ek alapján is. Elsődleges szempont volt a humán mikro-RNS-ek kiválasztása, az ismétlődő mikro-RNS-ek eltávolítása, valamint azon mikro-RNS-ek kiválasztása, amelyek a minták legalább 75%-ban fordulnak elő. Továbbá az egyes, már érett mikro-RNS-ek esetében átlagos expressziót számoltam (például az átlagos expresszióját vettem a hsa-miR-101-1 és a hsa-miR-101-2 mikro-RNS-eknek). A GSE10694 és a GSE31384-es vizsgálatok esetében az expressziós értékeket az eredeti logaritmikus skáláról lineáris skálára alakítottam az egységes expressziós skálázás érdekében.

#### **4.1.2.2 RNS szekvenálás alapú mikro-RNS expressziós adatok**

Az Illumina HiSeq 2000 eljárás alapján alapuló máj hepatocelluláris karcinóma mikro-RNS expressziós adatok esetében a már „reads per million mirna mapped” (RPM) eljárással normalizált mikro-RNS expressziós adatokat használtam fel. Kiszűrésre kerültek azon minták, amelyek esetében nem volt párosított normál minta ( $n = 1$ ), valamint ismétlődést mutattak ( $n = 3$ ). Minden további adatszűrési lépés megegyezett a microarray alapú mikro-RNS expressziós adatok szűrési lépéseivel.

#### **4.1.3 Szomatikus mutációk azonosítása**

A nem-kissejtes tüdőrákos betegek tumoros és a hozzá tartozó normál mintájának teljes exom szekvenálási adatát a GDC TCGA adatbázisból töltöttük le. A szomatikus mutációk azonosítására a MuTect algoritmust alkalmaztuk, az alapértelmezett beállítási paraméterek mellett (173), (174). Az azonosított variánsok közül csak azokat tartottuk meg, amelyek legalább 20-szoros lefedettségűek voltak, és legalább négy leolvasásban voltak benne. Referencia szekvenciaként, az emberi GRCh37 (hg19) verziójú genomot használtuk. Az azonosított variánsok annotálására a dbSNP (verzió: 139) és a COSMIC

(verzió: 68) adatbázisokat használtuk (175). A variánsok funkcionális annotálására az SNPeff (verzió: 3.5) programot használtuk (176).

#### 4.1.4 Kópiaszám eltérések azonosítása

A nem-kissejtes tüdőrákos betegek már feldolgozott kópiaszám eltérési (CNV) adatát szintén a GDC TCGA adatbázisból töltöttük le. Azokat az eltéréseket vettük figyelembe, ahol legalább 10 próba volt jelen egy pozícióban. Génamplifikáció esetén a szegmensek átlaga 0,2 feletti, gén deléció esetén pedig -0,2 alatti. A szegmensek annotálására a Human Gene Sets GTF annotációs fájlt használtuk fel, amelyet az Ensembl adatbázisból (GRCh37) töltöttünk le.

#### 4.1.5 TCGA programból elérhető mutációs adatok

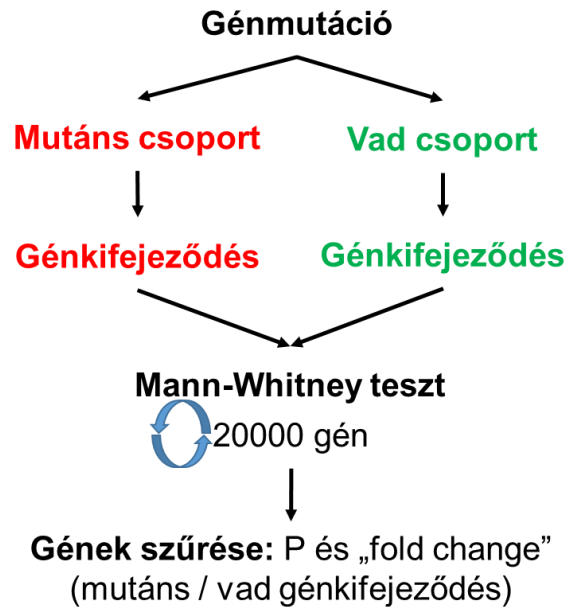
Az akut mieloid leukémia esetén a TCGA program már azonosított, feldolgozott mutációit használtam fel, az úgynevezett mutation annotated format (MAF) fájlok formájában. Mind a MuTect2, a MUSE, a VarScan és a SomaticSniper variánsazonosító algoritmusok eredményét felhasználtam és integráltam. Az adatok aggregálására és vizualizálására a „maftools” R Bioconductor programcsomagot (<http://bioconductor.org/packages/maftools/>) alkalmaztam.

## 4.2 Statisztikai számítások

### 4.2.1 A szomatikus mutációk és a génkifejeződés kapcsolata

A szomatikus mutációk génexpresszióra gyakorolt hatását Mann-Whitney teszttel vizsgáltuk. Az elemzés fő lépéseit a **3. ábra** mutatja. Első lépésben azon betegeket kell kiválasztani, amelyek rendelkeznek mind mutációs, mind pedig RNS szekvenálási adattal. Ezt követően a kiválasztott gén szomatikus mutációs státusza alapján a betegek RNS szekvenálási adatát két csoportra osztottuk: az egyik betegcsoport hordozta az adott génmutációt, a másik betegcsoport pedig nem. Mann-Whitney teszt segítségével megvizsgáltuk, hogy a mutáns és a vad betegcsoport között mely gének kifejeződése mutat szignifikáns változást. Minden gén esetén meghatároztuk a génexpressziós változás mértékét (fold change) a két betegcsoport között, a mutáns és a vad betegcsoport medián

génkifejeződési értékeinek hányadosával. A kapott génlistát a Mann-Whitney teszt  $P$ -értéke ( $P \leq 0,05$ ), valamint a fold change értékek ( $0,5 > \text{fold change} > 2$ ) alapján szűrtem.



**3. ábra.** A génmutációk és a génkifejeződés közötti összefüggés vizsgálata.

Azonban fontos kiemelni, hogy a munkám során Mann-Whitney tesztet nemcsak a szomatikus mutációk és a génexpresszió közötti összefüggés vizsgálatára alkalmaztam, hanem például a génkifejeződés és a vaszkuláris invázió közötti összefüggést is ezen statisztikai próbával vizsgálatunk máj hepatocelluláris karcinómában.

#### 4.2.1.1 A KRAS mutációk és kópiaszám eltérések hatása a génkifejeződésre nem-kissejtes tüdőrákban

Az elemzés során a KRAS génben található kódoló szomatikus mutációkat és kópiaszám eltéréseket használtuk fel. A Mann-Whitney teszt elvégzését követően kiválasztottam a legszignifikánsabb öt gént, amelyek génexpressziója összefüggött a KRAS génben található szomatikus mutációkkal. Az elemzésünk során ez az öt gén képezte az úgynevezett KRAS mutációkkal összefüggő génexpressziós mintázatot.

#### 4.2.1.2 Az NPM1 mutációk hatása a génkifejeződésre akut mieloid leukémiában

Az NPM1 gén szomatikus mutációk és a génexpresszió közötti kapcsolat vizsgálata megegyező módon történt a KRAS szomatikus mutációkkal. További analízis céljából azon szignifikáns gének kerültek kiválasztásra, amelyek esetében a mutáns és a vad betegcsoport medián génkifejeződési értékeinek hányadosa kisebb volt, mint 0,5 vagy nagyobb volt, mint 2.

#### 4.2.2 Mikro-RNS kifejeződés normál és tumoros májszövetben

A mikro-RNS kifejeződés összehasonlítására a párosított normál és tumoros máj hepatocelluláris karcinóma betegminták között, a Limma R Bioconductor programcsomagot alkalmaztam a *model.matrix()*, *lmFit()*, *eBays()* és a *topTable()* parancsok felhasználásával (177). Meghatároztam a mikro-RNS expressziós változás mértékét és irányultságát minden mikro-RNS esetén a tumoros és a hozzájuk tartozó normál mintapárok mikro-RNS expressziós értékeinek a hányadosaként.

#### 4.2.3 Túlélés elemzés

A génkifejeződés és a túlélés közötti kapcsolatot Cox regressziós modellel és Kaplan-Meier túlélés elemzéssel vizsgáltam. Az arányos túlélési kockázati feltételek igazolása elengedhetetlen előfeltétele a Cox regresszió alkalmazásának, amelyre a „survival” R programcsomagban (<https://CRAN.R-project.org/package=survival>) található *coxph()* függvényt alkalmaztuk. A Cox regressziós elemzéshez szintén a „survival” R programcsomagot használtuk. Az elemzések során a hazard ratio (HR), 95 % konfidencia intervallumok (CI) és a log-rank *P*-értékek kerültek kiszámításra. A túlélés elemzésből származó *P*-értékekre false discovery rate (FDR) korrekciót alkalmaztunk 5% -os FDR küszöb mellett (178). A munkám során uni- és multivariáns túlélés elemzést is végeztem, amelyeket a következő részben részletesen jellemezek. A túlélés elemzés során a Kaplan-Meier plotok készítésére a „survplot” R programcsomagot (<http://www.cbs.dtu.dk/~eklund/survplot/>) használtam.

#### 4.2.3.1 Univariáns túlélés elemzés

A nem-kissejtes tüdőrák esetén a KRAS mutációkkal összefüggő transzkriptomikus mintázat és a teljes túlélés kapcsolatát vizsgáltam. A transzkriptomikus mintázatban szereplő legjobb öt gén – amelyek kifejeződése a Mann-Whitney teszt során a legszignifikánsabb összefüggést mutatta a KRAS szomatikus mutációkkal – átlagos génexpresszióját korreláltam a teljes túléléssel. Vágópontként az expressziós értékek mediánját használtam.

Máj hepatocelluláris karcinómában az egyes mikro-RNS-ek expressziós értékei és a teljes túlélés közötti összefüggést vizsgáltam. A betegeket a kifejeződési értékük alapján történő két csoportra bontásához a medián vágópont helyett a túlélés elemzés szempontjából a legideálisabb vágópontot választottam ki, amelynek a számítása a következő képpen történt: minden gén esetén meghatároztam az egyes betegekben mért expressziós értékek alsó és felső kvartilis értékét. Mivel kis mintaszámú betegcsoporton végeztem az elemzést, az alsó kvartilis alatti és a felső kvartilis feletti értékek megbízhatatlan eredményekhez vezetnének, így ezeket az értékeket kihagytam a vizsgálatból. Ezt követően, a kvartilisek között, minden percentilis értéket a túlélés elemzés vágópontjaként használtam fel, és az FDR számítást követően a legszignifikánsabb *P*-értékhez tartozó vágópontot használtam a túlélés elemzés legjobb vágópontjaként (179).

Akut mieloid leukémiában az NPM1 szomatikus mutációkkal összefüggő gének expressziójának a hatását vizsgáltam a teljes túlélésre. A túléléselemzésben a fentebb leírt legjobb vágópontot használtam. Fontos kiemelni, hogy ebben az esetben – ellentétben a nem-kissejtes tüdőrákban végzett vizsgálatokkal – a túléléselemzést egyenként, minden génre elvégeztük.

Továbbá, szintén univariáns túlélés elemzéssel vizsgáltuk a máj hepatocelluláris karcinóma gének expressziójának prognosztikus hatását, medián vágópontot használva.

#### 4.2.3.2 Multivariáns túlélés elemzés

A nem-kissejtes tüdőrákban multivariáns Cox regresszióval vizsgáltam a túlélés és a KRAS génexpressziós mintázat kapcsolatát az életkor, a nem, a dohányzási állapot és a tumor stádiumának függvényében.



Továbbá, máj hepatocelluláris karcinómára is végeztünk multivariáns túlélés elemzést, ahol az egyes mikro-RNS-ek kifejeződésének, a betegek stádiumának és nemének az együttes hatását vizsgáltuk a teljes túlélésre.

#### **4.2.4 Daganat jellemző specifikus génexpressziós mintázat**

A daganat asszociált géneket nyolc fő csoportra lehet bontani, az alapján, hogy a daganatok kialakulásában milyen szerepet töltenek be (6), (180). A daganat jellemző specifikus génexpressziós mintázatnak, az adott daganat jellemző kialakításában részt vevő összes gén átlagos expresszióját vettük mind a 28 daganattípus esetén.

### **4.3 Klinikai minták feldolgoása**

Az *in vitro* validációra használt akut mieloid leukémiás klinikai minták a Semmelweis Egyetem I.sz. Patológiai és Kísérleti Rákkutató Intézetből származtak. A betegmintákkal történő munkára az Egészségügyi Tudományos Tanács Tudományos és Kutatásetikai Bizottság (ETT TUKEB) – 14383-2/2017/EKU engedélye jogosított fel. A minták NPM1 mutációs státuszának meghatározására Sanger szekvenálást, a génkifejeződés változásának vizsgálatára pedig kvantitatív polimeráz lánreakciót (qPCR) alkalmaztunk. A perifériális vér és csontvelő mintákból DNS-t preparáltunk a High Pure PCR Template kit segítségével (Roche, Basel, Switzerland). A DNS koncentrációját UV spektrofotométerrel ellenőriztük (NanoDrop, Thermo Fisher Scientific, Waltham, Massachusetts, USA).

#### **4.3.1 Az NPM1 gén Sanger szekvenálása**

Az NPM1 gén felsokszorosítását az AmpliTaqGold polimeráz mix (Thermo Fisher Scientific, Waltham, Massachusetts, USA) felhasználásával végeztük a PE 2720 GeneAmp PCR készülék (Perkin-Elmer, Waltham, Massachusetts, USA) segítségével. A forward (5'-TTCCATACATACTTAAAACCAA-3') és reverz (5'-TGGTTCCTTAACCACATTTCTTT-3') irányultságú oligonukleotidokat 25 ml-es végtérfogatban alkalmaztuk. A reakcióelegy 2x AmpliTaqGold polimerázt, 400-400 nM koncentrációjú forward és reverz oligonukleotidot, valamint 100 ng DNS-t tartalmazott. Az amplifikáció a következő lépésekből állt: 1) 10 perces denaturáció 95°C-on, 2) inkubálás 40 cikluson keresztül 30 másodpercig 95°C-on, 3) inkubálás 60 másodpercig

56°C-on, 4) inkubálás 60 másodpercig 72°C-on. A PCR termékek tisztításához a ExoSAP-IT PCR kítet (Affymetrix, Santa Clara, California USA) és a Big Dye Terminator kítet v3.1 használtuk (Thermo Fisher Scientific, Waltham, Massachusetts, USA). A szekvenálási reakciót az ABI 3500 Genetic Analyzer készülék (Thermo Fisher Scientific, Waltham, Massachusetts, USA) segítségével végeztük. A szekvenálási adatok kiértékeléséhez a SeqA6 szoftvert (Thermo Fisher Scientific, Waltham, Massachusetts, USA) használtuk.

#### 4.3.2 RNS izolálás

A perifériális vér és csontvelő mintákat két órán keresztül homogenizáltuk hemolízis oldattal, amely a következőket tartalmazta: 0,15 M NH<sub>4</sub>Cl, 10 M NH<sub>4</sub>HCO<sub>3</sub> és 0,1 M pH=4 EDTA (Sigma-Aldrich, St. Louis, MO, USA). A hemolízist követően a mintákat 10 percig 1800 fordulaton (RPM) centrifugáltuk, majd 1x foszfát pufferrel mostuk (PBS, Lonza, Basel, Switzerland). A teljes RNS-t a sejtekből TRIzol reagens (Invitrogen, Waltham, Massachusetts, USA) hozzáadásával nyertük ki. Az RNS koncentrációját UV spektrofotométerrel ellenőriztük (NanoDrop, Thermo Fisher Scientific, Waltham, Massachusetts, USA).

#### 4.3.3 Kvantitatív polimeráz láncreakció

A kvantitatív polimeráz láncreakció (qPCR) reakcióhoz minden mintából 1 µg teljes RNS-t írtunk át cDNS-sé 25 µl-es végtérfogatban a High-Capacity cDNA Reverse Transcription reagens kít (Thermo Fisher Scientific, Waltham, Massachusetts, USA) segítségével. A qPCR reakciót a CFX96 Real-Time PCR készülékkel (Bio-Rad Laboratories, Hercules, California, USA) és a SensiFAST SYBR No-ROX reagens kít (Bioline Reagents, London, UK) felhasználásával végeztük.

Az oligonukleotidokat exon-exon határokra terveztük lefedve az összes transzkript változatot az adott génre nézve (**I. táblázat**). Továbbá, minden oligonukleotid esetén az olvadáspont (T<sub>m</sub>) értéke 60°C volt. A GAPDH és a TBP géneket használtuk referencia génként.

**1. táblázat.** A tervezett oligonukleotid szekvenciák a kvantitatív PCR reakcióhoz.

Mutáció	Gén	NCBI nukleotid azonosító	Oligonukleotid szekvencia	Hosszúság (bázispár)
<b>IDH1</b>	<b>RASGRP3</b>	NM_015376.2	F: 5'-CAAGCCAACCTTCTGCGAAC-3' R: 5'-TGGCTCCACAGTCTTTGCAT-3'	83
<b>IDH2</b>	<b>NPDC1</b>	NM_015392.3	F: 5'-GACTACGCCACTGCGAAGG-3' R: 5'-CTTTATGCCGCTCCAGGCAC-3'	139
<b>NPM1</b>	<b>HOXA5</b>	NM_019102.3	F: 5'-AGCTGCACATAAGTCATGACAACA-3' R: 5'-TCAATCCTCCTTCTGCGGGT-3'	136
<b>NPM1</b>	<b>HOXB5</b>	NM_002147.3	F: 5'-AACTCCTTCTCGGGGCGTTAT-3' R: 5'-CATCCCATTGTAATTGTAGCCGT-3'	138
<b>NPM1</b>	<b>HOXA10</b>	NM_018951.3	F: 5'-GAGAGCAGCAAAGCCTCGC-3' R: 5'-CCAGTGTCTGGTGCTTCGTG-3'	127
<b>NPM1</b>	<b>ITM2A</b>	NM_001171581.1	F: 5'-TGTTGCTGGGGAAGTCTAT-3' R: 5'-GATATCTGCCACTCGCCAGTTT-3'	102
<b>NPM1</b>	<b>MEIS1</b>	NM_002398.2	F: 5'-CACGGGACTCACCATCCTTC-3' R: 5'-TGACTTACTGCTCGGTTGGAC-3'	99
<b>NPM1</b>	<b>PBX3</b>	NM_006195.5	F: 5'-CACACCTCAGCAACCCCTAC-3' R: 5'-ACCAATTGGATACCTGTGACACT-3'	90
-	<b>GAPDH</b>	NM_002046.6	F: 5'-AAATCAAGTGGGGCGATGCT-3' R: 5'-CAAATGAGCCCCAGCCTTCT-3'	86
-	<b>TBP</b>	NM_003194.4	F: 5'-GCACAGGAGCCAAGAGTGAA-3' R: 5'-TCACAGCTCCCCACCATGT-3'	127

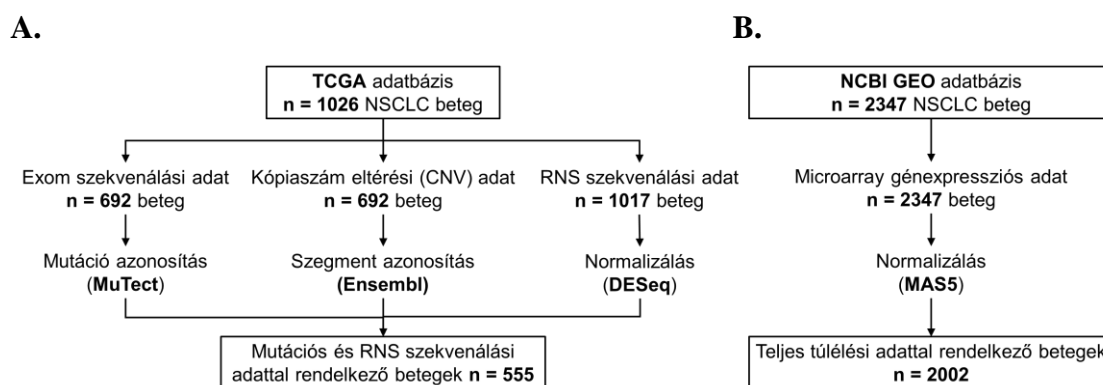
A reakciót 20 µl-es végtérfogatban végeztük, amely tartalmazott 1 µl cDNS-t és 125 nM oligonukleotidot. A kezdeti denaturációs szakaszt (2 perc – 95°C) egy 40 cikluson keresztül tartó háromlépéses folyamat követett: 1) inkubáció 15 másodpercig 95°C-on, 2) inkubáció 15 másodpercig 60°C-on, 3) inkubáció 30 másodpercig 72°C-on. Minden reakciót három szoros ismétlésben végeztünk és minden génre meghatároztuk az áttörési ciklusszám értékeket (Ct). A génexpressziós változás meghatározására a  $\Delta Ct$  módszert alkalmaztuk és a meghatároztuk a  $2^{(-\Delta Ct)}$  értékeket. Az adatok elemzését a WinSTAT program (<http://www.winstat.com/>) segítségével végeztük.

## 5 EREDMÉNYEK

### 5.1 KRAS mutáció függő génexpressziós változások vizsgálata nem-kissejtes tüdőrákban

#### 5.1.1 A nem-kissejtes tüdőrák adatbázis felépítése

A nem-kissejtes tüdőrák adatok a TCGA és az NCBI GEO adatbázisból származnak. A TCGA program, összesen 1026 nem-kissejtes tüdőrákos adatát tartalmazza. Teljes exom szekvenálási adat és kópiaszám változási adat 692, valamint RNS szekvenálási adat 1017 beteghez volt elérhető. Az elemzés során azokat a betegeket válogattam ki, amelyek mind a három adattípussal rendelkeznek, így a teljes mintaszám 555 beteg volt (**4. ábra A**).



**4. ábra.** A nem-kissejtes adatbázis felépítése. A TCGA adatbázisból (**A**) a teljes exom, az RNS szekvenálási és a kópiaszám eltérési adatok, míg az NCBI GEO adatbázisból (**B**) a microarray génexpressziós adatok származnak.

A TCGA klinikai és teljes túlélési adatok összesen 967 beteghez voltak elérhetőek, ahol a medián teljes túlélési idő 8,68 hónap volt. Megvizsgáltam az összefüggést az egyes klinikopatológiai paraméterek – tumor méret ( $P = 0,0001$ ), nyirokcsomó státusza ( $P = 2,03E-06$ ) és metasztázis státusza ( $P = 0,03$ ) – és a túlélés között, ahol mindegyik esetben szignifikáns összefüggést kaptam.

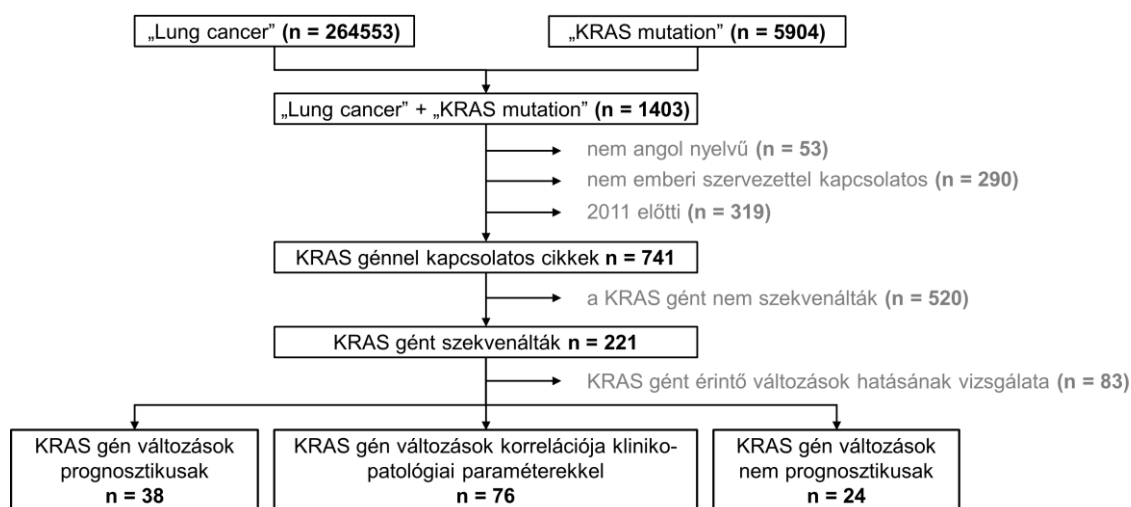
Az NCBI GEO adatbázis 2347 nem-kissejtes tüdőrákos beteg microarray génexpressziós adatát tartalmazta. Összesen 2002 beteghez tartozott teljes túlélési adat,

ahol a medián túlélési idő 41,6 hónap volt (**4. ábra B**). A túlélés elemzés során szignifikáns összefüggést kaptunk a teljes túlélés és a tumor mérete ( $P < 1E-16$ ), nyirokcsomó státusza ( $P < 1E-16$ ), metasztatizáló képessége ( $P = 0,003$ ) között.

Látható, hogy a microarray génexpressziós adatokhoz tartozó túlélési adat duplája az RNS szekvenáláshoz tartozó túlélési adatoknak. Ezért a túlélés elemzések során a megbízhatóbb eredmények érdekében a nagyobb mintaszámmal rendelkező microarray génexpressziós adatokat használtam fel.

### 5.1.2 A KRAS mutációk

A vizsgálatok első részében irodalmi keresést végeztem az NCBI PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>) adatbázisban, ahol a célom az volt, hogy azonosítsam azokat a közleményeket, amelyek vizsgálták a KRAS mutációk prognosztikus hatását. Az irodalmi keresés menetét az **5. ábra** mutatja.



**5. ábra.** KRAS mutációk prognosztikus hatásának vizsgálata irodalmi keresése alapján.

Az NCBI PubMed adatbázisban a „lung cancer” és a „KRAS mutation” keresőkifejezések alkalmazásával összesen 1403 közleményt azonosítottam, amelyben leírták a KRAS mutációk prognosztikus szerepét. A keresést 2016. január 31.-én végeztem, amely során elsősorban a 2011 és 2016 között megjelent, angol nyelvű és humán mintákat felhasználó közleményekre szűrtünk.

Összesen 221 tanulmány volt, amelyben a KRAS gént szekvenálták és 138 esetben korreláltatták a KRAS mutációkat túlélés végkimenetellel vagy egyéb klinikopatológiai paraméterekkel. Ebből, 38 közleményben a KRAS mutációk prognosztikusak voltak, és 76 cikk esetén a KRAS mutációk összefüggtek a páciens dohányzási státuszával, a kezelésre adott válaszával, valamint a tumor méretével. Továbbá, a tanulmányok 17%-ban nem volt a KRAS mutációknak prognosztikus hatása. A **2. táblázat** alapján látható, hogy az egyes tanulmányok nem írják le pontosan, hogy melyik nem-kissejtes tüdőrák altípust vizsgálták, illetve voltak olyanok is, ahol a vizsgálat több altípusra is kiterjedt.

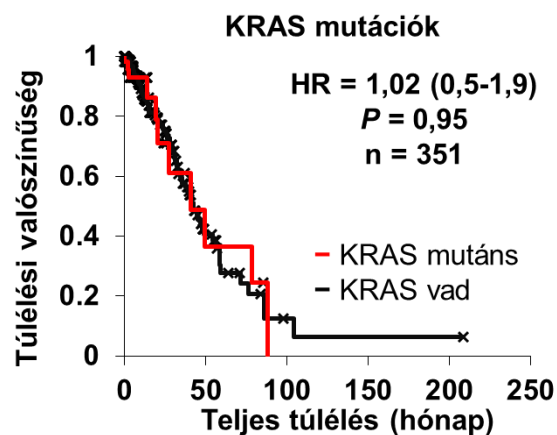
**2. táblázat.** Korábbi tanulmányok, amelyek a KRAS mutációk hatását vizsgálták különböző típusú tüdődaganatban.

	<b>Összes tanulmány n (%)</b>	<b>NSCLC altípus nem ismert n (%)</b>	<b>AC n (%)</b>	<b>SCC n (%)</b>	<b>LCC n (%)</b>	<b>LCN n (%)</b>	<b>SC n (%)</b>	<b>ASCC n (%)</b>
<b>A KRAS mutációk befolyásolták a túlélést</b>	38 (28%)	17 (12%)	17 (12%)	2 (1%)	2 (1%)	-	-	-
<b>A KRAS mutációk egyéb klinikai-patológiai paraméterrel függtek össze</b>	76 (55%)	27 (20%)	37 (27%)	4 (3%)	3 (2%)	2 (1%)	1 (1%)	2 (1%)
<b>A KRAS mutációknak nincsen szignifikáns hatásuk</b>	24 (17%)	13 (9%)	9 (7%)	1 (1%)	1 (1%)	-	1 (1%)	1 (1%)

54

**NSCLC** – Non–small cell lung cancer: Nem-kissejtes tüdőrák, **AC** – Adenocarcinoma: Tüdő adenokarcinóma, **SCC** – Squamous cell carcinoma: Tüdő laphámsejtes karcinóma, **LCC** – Large cell carcinoma: Nagy sejtes karcinóma, **LCN** – Large cell neuroendocrine carcinoma: Nagy sejtes neuroendokrin karcinóma, **SC** – Sarcomatoid carcinoma: Szarkomatoid karcinóma, **ASCC** – Adenosquamous carcinoma: Kevertsejtes karcinóma

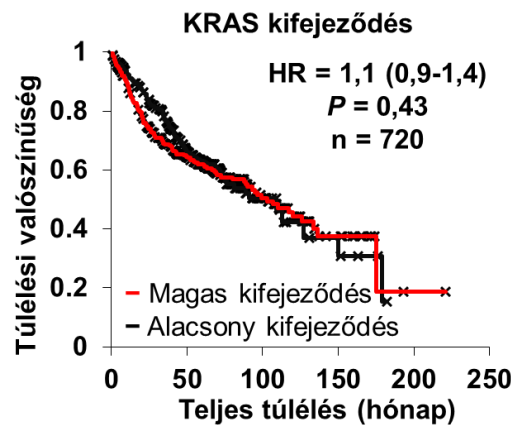
Az irodalmi keresést követően megvizsgáltam, hogy a KRAS szomatikus mutációk közvetlenül befolyásolják-e a betegek túlélését. Ehhez a TCGA adatbázis teljes exom szekvenálásból azonosított szomatikus mutációkat és a hozzájuk tartozó túlélési adatokat használtam fel. Összesen 61 nem-kissejtes tüdőrákos beteg hordozott szomatikus KRAS mutációt, ebből 60 beteg adenocarcinoma és mindössze egy beteg a laphámsejtes karcinóma altípusba sorolható. Az eredmények alapján nem tapasztaltunk szignifikáns összefüggést a KRAS mutációk és a teljes túlélés között **6. ábra**. Az összes további vizsgálatokat, az adenokarcinóma altípuson belül fogom bemutatni.



**6. ábra.** A KRAS szomatikus mutációk és a teljes túlélés összefüggése nem-kissejtes tüdőrákban.

A KRAS mutációk mellett a KRAS kifejeződés hatását is megvizsgáltuk a teljes túlélésre a microarray adatok felhasználásával **7. ábra**. A Kaplan-Meier ábrán látható, hogy KRAS génkifejeződés és a teljes túlélés között nem azonosítottam szignifikáns összefüggést. A microarray adatok mellett, az RNS szekvenálási adatok felhasználása mellett is megvizsgáltam a kapcsolatot a KRAS kifejeződés és a teljes túlélés között, azonban ebben az esetben sem kaptam szignifikáns összefüggést (HR = 1,3; P = 0,17).





**7. ábra.** A KRAS expresszió hatása a teljes túlélésre a microarray adatok alapján nem-kissejtes tüdőrákban.

### 5.1.3 A KRAS mutációkkal összefüggő transzkriptomikus változások

A vizsgálat során a céloom az volt, hogy az RNS szekvenálási adatok alapján olyan génextpressziós változásokat azonosítsak, amelyek összefüggést mutatnak a KRAS génben található szomatikus mutációkkal. Majd a legszignifikánsabb öt gént kiválasztva, a microarray génextpressziós adatok felhasználásával, vizsgálom az összefüggést a génextpresszió és a teljes túlélés között.

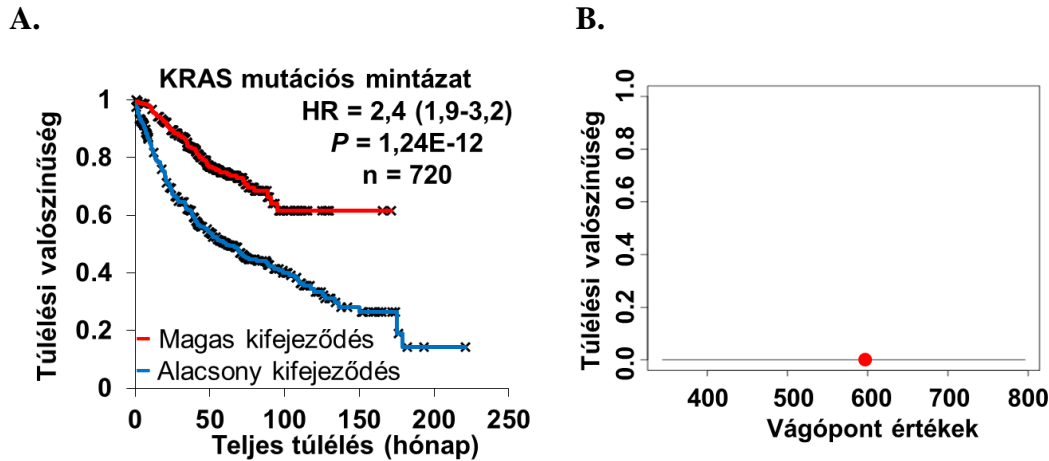
Első lépésben ellenőriztem, hogy milyen típusú KRAS mutációk találhatóak a betegekben. Elmondható, hogy a mutációk mind a 60 KRAS mutáns tüdő adenokarcinómás beteg esetén a gén kódoló részét érintették, továbbá volt hét olyan beteg, akik nem kódoló KRAS mutációt is hordoztak.

Mann-Whitney teszt segítségével összehasonlítottuk a 11500 gén kifejeződését a KRAS mutáns és vad típusú tüdő adenokarcinómás betegcsoportok között, az RNS szekvenálási adatok felhasználásával. Azonosítottam az öt legszignifikánsabb gént – FOXRED2, PEX3, KRAS, TOP1 és ABL2 –, amelyek kifejeződése korrelált a KRAS génben található szomatikus mutációkkal **3. táblázat**.

**3. táblázat.** A KRAS szomatikus mutációkkal összefüggő legszignifikánsabb öt gén.

Gének	<i>P</i> -értékek	Génkifejeződés irányultsága (átlagos mutáns / vad kifejeződés)	Átlag (mutáns csoport)	Átlag (vad csoport)
<b>FOXRED2</b>	1,14E-06	0,62	944	1530
<b>KRAS</b>	6,31E-06	1,29	4472	3462
<b>TOP1</b>	7,06E-06	1,29	10272	7979
<b>PEX3</b>	1,51E-05	0,77	499	647
<b>ABL2</b>	1,81E-05	1,22	2534	2077

Ezt követően kiszámoltam a KRAS szomatikus mutációkkal összefüggő transzkriptomikus mintázatot, ahol az öt gén átlagos kifejeződését vettem a micorarray génchip adatok alapján. A túlélés elemzés során, a medián vágópont használatával szignifikáns összefüggést azonosítottam a transzkriptomikus mintázat és a teljes túlélés között (HR = 2,4; 95% CI = 1,9–3,2;  $P = 1,24E-12$ ) (**8. ábra A**). Fontos kiemelni, hogy a Cox regressziós elemzést nemcsak a medián vágópont használatával végeztem el. A génexpressziós értékek alsó és felső kvartilis értékei között, minden percentilis értéket a túléléselemzés egy vágópontjának feleltettem meg és elmondható, hogy minden vágópont esetén szignifikáns összefüggést tapasztaltam a génexpresszió és a teljes túlélés között. A legszignifikánsabb *P*-értékhez tartozó vágópontot a piros pont jelöli (**8. ábra B**).

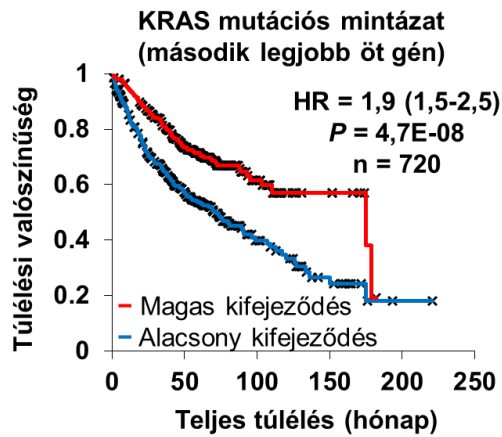


**8. ábra.** A KRAS mutációk transzkriptomikus mintázatának hatása a túlélésre (**A**) és a különböző génexpressziós vágópontok alkalmazásával mutatott prognosztikus ereje (**B**).

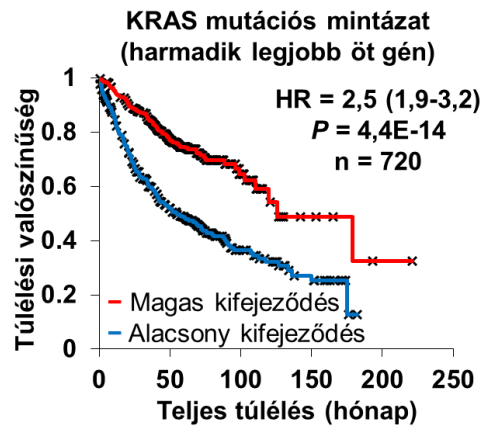
Multivariáns túlélés elemzés során megvizsgáltam, hogy a KRAS szomatikus mutációkkal összefüggő transzkriptomikus mintázat a betegek életkorával, nemével, dohányzási anamnézisével és a tumor stádiumával együtt, hogyan befolyásolja a teljes túlélést. A vizsgálat során csak a KRAS transzkriptomikus mintázat ( $P = 0,01$ ), az életkor ( $P = 0,01$ ) és a tumor stádium ( $P = 5E-07$ ) tényezők bizonyultak szignifikánsnak.

Az elemzés további részében megvizsgáltam, hogy a második valamint a harmadik leszignifikánsabb öt gén által képzett KRAS transzkriptomikus mintázatnak milyen hatása van a teljes túlélésre. Az eredmények azt mutatják, hogy mind a második (HR = 1,9; 95% CI = 1,5–2,5;  $P = 4,7E-08$ ) és harmadik (HR = 2,5; 95% CI = 1,9–3,2;  $P = 4,4E-14$ ) legjobb öt gén esetén, a transzkriptomikus mintázatnak erős prognosztikus hatása van (**9. ábra**).

A.



B.



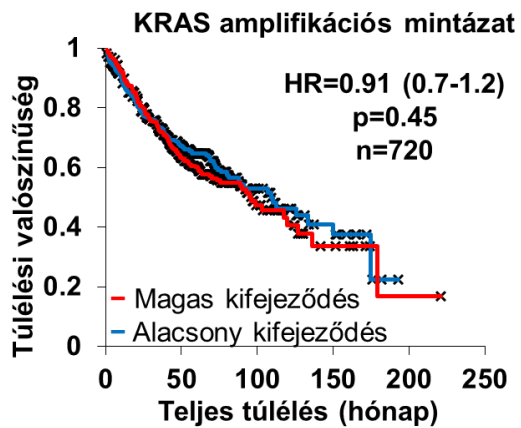
**9. ábra.** A második (A) és a harmadik (B) legszignifikánsabb KRAS szomatikus mutációkkal összefüggő öt gén hatása a túlélésre.

#### 5.1.4 A KRAS amplifikációval és delécióval összefüggő génexpressziós mintázat hatása a túlélésre

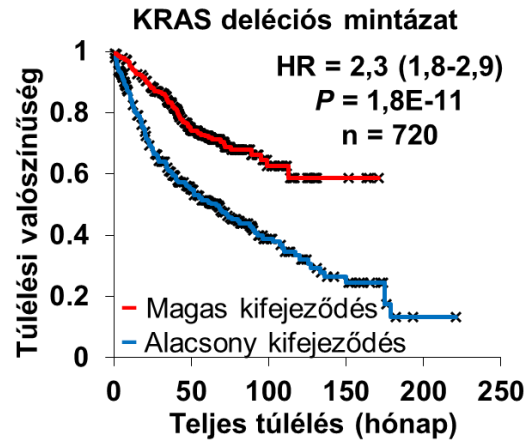
A munkám során a KRAS szomatikus mutációk mellett, a KRAS génben található amplifikációkkal és deléciókkal összefüggő transzkriptomikus mintázat prognosztikus hatását is vizsgáltam tüdő adenokarcinómában.

A KRAS amplifikációt 30, míg a KRAS deléciót 12 tüdő adenokarcinómás beteg hordozta. Vizsgálatom során a KRAS amplifikációkkal összefüggő transzkriptomikus mintázat nem mutatott szignifikáns összefüggést a teljes túléléssel (HR = 0,91; 95% CI = 0,7–1,2; P = 0,45) (10. ábra A).

A.



B.



**10. ábra.** A KRAS amplifikációval (A) és delécióval (B) összefüggő génexpressziós mintázat hatása a túlélésre.

Ezzel szemben a KRAS deléciókkal asszociált expressziós mintázat szignifikáns összefüggést mutatott a teljes túléléssel (HR = 2,3; 95% CI = 1,8–2,9;  $P = 1,8E-11$ ) (**10. ábra B**). A KRAS amplifikációkkal és deléciókkal összefüggő kifejeződést mutató gének a **4. táblázatban** láthatóak.

**4. táblázat.** A KRAS amplifikációval (A) és delécióval (B) összefüggő expressziókülönbséget mutató legszignifikánsabb öt gén.

**A.**

Gének	<i>P</i> -értékek	Génkifejeződés irányultsága (átlagos mutáns / vad kifejeződés)	Átlag (mutáns csoport)	Átlag (vad csoport)
<b>KRAS</b>	1,55E-09	1,94	6553	3379
<b>ETNK1</b>	5,29E-06	1,57	5865	3740
<b>FAM60A</b>	8,00E-06	1,77	7491	4243
<b>TMEM185B</b>	8,33E-06	1,29	1623	1258
<b>CLEC11A</b>	1,73E-05	0,55	298	544

**B.**

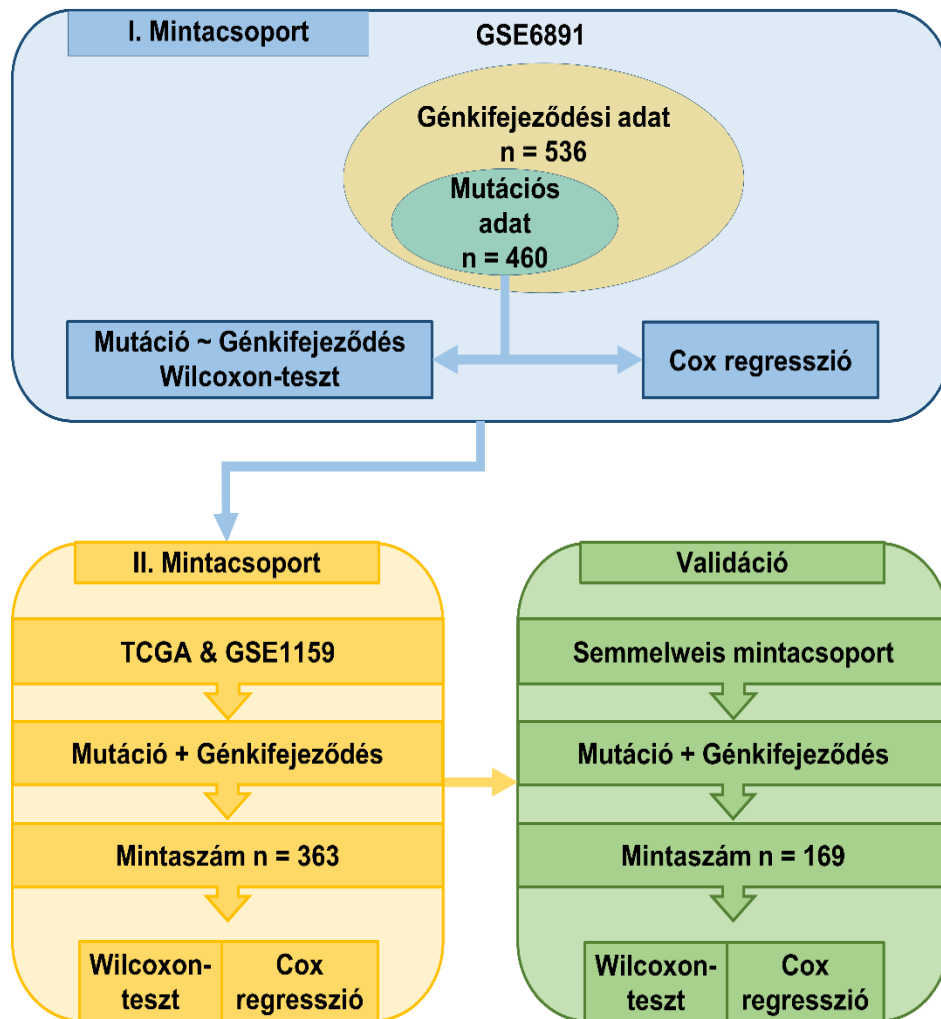
Gének	<i>P</i> -értékek	Génkifejeződés irányultsága (átlagos mutáns / vad kifejeződés)	Átlag (mutáns csoport)	Átlag (vad csoport)
<b>HSDL2</b>	5,91E-05	1,64	3923	2398
<b>GAK</b>	1,06E-04	0,63	3132	4982
<b>YARS2</b>	1,44E-04	0,61	517	846
<b>MCOLN1</b>	1,50E-04	0,63	733	1155
<b>CMAS</b>	1,65E-04	0,59	1095	1862

Továbbá, a KRAS deléciókkal összefüggő transzkriptomikus mintázat esetén is végeztem multivariáns túlélés elemzést, azonban ebben az esetben csak az életkor ( $P = 0,01$ ) és a tumor stádiuma ( $P = 3E-07$ ) változók mutattak szignifikáns összefüggést a teljes túléléssel.

## 5.2 Az NPM1 mutációk transzkriptomikus lenyomata és validálása akut mieloid leukémiában

### 5.2.1 Akut mieloid leukémia adatbázis

Az akut mieloid leukémia adatbázis négy, egymástól független génexpressziós, mutációs és klinikai adatokkal rendelkező vizsgálati betegcsoporton alapszik (11. ábra).



11. ábra. Az akut mieloid leukémia adatbázis felépítése.

5. táblázat. Az adatbázis klinikai-patológiai jellemzői.

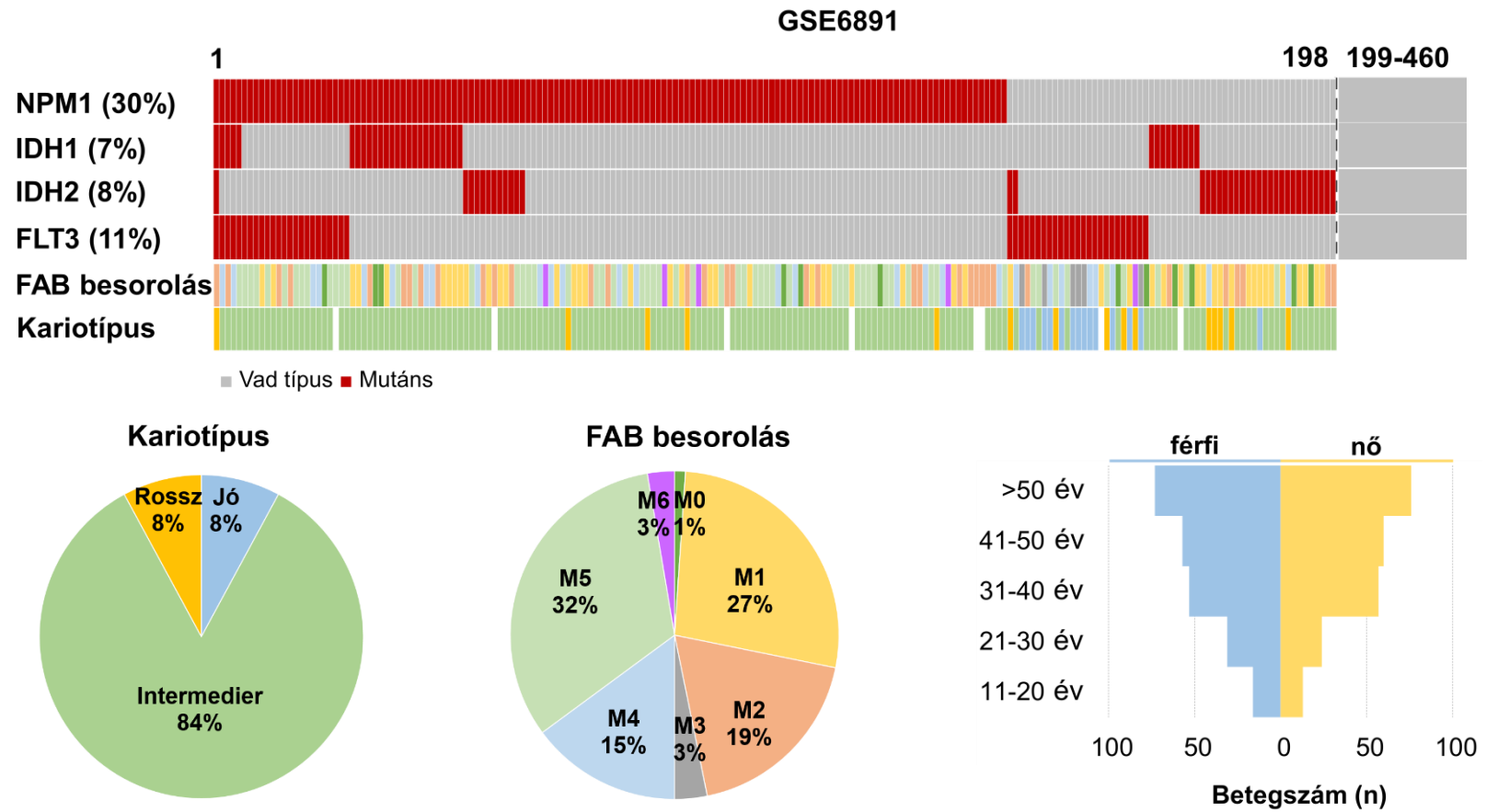
	<b>GSE6859</b>	<b>TCGA</b>	<b>GSE1159</b>	<b>Semmelweis mintacsoport</b>
<b>Összes minta (n)</b>	536	200	293	169
<b>Mutációs és génkifejeződési adattal rendelkező minták (n)</b>	460	116	247	169
<b>Életkor tartomány (medián)</b>	15-60 (43)	18-89 (58)	15-60 (42)	0-85 (59)
<b>Nem (nő/férfi)</b>	230/230	91/109	128/119	84/85
<b>Medián túlélési idő (hónap)</b>	18,7	12	17	6,92
<b>Kariotípus (jó/intermedier/rossz/ismeretlen)</b>	97/261/92/86	-	60/136/48/49	12/97/25/35
<b>FAB altípus (M0/M1/M2/M3/M4/M5/M6)</b>	16/95/105/24/84/104/6	-	6/55/54/17/43/62/3	-



Az első vizsgálati csoport a GSE6891-es betegcsoport (168) volt, amely 536 betegmintát tartalmazott. A vizsgálat során ezt a betegcsoportot elsődleges vizsgálati halmazként, úgynevezett tanulóhalmazként használtam fel. Minden betegminta rendelkezett microarray génexpressziós adattal, amelyből 460 páciens esetében ismert volt az NPM1 gén szomatikus mutációs státusza. A klinikai adatok tekintetében 536 beteg rendelkezett teljes túlélési adattal, ahol a medián teljes túlélési idő 18,7 hónap volt. További klinikai-patológiai paraméterek közül ismert volt a betegek neme, életkora, kariotípusa, valamint az French-American-British (FAB) osztályozáson alapuló altípus besorolása (181) (**5. táblázat és a 12. ábra**). A betegek körében az NPM1 volt a leggyakrabban mutált gén, csaknem a betegek 30%-a hordozta ezen gén mutációit (**12. ábra**).

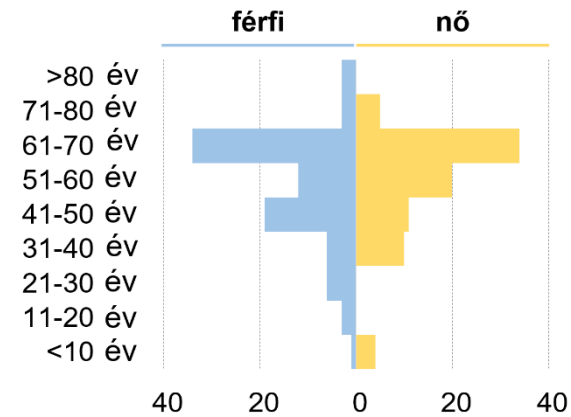
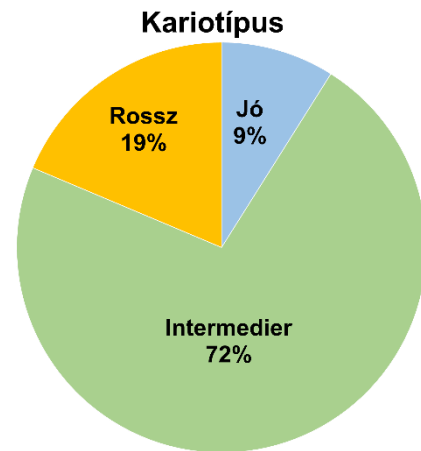
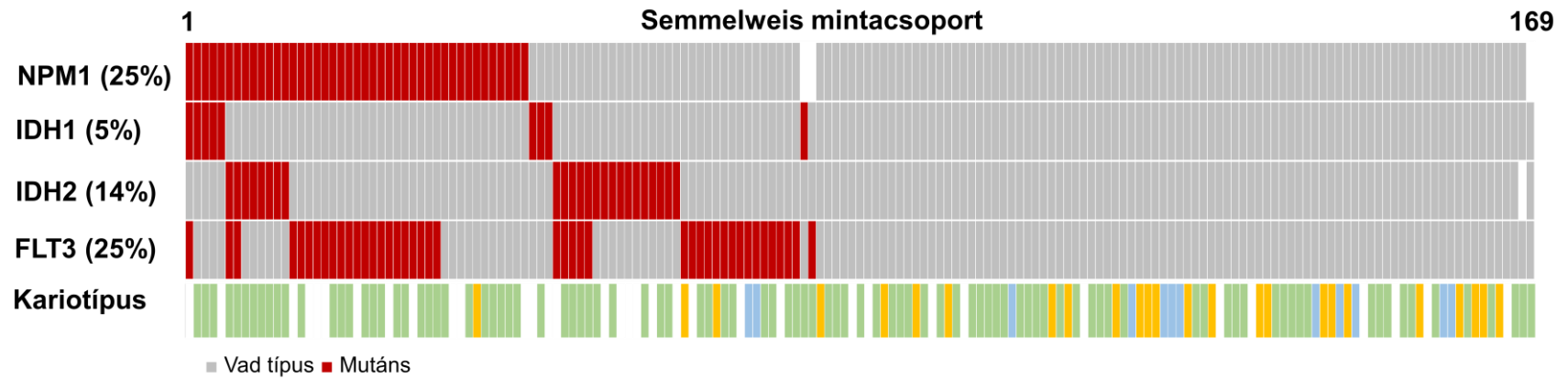
A TCGA program adatai képezték a második vizsgálati csoportot, vagy úgynevezett tanulóhalmazt, amely 200 akut mieloid leukémiás beteg RNS szekvenálási adatát tartalmazta (**5. táblázat**). A 200 betegből, mintegy 116 rendelkezett szomatikus mutációs adattal is. Azonban teljes túlélési adat, az RNS szekvenálási és mutációs adattal is rendelkező betegek kevesebb, mint a feléhez tartozott, így nem végeztem túlélés elemzést ezen adatok felhasználásával. A betegek 17%-a hordozott NPM1 szomatikus mutációt, amelyek döntő többsége úgynevezett kereteltolódási (frame shift) inszerció volt.

A harmadik vizsgálati csoportot a GSE1159 tanulmány (169) képezte, amely 293 akut mieloid leukémiás beteg microarray génexpressziós adatát tartalmazta (**5. táblázat**). A teljes túlélési adat 260 beteg esetén volt elérhető, valamint 247 páciens rendelkezett génexpressziós és NPM1 gén mutációs státusszal, ebből 77 beteg az NPM1 mutációra pozitív volt.



12. ábra. A GSE6891-es betegcsoport klinikai adatainak ábrázolása.

Az utolsó mintacsoport, a Semmelweis mintacsoport volt, amelyet az elemzésünkben validációs célra használtam fel. Összesen 169 akut mieloid leukémiás minta állt rendelkezésünkre, ahol a Sanger szekvenálás során az NPM1 gén a betegek 25%-ban tartalmazott mutációt. Szintén a betegek 25%-a hordozott FLT3 mutációt, 14%-a IDH2 mutációt és csaknem 5%-a hordozott IDH1 mutációt. A minták 52,6%-a csontvelői, 47,4%-a perifériális vérből származott. Minden minta rendelkezett teljes túlélési adattal és a medián túlélési idő 6,92 hónap volt. További klinikai-patológia adatok közül kiemelném, hogy a betegek kariotípusa esetén a citogenetikai kockázat hasonlóan a GSE6891-es betegcsoporthoz nagyrészt intermedier volt. A betegek döntő többségét a 60 és 70 év közötti korosztály képviselte (**13. ábra**).



13. ábra. A Semmelweis mintacsoport jellemzése.

## 5.2.2 Az NPM1 mutációkkal összefüggő gének

Az első vizsgálati csoporton (GSE6891) végzett Mann-Whitney elemzés során 85 gént azonosítottam, amelyek expressziója különbözött az NPM1 mutáns és vad típusú betegek között ( $P \leq 0,05$  és  $0,5 > FC > 2$ ). Ezen gének kifejeződése közül 57 gén felül, míg 28 gén alul szabályozódott (**6. táblázat A**). Ezt követően, Cox regressziós elemzés során 47 gén expressziója mutatott szignifikáns összefüggést a teljes túléléssel 10%-os FDR hibahatár alatt (**6. táblázat B**).

**6. táblázat.** Az NPM1 mutáns és vad típusú betegcsoport között eltérő kifejeződést mutató gének (**A**) és hatásuk a teljes túlélésre (**B**).

**A.**

Gének	Mutáns medián	Vad medián	FC	P-érték					
					<i>ITM2C</i>	834,5	2579	0,32	2,45E-27
<i>HOXB3</i>	598,5	189	3,17	5,12E-45	<i>CD200</i>	77,5	664,5	0,12	3,38E-27
<i>HOXA5</i>	2799	100	27,99	1,87E-44	<i>H2AFY2</i>	588,5	235,5	2,5	1,41E-25
<i>HOXB2</i>	2282	220,5	10,35	2,85E-43	<i>CCND2</i>	2266,5	4802,5	0,47	2,54E-24
<i>HOXB6</i>	1017	83,5	12,18	4,55E-43	<i>GYPC</i>	803,5	2440,5	0,33	5,68E-23
<i>HOXA10</i>	2952	683,5	4,32	2,22E-39	<i>RASGRP3</i>	1022,5	278,5	3,67	2,54E-22
<i>PBX3</i>	3544,5	654	5,42	5,45E-39	<i>JUP</i>	702	1944	0,36	6,90E-22
<i>MEIS1</i>	2264,5	431	5,25	1,12E-38	<i>PRKAR2B</i>	2554	871,5	2,93	5,88E-21
<i>HOXB5</i>	840,5	321,5	2,61	1,35E-38	<i>TSPAN13</i>	343,5	1157,5	0,3	1,59E-20
<i>PDGFD</i>	665,5	227,5	2,93	2,30E-33	<i>MANIA1</i>	1746,5	3552,5	0,49	2,11E-20
<i>SMC4</i>	4415	2043,5	2,16	2,75E-32	<i>ITM2A</i>	977,5	2989	0,33	3,81E-20
<i>COL4A5</i>	1342,5	100,5	13,36	1,00E-31	<i>HIF0</i>	562,5	2117	0,27	1,45E-18
<i>DMXL2</i>	4371,5	1398	3,13	3,00E-31	<i>C3ARI</i>	1880	831,5	2,26	2,43E-18
<i>PLA2G4A</i>	593,5	262,5	2,26	6,11E-29	<i>BAHCCI</i>	1864	770	2,42	2,77E-18
<i>CD34</i>	257,5	1854	0,14	7,04E-29	<i>LPAR6</i>	318	964	0,33	3,72E-18
<i>APP</i>	49	839	0,06	3,44E-28	<i>IFITM1</i>	1370	2974,5	0,46	4,47E-18
<i>BAALC</i>	78,5	611	0,13	3,49E-28	<i>SELIL3</i>	1668,5	766,5	2,18	2,28E-17

<i>LGALS3BP</i>	2999,5	794	3,78	3,47E-17	<i>PF4</i>	514,5	197	2,61	1,17E-08
<i>MEST</i>	986	3028	0,33	3,88E-17	<i>PROM1</i>	320	1699,5	0,19	1,96E-08
<i>HIST2H2BE</i>	3068	1500	2,05	5,65E-16	<i>LILRB2</i>	976	382,5	2,55	2,19E-08
<i>CPVL</i>	1442,5	553,5	2,61	1,03E-15	<i>CYTL1</i>	342,5	751,5	0,46	3,27E-08
<i>SLC38A1</i>	818,5	1878,5	0,44	2,49E-15	<i>NPR3</i>	479,5	1440	0,33	3,50E-08
<i>EGFL7</i>	276,5	728	0,38	3,33E-15	<i>SERPINA1</i>	4521	1940,5	2,33	8,33E-08
<i>PRKD3</i>	331	805	0,41	6,67E-15	<i>HK3</i>	1125	432,5	2,6	3,45E-07
<i>VNN1</i>	1144	261	4,38	9,17E-15	<i>TMEM176B</i>	744	263	2,83	4,79E-07
<i>TLR4</i>	1193	524	2,28	3,39E-14	<i>SLC4A1</i>	470	1161,5	0,4	6,02E-07
<i>CTSG</i>	3670	948,5	3,87	1,66E-13	<i>HBB</i>	6031	19089	0,32	1,43E-06
<i>JAG1</i>	1095,5	480,5	2,28	2,63E-13	<i>VCAN</i>	2036	491,5	4,14	1,81E-06
<i>TNFAIP2</i>	2286,5	1114	2,05	5,73E-13	<i>TMEM176A</i>	619,5	302,5	2,05	3,33E-06
<i>CD36</i>	2778	1155	2,41	2,74E-12	<i>BASPI</i>	2885	1120	2,58	3,68E-06
<i>CCNA1</i>	1382,5	476,5	2,9	7,85E-12	<i>MPO</i>	6784,5	15838	0,43	4,05E-06
<i>TARP</i>	4965,5	2317,5	2,14	1,03E-11	<i>CPA3</i>	3423,5	1255,5	2,73	1,83E-05
<i>PPBP</i>	1487,5	332	4,48	1,08E-11	<i>MYCN</i>	839	390,5	2,15	2,42E-05
<i>EREG</i>	1391,5	255	5,46	1,39E-11	<i>MYOF</i>	736,5	303,5	2,43	3,17E-05
<i>EMPI</i>	433	1063	0,41	2,96E-11	<i>IFI30</i>	4928	1872,5	2,63	3,24E-05
<i>SPINK2</i>	2270	589,5	3,85	3,75E-11	<i>CAI</i>	764,5	1800	0,42	2,42E-04
<i>CX3CRI</i>	2901,5	893	3,25	5,75E-11	<i>FCNI</i>	2595,5	869	2,99	4,39E-04
<i>MARCKS</i>	1786,5	635,5	2,81	9,32E-11	<i>FGL2</i>	2020	893	2,26	7,20E-04
<i>TREMI</i>	1000,5	447	2,24	1,19E-10	<i>FPRI</i>	1097	478,5	2,29	9,26E-04
<i>BCL2A1</i>	993	446	2,23	1,35E-09	<i>C5ARI</i>	1231,5	609	2,02	1,48E-03
<i>WASF1</i>	452	911,5	0,5	2,60E-09	<i>ELANE</i>	2086,5	4984	0,42	2,26E-03
<i>PTX3</i>	766	368,5	2,08	2,63E-09	<i>CD14</i>	1211	359	3,37	5,38E-03
<i>MAFB</i>	1597,5	385,5	4,14	6,14E-09	<i>SI00A12</i>	765	358	2,14	2,23E-02

## B.

Gének	HR	P-érték	q-érték				
				<i>PBX3</i>	0,69	4,76E-03	1,50E-02
<i>MPO</i>	2,17	2,85E-07	2,42E-05	<i>BAALC</i>	0,69	7,04E-03	2,14E-02
<i>HOXA5</i>	0,55	1,15E-05	4,41E-04	<i>IFI30</i>	0,68	7,80E-03	2,24E-02
<i>HOXA10</i>	0,54	1,56E-05	4,41E-04	<i>CPVL</i>	0,71	8,09E-03	2,24E-02
<i>CD34</i>	0,55	2,78E-05	5,71E-04	<i>VNNI</i>	0,69	8,18E-03	2,24E-02
<i>TARP</i>	0,61	3,36E-05	5,71E-04	<i>CD14</i>	0,71	8,83E-03	2,34E-02
<i>SPINK2</i>	0,63	6,59E-05	9,34E-04	<i>HOXB5</i>	0,73	9,86E-03	2,54E-02
<i>MYOF</i>	0,62	2,27E-04	2,76E-03	<i>LGALS3BP</i>	0,72	1,13E-02	2,81E-02
<i>MEIS1</i>	0,59	3,12E-04	3,31E-03	<i>TNFAIP2</i>	0,72	1,21E-02	2,88E-02
<i>SELIL3</i>	0,61	3,63E-04	3,43E-03	<i>SLC38A1</i>	0,74	1,22E-02	2,88E-02
<i>PRKAR2B</i>	0,66	5,22E-04	4,44E-03	<i>CD200</i>	0,73	1,38E-02	3,16E-02
<i>H2AFY2</i>	0,67	8,56E-04	6,62E-03	<i>GYPC</i>	1,34	1,41E-02	3,16E-02
<i>PRKD3</i>	0,66	1,10E-03	7,81E-03	<i>MYCN</i>	0,73	1,48E-02	3,23E-02
<i>PPBP</i>	0,68	1,35E-03	8,85E-03	<i>COL4A5</i>	0,75	1,54E-02	3,27E-02
<i>MEST</i>	1,53	2,10E-03	1,25E-02	<i>HOXB6</i>	0,76	1,75E-02	3,59E-02
<i>PF4</i>	0,68	2,21E-03	1,25E-02	<i>FPR1</i>	0,72	1,77E-02	3,59E-02
<i>SMC4</i>	0,7	2,75E-03	1,25E-02	<i>RASGRP3</i>	0,76	1,90E-02	3,75E-02
<i>PLA2G4A</i>	0,7	2,81E-03	1,25E-02	<i>EREG</i>	0,76	2,12E-02	4,10E-02
<i>ELANE</i>	1,54	2,91E-03	1,25E-02	<i>MAFB</i>	0,73	2,22E-02	4,19E-02
<i>BASPI</i>	0,66	2,94E-03	1,25E-02	<i>EMP1</i>	0,73	2,61E-02	4,83E-02
<i>MARCKS</i>	0,69	3,31E-03	1,25E-02	<i>HOXB3</i>	0,77	2,71E-02	4,90E-02
<i>LILRB2</i>	0,66	3,34E-03	1,25E-02	<i>CTSG</i>	0,76	3,22E-02	5,71E-02
<i>HIF0</i>	0,68	3,36E-03	1,25E-02	<i>CYTL1</i>	1,35	3,33E-02	5,78E-02
<i>JUP</i>	1,5	3,38E-03	1,25E-02	<i>HOXB2</i>	0,77	4,19E-02	7,02E-02
<i>TSPAN13</i>	0,69	3,83E-03	1,36E-02	<i>EGFL7</i>	0,76	4,21E-02	7,02E-02
<i>FCNI</i>	0,71	4,58E-03	1,50E-02	<i>IFITM1</i>	0,77	4,36E-02	7,09E-02
<i>ITM2A</i>	1,46	4,65E-03	1,50E-02	<i>MANIA1</i>	1,28	4,42E-02	7,09E-02

A második (TCGA) és a harmadik (GSE1159) tanulókör adatait felhasználva a korábban azonosított 85 génből 49 gén esetén azonosítottam szignifikáns géneexpressziós eltérést az NPM1 mutáns és a vad betegcsoport között (**7. táblázat**).

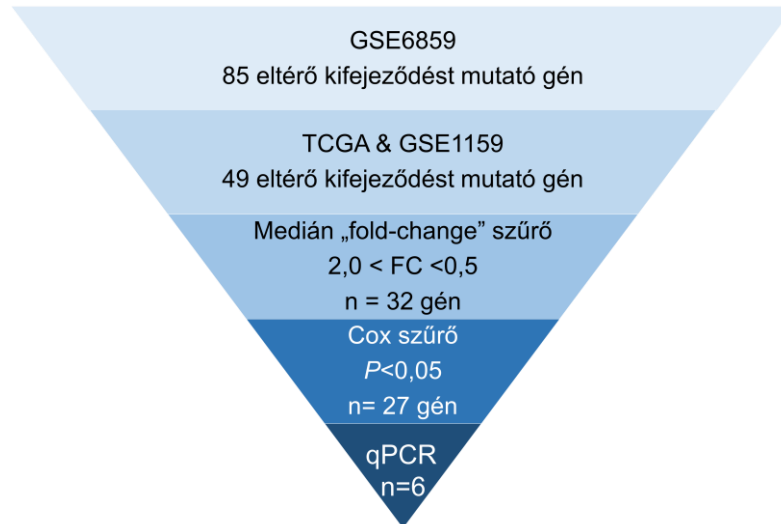
**7. táblázat.** Az NPM1 mutáns és vad betegcsoport között eltérő kifejeződést mutató gének a TCGA és a GSE1159 betegcsoport adatai alapján.

Gének	TCGA				GSE1159			
	Mutáns medián	Vad medián	FC	P-érték	Mutáns medián	Vad medián	FC	P-érték
<i>BAALC</i>	41,5	1010	0,04	4,75E-06	105	527	0,2	1,20E-14
<i>HOXA5</i>	1651,5	175,5	9,41	1,15E-05	3320	167,5	19,82	3,79E-26
<i>CD34</i>	89	9587	0,01	1,18E-05	310	1862	0,17	1,02E-13
<i>GYPC</i>	752,5	2596,5	0,29	1,26E-05	814	2218,5	0,37	7,62E-12
<i>HOXB3</i>	6453	729	8,85	1,54E-05	395	93	4,25	1,19E-25
<i>HOXB5</i>	426,5	5	85,3	2,75E-05	687	245,5	2,8	1,53E-23
<i>HOXB6</i>	714	7,5	95,2	3,61E-05	952	14,5	65,66	1,06E-22
<i>RASGRP3</i>	2853	693,5	4,11	5,14E-05	743	197	3,77	1,93E-11
<i>MANIA1</i>	1577,5	4319,5	0,37	5,91E-05	1025	2469,5	0,42	7,51E-11
<i>PBX3</i>	3952	895,5	4,41	6,29E-05	3406	647	5,26	1,58E-22
<i>HOXB2</i>	750,5	199	3,77	6,48E-05	2268	245	9,26	2,71E-23
<i>CD200</i>	37	869	0,04	7,68E-05	69	538	0,13	1,63E-16
<i>PDGFD</i>	377	85,5	4,41	1,10E-04	573	205,5	2,79	2,23E-18
<i>COL4A5</i>	1769	54	32,76	1,26E-04	1161	99	11,73	1,03E-17
<i>PROM1</i>	118	3421	0,03	1,26E-04	288	1468	0,2	2,68E-05
<i>HOXA10</i>	1164,5	318,5	3,66	1,46E-04	1842	304,5	6,05	8,81E-23
<i>DMXL2</i>	9338	4220	2,21	1,51E-04	3644	1164,5	3,13	1,03E-17
<i>MEIS1</i>	4178	1235	3,38	1,96E-04	1761	352,5	5	4,04E-21
<i>SMC4</i>	5938,5	3471	1,71	2,14E-04	3502	1565,5	2,24	5,65E-20
<i>NPR3</i>	561,5	3175	0,18	3,67E-04	440	1493	0,29	7,34E-06



<i>ITM2C</i>	2335	3929	0,59	4,83E-04	712	2538	0,28	3,27E-17
<i>MEST</i>	678	1710	0,4	1,27E-03	948	2877	0,33	1,20E-12
<i>BAHCCI</i>	14302	5990	2,39	1,49E-03	2543	1273	2	1,15E-08
<i>TSPAN13</i>	133,5	405	0,33	2,20E-03	252	732	0,34	4,82E-11
<i>TMEM176B</i>	28,5	105,5	0,27	2,90E-03	651	170	3,83	2,21E-03
<i>TMEM176A</i>	17	65,5	0,26	3,04E-03	831	435,5	1,91	8,10E-04
<i>JUP</i>	2023	4307	0,47	3,22E-03	510	1762,5	0,29	1,40E-14
<i>APP</i>	230	4225,5	0,05	4,07E-03	43	335,5	0,13	1,90E-14
<i>PTX3</i>	177	99	1,79	5,66E-03	722	286	2,52	3,17E-07
<i>PLA2G4A</i>	845,5	542,5	1,56	7,47E-03	400	187,5	2,13	3,73E-15
<i>CTSG</i>	3846,5	891	4,32	7,55E-03	3909	837	4,67	1,05E-08
<i>IFITM1</i>	208,5	405	0,51	8,51E-03	1295	2301	0,56	1,06E-08
<i>LPAR6</i>	330	649	0,51	8,60E-03	220	805	0,27	1,24E-11
<i>CCND2</i>	4057,5	6980	0,58	8,98E-03	2137	5490,5	0,39	3,31E-16
<i>SEL1L3</i>	2942,5	1823,5	1,61	1,41E-02	1650	791	2,09	2,16E-09
<i>ITM2A</i>	730,5	2173	0,34	1,45E-02	647	1967	0,33	1,75E-11
<i>SLC38A1</i>	2730,5	5749,5	0,47	1,70E-02	831	1893,5	0,44	4,41E-10
<i>EMPI</i>	478	698	0,68	1,93E-02	281	906,5	0,31	3,13E-09
<i>EGFL7</i>	799	1628	0,49	2,28E-02	376	965	0,39	1,51E-09
<i>JAG1</i>	1032	701,5	1,47	2,56E-02	888	403	2,2	8,53E-08
<i>CCNA1</i>	866,5	392,5	2,21	2,60E-02	1514	583,5	2,59	2,04E-05
<i>ELANE</i>	2815	1644	1,71	3,66E-02	2466	5811	0,42	1,27E-02
<i>TREM1</i>	1238	565,5	2,19	4,07E-02	1158	597,5	1,94	2,99E-07
<i>TNFAIP2</i>	5289	3448	1,53	4,25E-02	2196	1215,5	1,81	2,72E-08
<i>SLC4A1</i>	255	1005,5	0,25	4,29E-02	284	794	0,36	4,45E-04
<i>PRKD3</i>	898	1510,5	0,59	4,33E-02	269	558,5	0,48	6,82E-08
<i>LGALS3BP</i>	5023,5	1190	4,22	4,60E-02	2623	996,5	2,63	2,70E-09
<i>TARP</i>	1053,5	503	2,09	4,72E-02	5095	2815,5	1,81	3,96E-05
<i>HBB</i>	3253	11122,5	0,29	4,72E-02	4514	21338,5	0,21	5,18E-05

A klinikai mintákon (Simmelweis mintacsoport) végzett kvantitatív-PCR vizsgálathoz a gének kiválasztása a **14. ábra** alapján történt.



**14. ábra.** A gének kiválasztása a qPCR validációhoz.

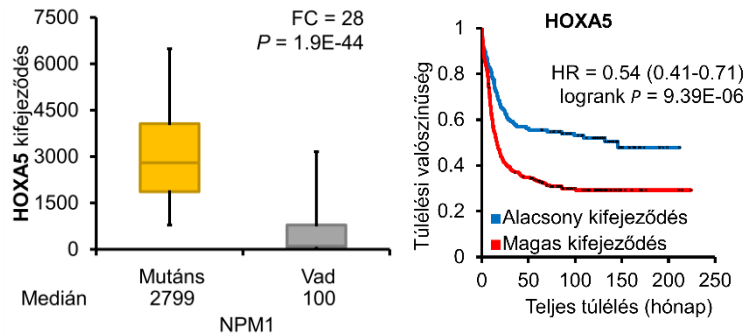
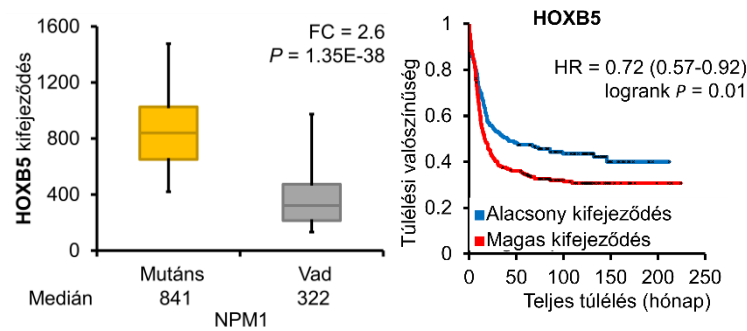
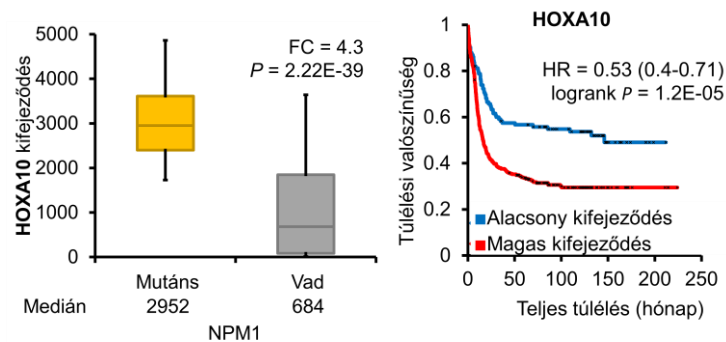
Elsőként a TCGA és a GSE1159-es mintacsoportok alapján, a 49 eltérő expressziót mutató génből 32 gént szűrtem ki, amelyek esetében az NPM1 mutáns és a vad betegcsoportok medián kifejeződési értékeinek a hányadosai nagyobbak voltak, mint 2, vagy kisebbek voltak, mint 0,5 (**14. ábra**). Továbbá ezen gének közül kiválasztottam azt a 27 gént, amelyeknél a génkifejeződés és a teljes túlélés között a Cox regresszió során szignifikáns összefüggést tapasztaltam (**8. táblázat**). A TCGA mintacsoport esetén nem végeztünk túlélés elemzést azért, mert kevesebb, mint 100 beteghez tartozott túlélési adat.

**8. táblázat.** A teljes túléléssel összefüggő gének a GSE1159-es mintacsoport adatai alapján.

Gének	HR	P-érték	q-érték		HR	P-érték	q-érték
				<i>MEIS1</i>	0,49	2,85E-04	2,76E-03
<i>HOXA10</i>	0,48	1,63E-05	7,99E-04	<i>ITM2A</i>	1,96	3,38E-04	2,76E-03
<i>TARP</i>	0,53	1,31E-04	2,22E-03	<i>PLA2G4A</i>	0,59	1,19E-03	8,33E-03
<i>HOXA5</i>	0,51	1,69E-04	2,22E-03	<i>ELANE</i>	1,8	1,39E-03	8,51E-03
<i>SELIL3</i>	0,53	1,81E-04	2,22E-03	<i>MEST</i>	1,77	2,06E-03	1,12E-02

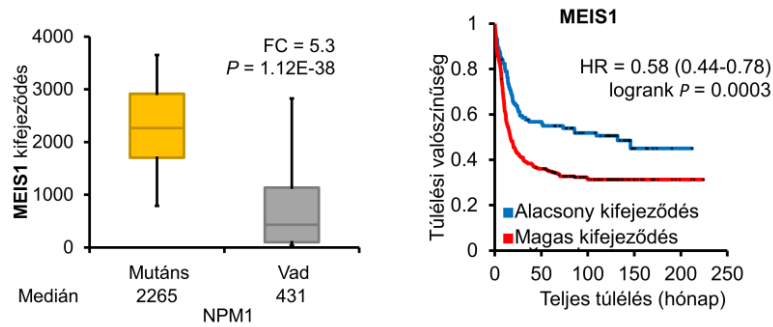
<i>CD34</i>	0,58	2,48E-03	1,22E-02	<i>TSPAN13</i>	0,65	1,03E-02	2,66E-02
<i>JUP</i>	1,76	3,73E-03	1,58E-02	<i>SLC38A1</i>	0,67	1,26E-02	3,09E-02
<i>GYPC</i>	1,57	3,86E-03	1,58E-02	<i>IFITM1</i>	1,49	1,49E-02	3,48E-02
<i>LGALS3BP</i>	0,62	4,49E-03	1,67E-02	<i>HOXB2</i>	0,68	1,98E-02	4,41E-02
<i>SMC4</i>	0,62	4,76E-03	1,67E-02	<i>RASGRP3</i>	0,71	2,79E-02	5,94E-02
<i>MAN1A1</i>	1,54	5,38E-03	1,76E-02	<i>CCND2</i>	1,38	4,21E-02	8,24E-02
<i>PBX3</i>	0,65	6,02E-03	1,84E-02	<i>LPAR6</i>	1,41	4,42E-02	8,24E-02
<i>HOXB5</i>	0,63	7,11E-03	2,05E-02	<i>HOXB3</i>	0,72	4,45E-02	8,24E-02
<i>CTSG</i>	0,63	8,96E-03	2,44E-02	<i>EGFL7</i>	1,46	4,54E-02	8,24E-02

Ezt követően a klinikai mintákon történő validációra a 27 génből azt a hat gént választottuk ki, amelyek mind a túlélés elemzés során, mind pedig a Mann-Whitney elemzés során a legszignifikánsabb összefüggést mutatták. Ezek a gének a HOXA5, HOXB5, HOXA10, PBX3, MEIS1 és az ITM2A gének voltak. A **15. ábra** és a **16. ábra** mutatja a kvantitatív-PCR validációra kiválasztott hat gén kifejeződését az NPM1 mutáns és vad típusú mintákban, valamint a teljes túléléssel való összefüggését a GSE6891-es betegcsoport adatai alapján. Ezen gének között az ITM2A volt az egyedüli gén, amelynek a kifejeződése csökkent NPM1 mutáns tumorokban. Kaplan-Meier analízis alapján a gének magas kifejeződése rosszabb túléléssel társult, szemben az ITM2A génnel, amely esetében a betegek rossz túlélésével az alacsony génkifejeződés korrelált (**16. ábra C**).

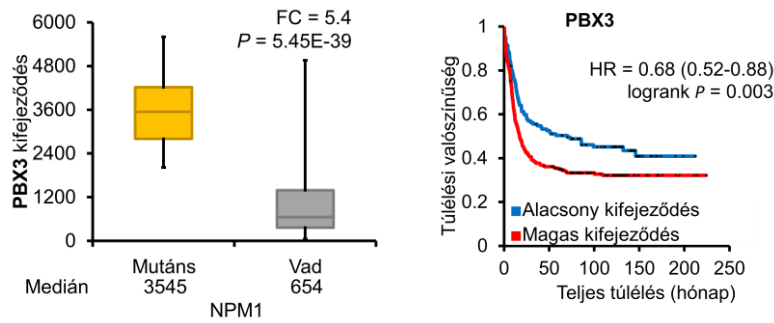
**A.****B.****C.**

**15. ábra. A-C.** A HOXA5, HOXB5 és HOXA10 gének kifejeződése az NPM1 mutáns és vad típusú mintákban és a teljes túléléssel való összefüggésük a GSE6891-es betegcsoport adatai alapján.

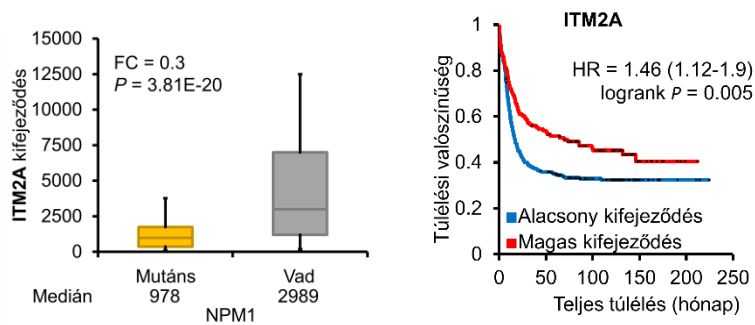
A.



B.



C.

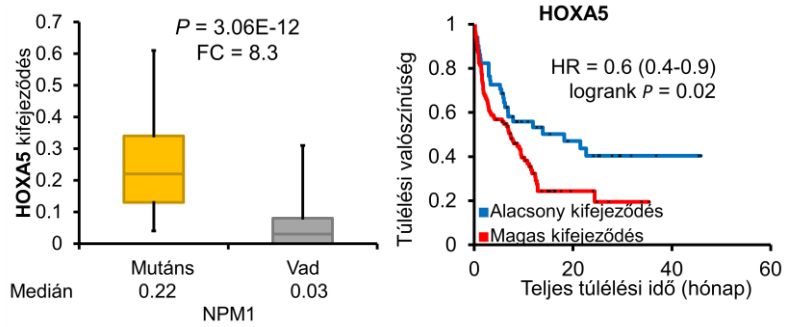


**16. ábra. A-C.** A MEIS1, PBX3 és az ITM2A gének kifejeződése az NPM1 mutáns és vad típusú mintákban és a teljes túléléssel való összefüggésük a GSE6891-es betegcsoport adatai alapján.

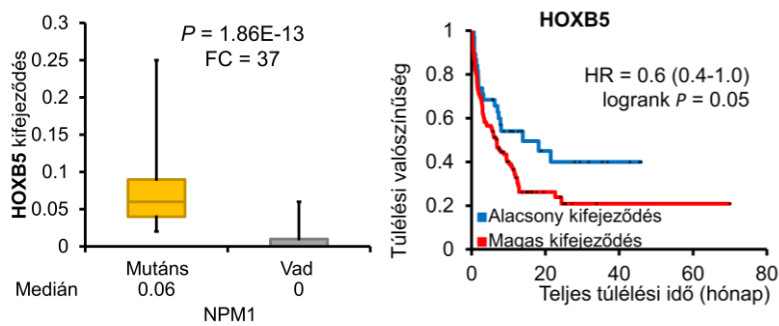
### **5.2.3 Az NPM1 mutációk transzkriptomikus lenyomatának igazolása klinikai mintákon**

A kiválasztott hat gén expressziós eltéréseinek igazolása az NPM1 mutáns és a vad akut mieloid leukémia betegcsoport között, kvantitatív-PCR reakcióval történt. A vizsgálat során a HOXA5, HOXA10, HOXB5, MEIS1 és a PBX3 gének kifejeződése nőtt (**17. ábra és 18. ábra**), míg az ITM2A gén kifejeződése szignifikánsan csökkent az NPM1 mutáns daganatokban (**18. ábra C**). Túlélés elemzés során igazoltuk a kiválasztott gének kifejeződése és teljes túlélés közötti szignifikáns kapcsolatot a klinikai mintacsoport felhasználásával (**17. ábra és 18. ábra**).

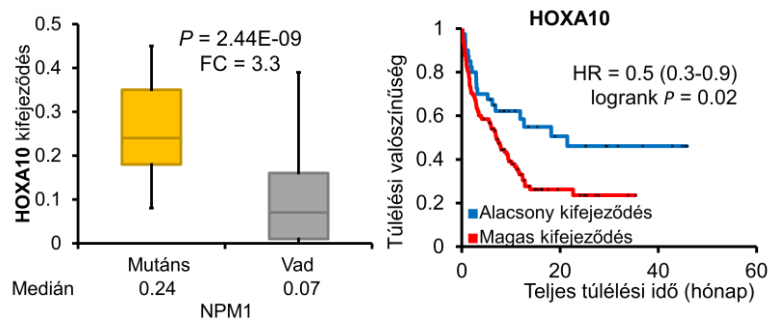
A.



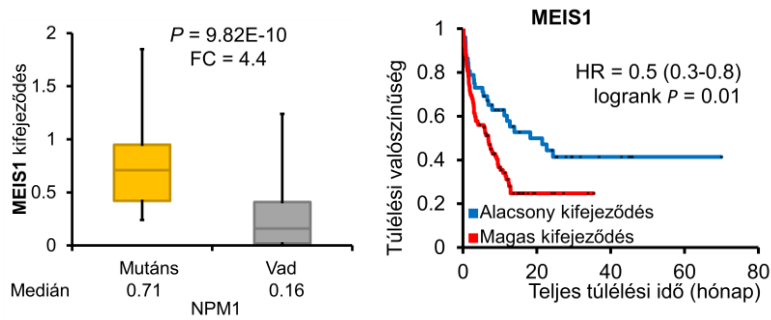
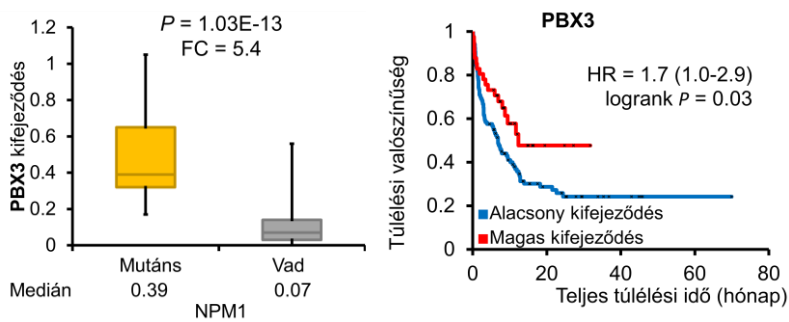
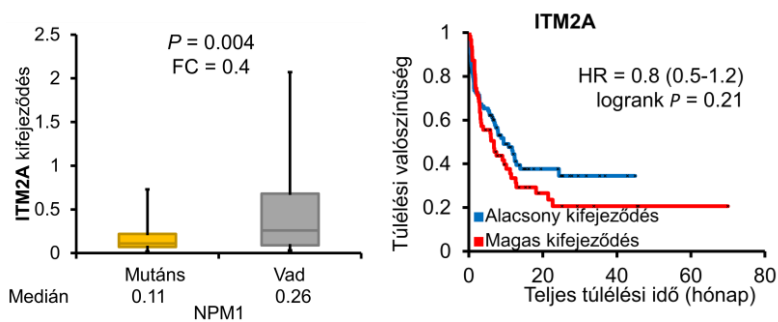
B.



C.



**17. ábra. A-C.** A HOXA5, HOXB5 és HOXA10 gének kifejeződésének az igazolása az NPM1 mutáns daganatokban és hatásuk a teljes túlélésre a Semmelweis mintacsoport alapján.

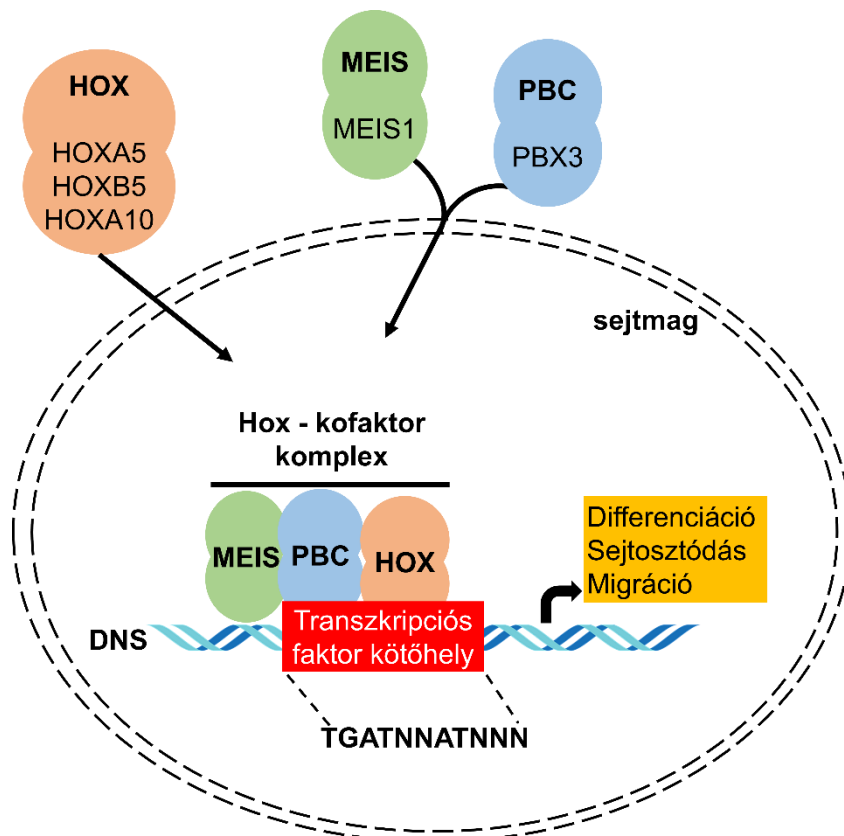
**A.****B.****C.**

**18. ábra. A-C.** A MEIS1, PBX3 és az ITM2A gének kifejeződésének az igazolása az NPM1 mutáns daganatokban és hatásuk a teljes túlélésre a Semmelweis mintacsoport alapján.



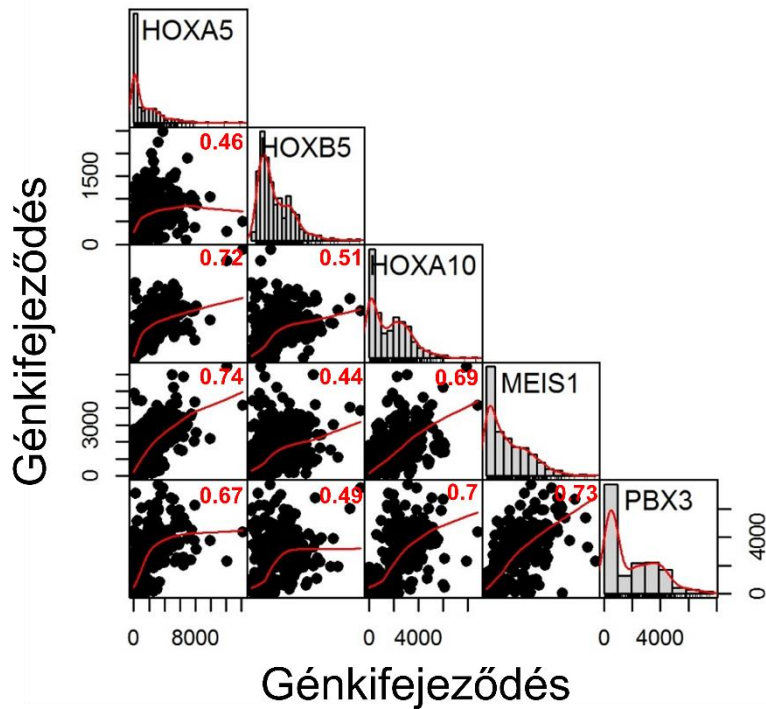
#### 5.2.4 A HOX gének és a kofaktoraik összefüggése

Az elemzések során részben olyan géneket azonosítottam, amelyek a HOX géncsalád egyes génjei esetén kofaktor szerepet töltenek be (182). Ezen gének voltak a MEIS és a PBX géncsalád tagjai, amelyeknek a HOX génekkel történő sejten belüli összefüggést a **19. ábra** szemléltet.



**19. ábra.** A HOX gének és kofaktoraik sejten belüli működése.

A GSE6891-es betegcsoport adatait felhasználva Pearson-féle rang korrelációval megvizsgáltam, hogy az egyes HOX gének és a kofaktorok génkifejeződése hogyan korrelálnak egymással. Az elemzés során az összes esetben a  $P$  érték,  $2,2E-16$ -tól kisebb volt. A legnagyobb korrelációt a HOXA5 és HOXA10, a HOXA5 és MEIS1, a HOXA10 és MEIS1, a HOXA10 és PBX3, valamint a MEIS1 és PBX3 gének között tapasztaltam (**20. ábra**).



**20. ábra.** A HOX gének és a kofaktoraik kifejeződésének korrelációja.

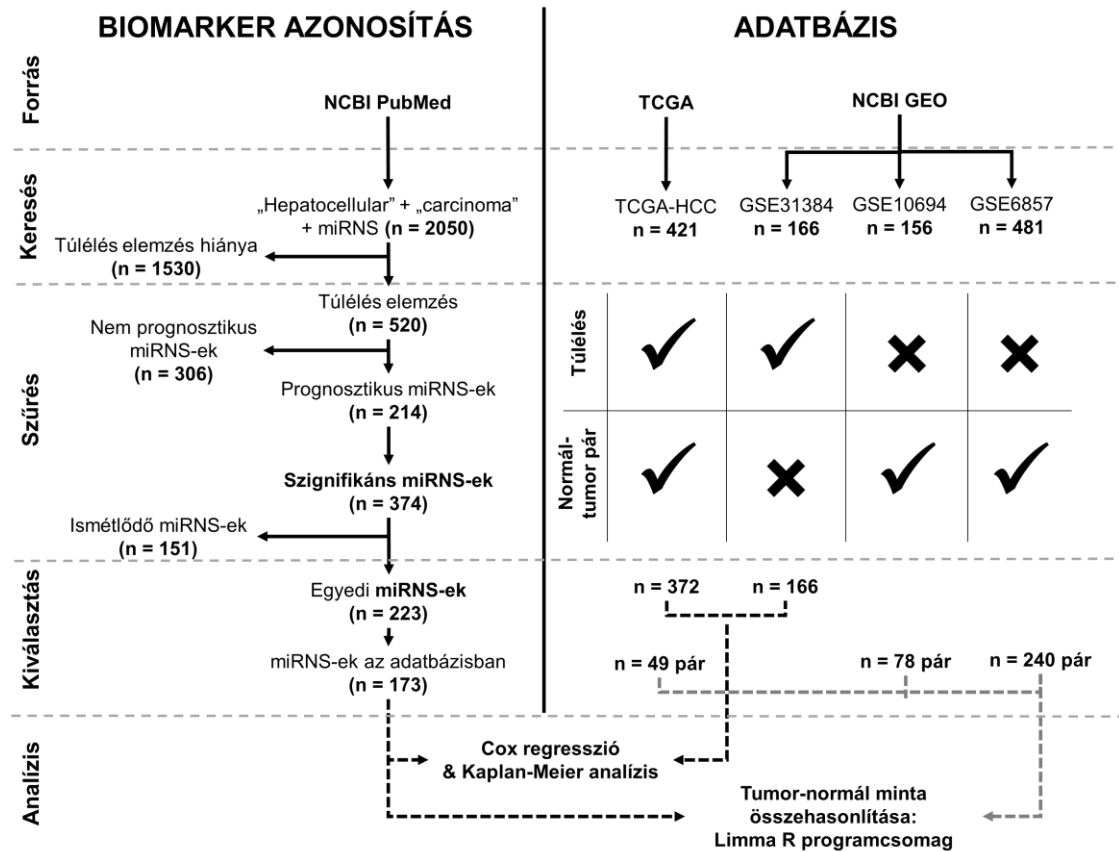
### 5.3 Prognosztikus biomarkerek azonosítása máj hepatocelluláris karcinómában

#### 5.3.1 Adatbázis felépítése

A máj hepatocelluláris karcinóma esetén mRNA és mikro-RNS expressziós adatokat használtunk fel. Mind a két adattípus adatai az NCBI GEO adatbázisból és a TCGA programból származtak.

##### 5.3.1.1 mikro-RNS expressziós adatbázis

A mikro-RNS expressziós adatbázis összesen négy mintacsoport transzkriptomikus adatát tartalmazta: az NCBI GEO adatbázisból a GSE31384 (100), a GSE10694 (171) és GSE6857 (172) vizsgálatok microarray mikro-RNS expressziós adatát, valamint a TCGA programból a mikro-RNS szekvenálási adatokat tartalmazta (**21. ábra**).



**21. ábra.** Prognosztikus mikro-RNS-ek azonosítása irodalmi keresés alapján és a máj mikro-RNS adatbázis felépítése.

A GSE6857-es vizsgálat összesen 241 tumoros mintát tartalmazott, amelyből 240 mintához tartozott normál szöveti minta is (**21. ábra**). A GSE10694-es tanulmány, 78 tumoros és normál szöveti mintapárt tartalmazott, a GSE31384-es vizsgálat 166 tumoros mintát tartalmazott, normál szöveti mintapárok nélkül (**21. ábra**). A TCGA program, 372 máj hepatocelluláris karcinómás beteg mikro-RNS szekvenálási adatát tartalmazta, amelyekből mintegy 49 pácienshez tartozott normál szöveti minta pár is (**21. ábra**).

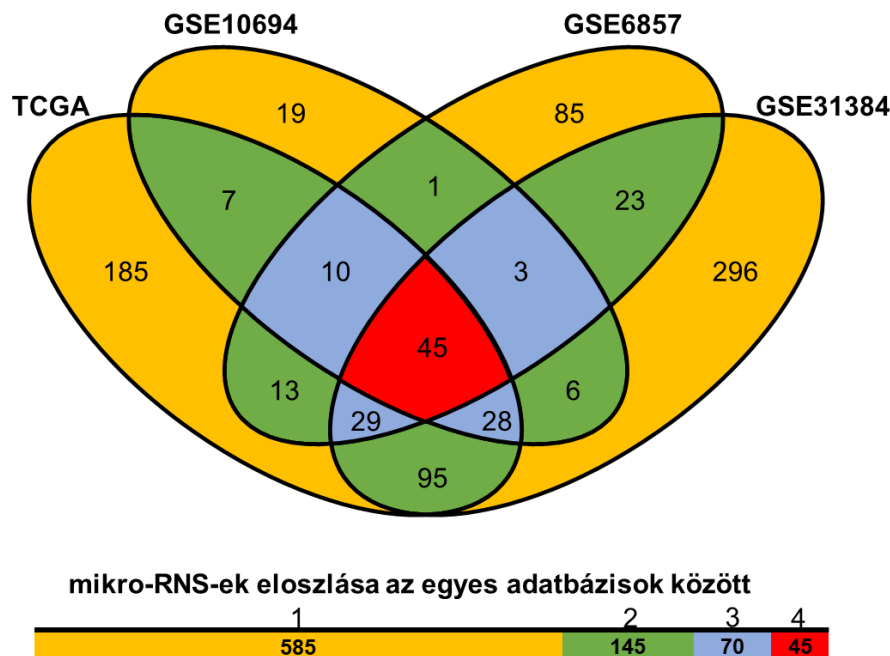
A túlélési adatok tekintetében a mintákhoz teljes túlélési, valamint a betegség-mentes túlélési adatok voltak elérhetőek. A TCGA program betegei esetén teljes túlélési adat 372, a GSE31384 adatoknál 166 és a GSE10694 adatoknál pedig 77 beteghez voltak elérhetőek. A teljes máj mikro-RNS adatbázisban a medián teljes túlélési idő 20,2 hónap volt, az egyes vizsgálatokra lebontva pedig a következő: TCGA – 19,6 hónap, GSE31384 – 34,0 hónap és a GSE10694 – 15,8 hónap. Betegség-mentes túlélési adat csak a GSE31384-es tanulmányhoz tartozott, ahol a 166 beteg esetén a medián betegség-mentes

túlélési idő 17,5 hónap volt. A GSE10694-es adatok közül csak 54 páciens rendelkezett betegség-mentes túlélési adattal, így ezen beteghalmazt a kis mintaszám miatt kihagytam a túlélés elemzésből.

További klinikai adatok, úgymint a daganat stádiuma a TCGA és a GSE10694 vizsgálatokhoz, áttétképzésre vonatkozó adat a GSE6857 és a GSE10694 vizsgálatokhoz, továbbá a páciensek neme a TCGA, a GSE6857 és a GSE10694 adatokhoz voltak elérhetőek. Túlélési elemzés során csak a daganat stádiuma és a teljes túlélés között figyeltem meg szignifikáns összefüggést a TCGA adatbázis esetén ( $P = 3,01E-06$ ).

A TCGA program betegeihez továbbá rendelkezésre állt etnikai hovatartozásra irányuló adat is. A páciensek között 182 kaukázusi, 17 afro-amerikai, 161 ázsiai és 2 amerikai őslakos beteg volt, továbbá 10 páciensnek ismeretlen volt az etnikai besorolása. Az adatbázist többségében ázsiai és kaukázusi betegek alkotják, és megfigyeltük, hogy a két csoport között teljes túlélés függvényében nincs különbség ( $P = 0,32$ ).

A teljes máj hepatocelluláris karcinóma mikro-RNS adatbázis 845 mikro-RNS-t tartalmaz, amelyből 585 mikro-RNS kifejeződését csak egyetlen vizsgálatban, míg 45 mikro-RNS kifejeződését pedig minden vizsgálatban mérték (22. ábra).



**22. ábra.** A mikro-RNS-ek előfordulása az egyes vizsgálati csoportok között.

### 5.3.1.2 mRNS expressziós adatok

mRNS expressziós adatok a máj hepatocelluláris karcinóma esetén szintén az NCBI GEO adatbázisból, valamint a TCGA programból származtak. Az adatbázis, a TCGA program összesen 372 máj hepatocelluláris karcinómás beteg RNS szekvenálási adatát tartalmazza. Ezen adatok klinikai-patológiai jellemzésére nem térnek ki, mivel ezek a betegek mikro-RNS kifejeződési adattal is rendelkeznek, és azokat az előző részben már részletesen ismertettem. Az adatbázis TCGA program betegmintáin kívül további két vizsgálati csoport adatát, a GSE9843 (183) 91 páciensét és a GSE20017 (103) 135 betegének microarray génexpressziós adatát tartalmazta. Mind a két vizsgálati csoport számos klinikai-patológia jellemzővel rendelkezett, úgymint a betegek neme, etnikai besorolása, vaszkuláris invázió és a teljes túlélés (**9. táblázat**).

**9. táblázat.** Az NCBI GEO adatbázisból származó betegcsoportok klinikai-patológiai jellemzői.

	<b>GSE20017</b>	<b>GSE9843</b>
Betegminták száma	135	91
Nem (férfi/nő/nincs adat)	102/33/-	54/27/10
Daganat stádiuma (I./II./III./IV./nincs adat)	-/-/-/135	9/51/7/8/11
Etnikum (Kaukázusi/Afro-Amerikai/Ázsiai)	102/4/28	72/3/4
Vaszkuláris inváziót mutat	40	45
Halálozási események száma	32	23

### 5.3.2 Prognosztikus mikro-RNS-ek vizsgálata máj hepatocelluláris karcinómában

#### 5.3.2.1 Korábban publikált prognosztikus mikro-RNS-ek

Az NCBI PubMed adatbázisban a „hepatocellular”, a „carcinoma” és a „miRNA” keresőszavak alkalmazásával, 214 közleményt sikerült azonosítanom, amelyek prognosztikus mikro-RNS-eket írtak le máj hepatocelluláris karcinómában (**21. ábra**).

A publikációk összesen 374 prognosztikus mikro-RNS-t írtak le máj hepatocelluláris karcinómában, amelyek között 151 mikro-RNS-t egyszerre több közlemény is tartalmazott. Az ismétlődő mikro-RNS-ek kiszűrését követően 223 egyedi mikro-RNS-t kaptam, amelyekből 173 mikro-RNS-t az általam létrehozott mikro-RNS adatbázisban (**21. ábra**).

### 5.3.2.2 Prognosztikus mikro-RNS-ek validálása

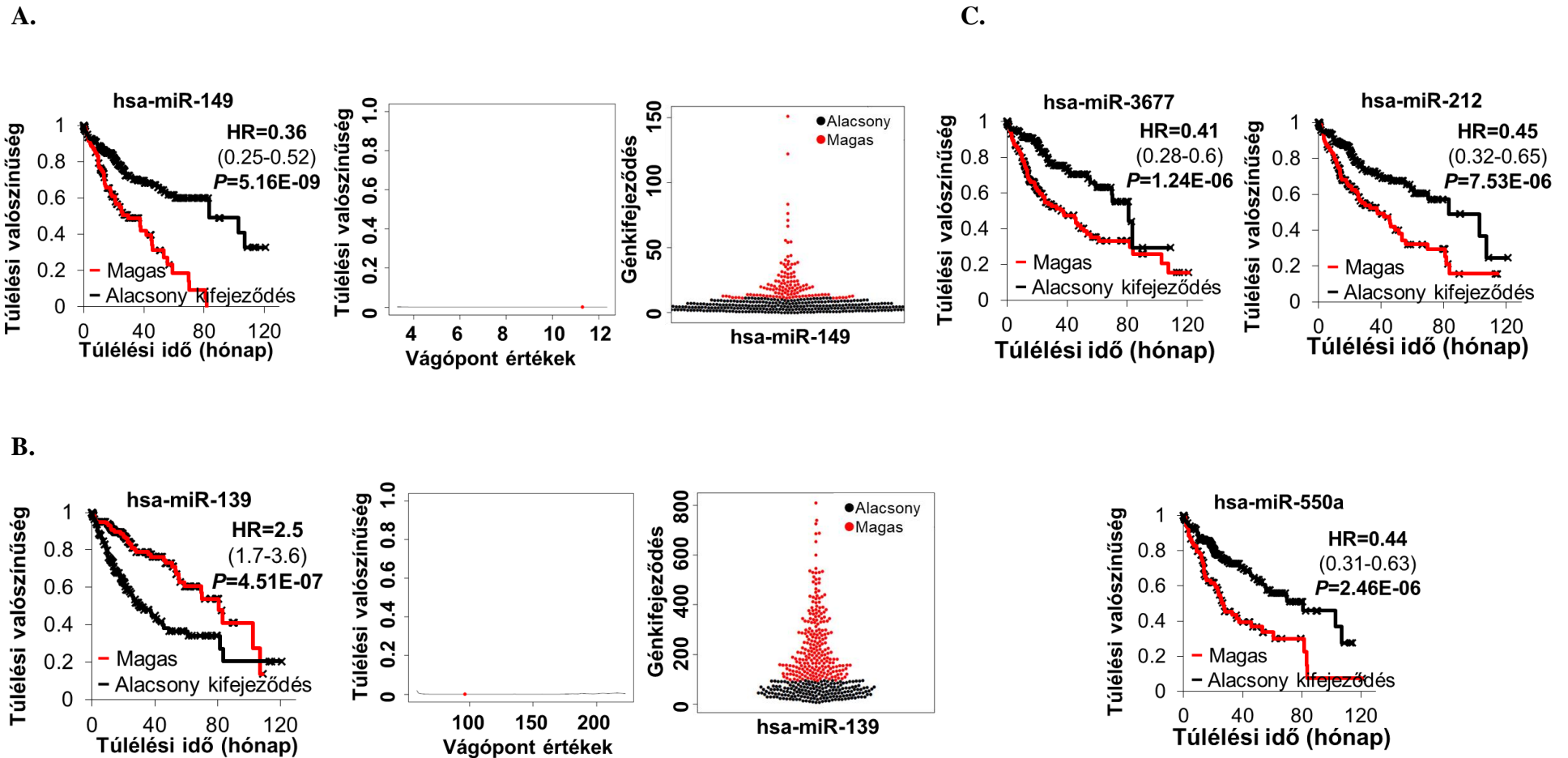
A 173 mikro-RNS alapú prognosztikus biomarkerből 55 mikro-RNS esetében tapasztaltam szignifikáns összefüggést a génexpresszió és a teljes túlélés között a TCGA adatok alapján (**10. táblázat**).

**10. táblázat.** Az univariáns és multivariáns túlélés elemzés során szignifikáns mikro-RNS-ek listája a TCGA program adatai alapján.

mikro-RNS	HR	<i>P</i> -érték (univ.)	<i>P</i> -érték (multiv.)				
				hsa-miR-9	0,54	6,00E-04	2,52E-03
hsa-miR-149	0,36	5,19E-09	2,24E-06	hsa-miR-15b	0,55	6,00E-04	1,08E-02
hsa-miR-139	2,48	4,50E-07	6,02E-04	hsa-miR-25	0,54	6,00E-04	3,82E-03
hsa-miR-3677	0,41	1,24E-06	2,39E-04	hsa-miR-34a	1,78	1,10E-03	3,27E-03
hsa-miR-550a	0,44	2,45E-06	5,95E-04	hsa-miR-126	1,76	1,20E-03	8,52E-02
hsa-miR-212	0,45	7,53E-06	1,92E-04	hsa-miR-132	0,58	1,60E-03	7,67E-03
hsa-miR-421	0,48	2,66E-05	5,20E-03	hsa-miR-301b	0,56	1,80E-03	6,71E-04
hsa-miR-3607	2,09	4,29E-05	5,42E-03	hsa-miR-24	0,57	2,10E-03	1,09E-02
hsa-miR-326	0,5	1,00E-04	3,04E-04	hsa-miR-148a	1,74	2,30E-03	1,12E-02
hsa-miR-940	0,49	1,00E-04	2,86E-04	hsa-miR-194	1,72	2,50E-03	3,67E-02
hsa-miR-100	1,96	2,00E-04	3,74E-03	hsa-miR-222	0,54	2,80E-03	6,00E-03
hsa-miR-215	0,53	2,00E-04	1,03E-02	hsa-miR-101	1,68	3,30E-03	3,88E-02
hsa-miR-454	0,52	2,00E-04	6,16E-03	hsa-miR-125b	1,85	4,30E-03	1,65E-02
hsa-miR-301a	0,53	3,00E-04	1,52E-03	hsa-miR-744	0,58	5,00E-03	7,20E-04
hsa-miR-330	0,52	3,00E-04	4,55E-03	hsa-miR-141	0,61	5,70E-03	3,87E-02
hsa-miR-22	1,92	3,00E-04	8,52E-04	hsa-miR-33a	0,6	5,70E-03	5,29E-03
hsa-miR-491	0,54	5,00E-04	4,67E-03	hsa-miR-33b	0,6	6,20E-03	1,34E-02

hsa-miR-99a	1,64	6,20E-03	5,84E-02	hsa-miR-106b	0,59	1,59E-02	6,20E-02
hsa-miR-21	0,61	6,80E-03	7,93E-02	hsa-miR-122	1,56	1,65E-02	2,83E-01
hsa-miR-20b	1,66	6,80E-03	9,26E-02	hsa-miR-221	0,66	1,72E-02	6,73E-03
hsa-miR-429	0,62	6,90E-03	1,33E-02	hsa-miR-183	0,67	2,60E-02	2,67E-01
hsa-miR-1	1,67	7,20E-03	6,78E-03	hsa-miR-204	1,47	3,09E-02	2,23E-01
hsa-miR-23a	0,61	7,60E-03	8,88E-02	hsa-miR-145	1,46	3,21E-02	3,44E-02
hsa-miR-186	0,63	9,90E-03	3,79E-02	hsa-miR-203a	1,58	3,57E-02	2,42E-01
hsa-miR-10b	0,63	1,15E-02	8,67E-03	hsa-miR-29c	1,45	4,13E-02	6,97E-02
hsa-miR-182	0,64	1,30E-02	3,32E-01	hsa-miR-223	0,68	4,27E-02	1,07E-01
hsa-miR-195	1,54	1,43E-02	6,35E-02	hsa-let-7g	1,44	4,55E-02	1,06E-01
hsa-miR-200c	0,65	1,58E-02	5,82E-02	hsa-miR-135a	0,7	4,71E-02	1,01E-02

Ezek közül az öt legszignifikánsabb mikro-RNS – a hsa-miR-149, hsa-miR-139, hsa-miR-3677, hsa-miR-550a és a hsa-miR-212 – Kaplan-Meier túlélési görbéit a **23. ábra** mutatja. A mikro-RNS-ek expressziós értékeinek alsó és felső kvartilis értékei között percentilisenként Cox regressziót végeztünk, ahol minden percentilis érték a túlélés elemzés vágópontjának felelt meg. Mind az öt mikro-RNS esetén elmondható, hogy bármely vágópontot használva szignifikáns összefüggést tapasztaltam a mikro-RNS kifejeződése és a teljes túlélés között (**23. ábra**).



23. ábra. Mikro-RNS-ek hatása a teljes túlélésre a TCGA adatok alapján.



Multivariáns túlélés elemzés során megvizsgáltam az 55 szignifikáns mikro-RNS génextpressziójának, a tumor stádiumának és a páciensek nemének együttes hatását a teljes túlélésre, amely során 39 mikro-RNS esetében tapasztaltam szignifikáns összefüggést a teljes túléléssel (**10. táblázat**).

A 173 mikro-RNS expressziójának a hatását a teljes túlélésre a GSE31384 microarray adatok felhasználásával is vizsgáltam. Fontos kiemelni, hogy mivel teljes túlélési és betegség-mentes túlélési adat is elérhető volt ezen vizsgálatok esetén, így az egyes mikro-RNS-ek túlélés elemzését a közleményben megjelölt túlélési idő felhasználásával végeztem el. A mikro-RNS-ek prognosztikus hatását uni- és multivariáns túlélés elemzéssel is vizsgáltam, ahol szignifikáns mikro-RNS-ek listája a **11. táblázat**-ban látható. A teljes túlélés esetén 84, míg a betegség-mentes túlélés esetén 25 mikro-RNS kifejeződése mutatott szignifikáns összefüggést. Továbbá elmondható, hogy összesen 29 mikro-RNS-t azonosítottunk, amelyek kifejeződései mind a TCGA és a GSE31384 adatbázis esetén szignifikáns összefüggést mutatnak a teljes túléléssel.

**11. táblázat.** Mikro-RNS expresszió hatása a teljes (**A**), valamint a betegség-mentes (**B**) túlélésre a GSE31384 vizsgálat adatai alapján.

**A)**

<b>mikro-RNS</b>	<b>HR</b>	<b>P-érték (univ.)</b>	<b>P-érték (multiv)</b>				
				hsa-miR-24	2,42	1,08E-04	9,37E-01
hsa-miR-146b-3p	3,05	7,76E-07	6,94E-02	hsa-miR-330-3p	0,41	1,15E-04	1,11E-01
hsa-miR-584	0,33	1,38E-06		hsa-miR-675	0,4	1,17E-04	
hsa-miR-612	0,37	1,47E-05	4,53E-01	hsa-miR-105	2,53	1,22E-04	5,41E-02
hsa-miR-28-5p	2,69	1,58E-05		hsa-miR-489	0,41	1,25E-04	
hsa-miR-139-5p	2,87	1,85E-05		hsa-miR-31	2,42	1,49E-04	4,21E-01
hsa-miR-486-3p	2,6	3,72E-05	4,41E-03	hsa-miR-106b	2,93	1,74E-04	
hsa-miR-29c	2,56	5,63E-05		hsa-miR-99a	0,42	1,76E-04	
hsa-miR-34b	3,46	6,04E-05		hsa-miR-15b	2,79	2,17E-04	3,20E-01
hsa-miR-496	3,33	6,23E-05	4,63E-02	hsa-miR-548b-3p	0,38	2,42E-04	
hsa-miR-126	3	9,76E-05		hsa-miR-185	2,3	2,88E-04	1,48E-01
				hsa-miR-302d	2,28	3,41E-04	7,06E-01

hsa-miR-221	0,44	3,75E-04	9,95E-01	hsa-miR-222	0,5	3,91E-03	7,81E-01
hsa-miR-30a	2,31	3,86E-04	3,35E-01	hsa-miR-132	1,99	4,01E-03	
hsa-miR-92a	0,44	4,41E-04	9,59E-01	hsa-miR-492	0,51	4,17E-03	
hsa-miR-301b	2,25	4,42E-04		hsa-miR-339-5p	0,52	4,78E-03	5,12E-01
hsa-miR-661	0,43	4,52E-04		hsa-miR-141	1,95	5,45E-03	
hsa-miR-940	0,44	4,60E-04	7,99E-01	hsa-miR-886-5p	1,92	5,84E-03	8,93E-01
hsa-miR-509-3p	0,41	6,98E-04		hsa-miR-183	1,89	6,13E-03	
hsa-miR-769-5p	0,45	7,54E-04	6,95E-02	hsa-miR-100	0,52	6,66E-03	1,93E-01
hsa-miR-625	0,36	7,84E-04		hsa-miR-622	0,54	8,53E-03	5,55E-01
hsa-miR-122	2,22	8,10E-04		hsa-miR-601	0,44	8,58E-03	5,15E-02
hsa-miR-182	2,18	8,57E-04		hsa-miR-26a	0,55	8,83E-03	2,42E-01
hsa-miR-137	0,46	8,69E-04		hsa-miR-326	0,55	9,79E-03	
hsa-miR-29a	0,33	1,00E-03	1,43E-02	hsa-miR-138	1,85	1,12E-02	1,04E-01
hsa-miR-181b	2,14	1,01E-03	2,48E-01	hsa-miR-203	0,54	1,18E-02	4,24E-01
hsa-miR-155	2,15	1,02E-03	3,68E-01	hsa-miR-216b	1,81	1,19E-02	7,96E-01
hsa-miR-503	0,46	1,02E-03	5,99E-01	hsa-miR-331-3p	0,56	1,19E-02	
hsa-miR-9	0,47	1,14E-03		hsa-miR-96	0,55	1,23E-02	
hsa-miR-608	0,47	1,24E-03	4,66E-01	hsa-miR-25	0,49	1,24E-02	1,62E-02
hsa-miR-421	2,1	1,54E-03		hsa-miR-610	0,56	1,27E-02	4,54E-01
hsa-miR-325	0,38	1,66E-03		hsa-miR-148a	0,55	1,45E-02	
hsa-miR-522	0,48	1,93E-03	1,04E-01	hsa-miR-186	0,55	1,57E-02	4,23E-01
hsa-miR-142-3p	2,1	2,03E-03		hsa-miR-451	1,7	2,41E-02	6,65E-01
hsa-miR-638	0,48	2,18E-03		hsa-miR-135a	1,72	2,57E-02	
hsa-miR-29b	2,05	2,26E-03		hsa-miR-21	1,79	2,97E-02	9,07E-01
hsa-miR-876-5p	0,48	2,30E-03	8,43E-01	hsa-miR-218	1,65	3,49E-02	
hsa-miR-520h	0,44	2,39E-03	5,93E-01	hsa-miR-634	0,52	3,62E-02	3,51E-01
hsa-miR-22	2	3,14E-03	6,72E-01	hsa-miR-630	0,6	3,75E-02	1,93E-02
hsa-miR-10b	2	3,20E-03	2,20E-02	hsa-miR-590-3p	0,61	3,84E-02	3,33E-01
hsa-miR-219-5p	2,3	3,33E-03		hsa-miR-145	1,61	4,41E-02	

hsa-miR-140-5p	1,6	4,70E-02	3,92E-02	hsa-miR-383	1,59	4,87E-02
hsa-miR-212	0,62	4,76E-02	9,30E-01	hsa-miR-195	1,53	7,30E-02 3,04E-02

**B)**

<b>mikro-RNS</b>	<b>HR</b>	<b>P-érték (univ.)</b>	<b>P-érték (multiv.)</b>			
hsa-miR-126	2,69	4,9E-05		hsa-miR-381	0,6	0,009009
hsa-miR-122	2,3	5,21E-05		hsa-miR-92a	0,6	0,016594 0,01427
hsa-miR-106b	2,23	5,56E-05		hsa-miR-24	1,61	0,018346 0,696488
hsa-miR-195	1,93	0,000806	0,001688	hsa-miR-19b	1,78	0,019558
hsa-miR-30a	1,92	0,000919		hsa-miR-630	0,62	0,022973
hsa-miR-7	0,44	0,000988		hsa-miR-26a	0,64	0,023252 0,457265
hsa-miR-139-5p	2,1	0,002976		hsa-miR-103	1,54	0,02937
hsa-miR-486-3p	1,83	0,003682	0,166657	hsa-miR-221	0,65	0,030609 0,660817
hsa-miR-142-3p	1,82	0,003897		hsa-miR-9	0,64	0,032165
hsa-miR-29b	1,75	0,004775		hsa-miR-15b	1,54	0,039052 0,245006
hsa-miR-378	1,73	0,005122		hsa-miR-182	1,5	0,040979
hsa-miR-99a	0,56	0,005264		hsa-miR-451	1,51	0,042946 0,52744
				hsa-miR-22	1,52	0,044638 0,672517

A tumor stádiumával és a betegek nemével végzett multivariáns túlélés elemzés során mindössze két mikro-RNS – hsa-miR-10b és a hsa-miR-195 – mutatott szignifikáns összefüggést a teljes túléléssel, mind a TCGA és a GSE31384-es adatokkal végzett multivariáns túlélés elemzés esetén.

### 5.3.2.3 Eltérő kifejeződést mutató mikro-RNS-ek a normál és a tumoros májszövet között

Összehasonlítottam a 173 prognosztikus mikro-RNS kifejeződését a normál és a tumoros májszövet között a TCGA, a GSE10694 és a GSE6857 vizsgálat expressziós adatainak a felhasználásával. Összesen 113 mikro-RNS mutatott eltérő expressziót a tumor és a normál májszövet között (**12. táblázat**).

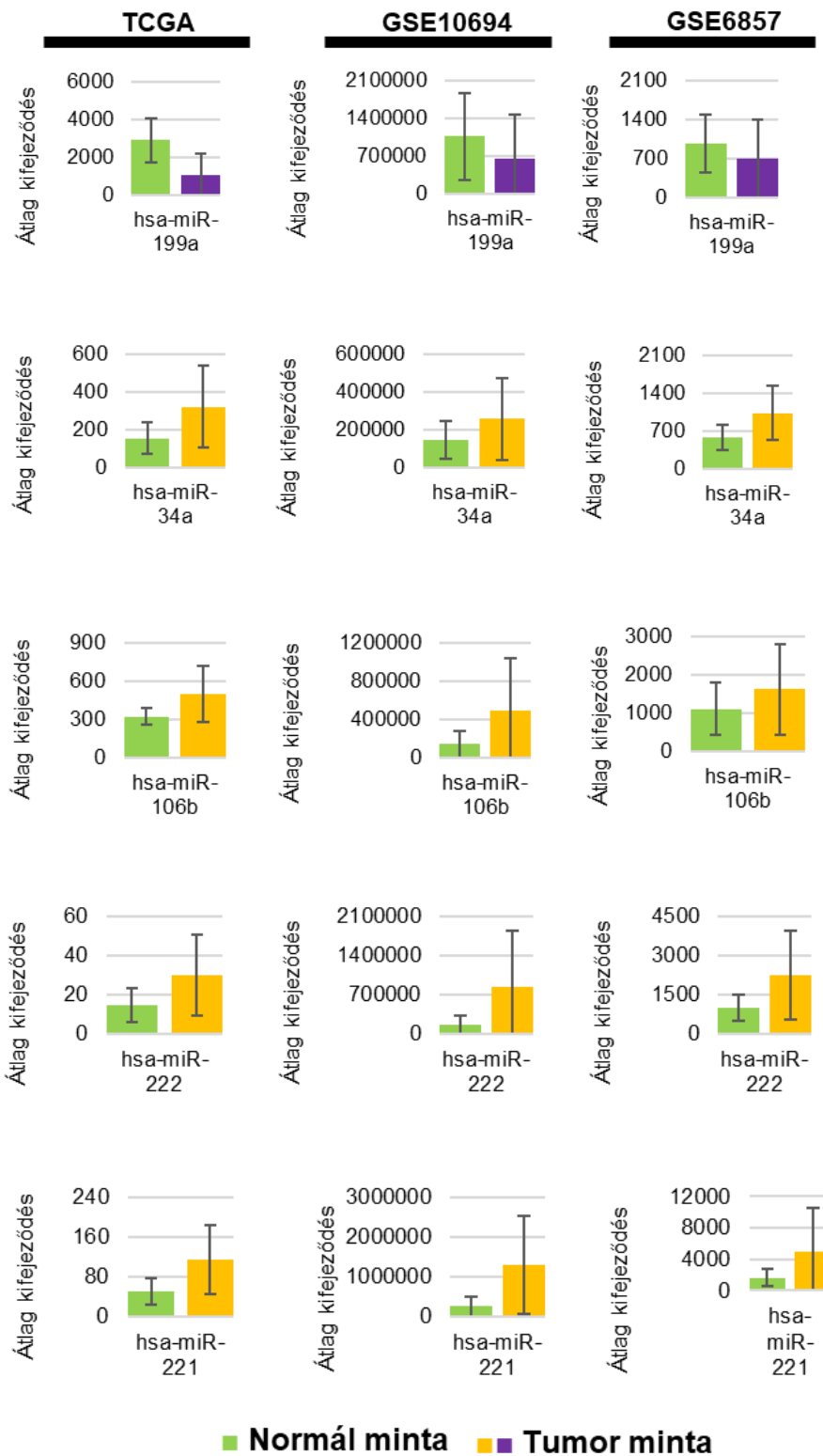
**12. táblázat.** Mikro-RNS-ek kifejeződése tumoros és normál májszövet között.

<b>mikro-RNS</b>	<b>Dataset</b>	<b>P-érték</b>	<b>FC</b>				
				hsa-miR-130a	GSE10694	7,02E-07	0,64
hsa-let-7g	TCGA	6,26E-03	0,89	hsa-miR-132	TCGA	1,07E-03	1,42
hsa-miR-1	GSE6857	P<1E-10	1,4	hsa-miR-135a	GSE6857	P<1E-10	1,74
hsa-miR-100	GSE10694	3,50E-09	0,55	hsa-miR-135a	TCGA	9,62E-05	9,53
hsa-miR-100	GSE6857	1,33E-02	0,97	hsa-miR-137	GSE6857	1,21E-02	1,2
hsa-miR-101	TCGA	P<1E-10	0,47	hsa-miR-138	GSE6857	1,04E-04	0,99
hsa-miR-101	GSE6857	P<1E-10	1,2	hsa-miR-139	TCGA	P<1E-10	0,28
hsa-miR-101	GSE10694	3,00E-10	0,38	hsa-miR-139	GSE6857	1,20E-09	0,87
hsa-miR-103	GSE10694	7,40E-09	1,69	hsa-miR-140	GSE6857	P<1E-10	1,83
hsa-miR-103	GSE6857	1,37E-07	1,32	hsa-miR-140	TCGA	7,15E-05	1,36
hsa-miR-105	GSE6857	P<1E-10	1,39	hsa-miR-141	GSE6857	P<1E-10	1,33
hsa-miR-106b	GSE10694	P<1E-10	3,53	hsa-miR-141	TCGA	3,73E-05	1,57
hsa-miR-106b	GSE6857	P<1E-10	1,47	hsa-miR-145	TCGA	1,43E-07	0,49
hsa-miR-106b	TCGA	3,21E-08	1,56	hsa-miR-145	GSE6857	3,92E-06	0,93
hsa-miR-107	GSE6857	P<1E-10	1,8	hsa-miR-145	GSE10694	4,32E-03	0,79
hsa-miR-107	GSE10694	6,00E-10	1,78	hsa-miR-1468	TCGA	1,18E-04	0,63
hsa-miR-10b	TCGA	P<1E-10	14,18	hsa-miR-148	GSE6857	3,26E-07	0,83
hsa-miR-122	TCGA	5,63E-04	0,7	hsa-miR-148a	GSE10694	1,36E-08	0,67
hsa-miR-125a	GSE6857	3,90E-09	0,83	hsa-miR-148a	TCGA	2,69E-03	0,76
hsa-miR-125a	TCGA	2,69E-03	0,7	hsa-miR-149	GSE6857	6,00E-10	1,17
hsa-miR-125a	GSE10694	2,83E-02	0,82	hsa-miR-155	GSE6857	1,51E-08	1,78
hsa-miR-125b	GSE6857	P<1E-10	1,66	hsa-miR-155	GSE10694	7,33E-04	0,68
hsa-miR-125b	TCGA	1,00E-10	0,52	hsa-miR-15a	GSE6857	P<1E-10	1,73
hsa-miR-125b	GSE10694	3,00E-10	0,55	hsa-miR-15b	GSE10694	3,00E-10	3,16
hsa-miR-126	GSE6857	2,16E-06	1,24	hsa-miR-15b	TCGA	1,65E-02	1,4
hsa-miR-126	TCGA	2,10E-03	0,73	hsa-miR-17-5p	GSE10694	2,60E-09	2,13
hsa-miR-126	GSE10694	2,08E-02	0,89	hsa-miR-181b	GSE6857	P<1E-10	1,36
hsa-miR-130a	TCGA	P<1E-10	0,31	hsa-miR-181b	TCGA	2,85E-02	1,52

hsa-miR-181b	GSE10694	3,53E-02	0,72	hsa-miR-212	GSE6857	P<1E-10	1,26
hsa-miR-182	GSE6857	P<1E-10	1,55	hsa-miR-212	TCGA	1,27E-02	1,4
hsa-miR-182	TCGA	1,80E-09	10,18	hsa-miR-214	GSE6857	P<1E-10	0,82
hsa-miR-183	GSE6857	P<1E-10	1,36	hsa-miR-214	TCGA	7,17E-08	0,39
hsa-miR-183	TCGA	1,00E-10	15,47	hsa-miR-214	GSE10694	1,92E-03	0,92
hsa-miR-185	GSE6857	P<1E-10	1,21	hsa-miR-215	GSE6857	3,00E-07	0,8
hsa-miR-185	GSE10694	3,90E-09	2,16	hsa-miR-216a	TCGA	7,18E-04	10,5
hsa-miR-185	TCGA	5,95E-03	1,29	hsa-miR-216b	TCGA	2,79E-03	11,16
hsa-miR-186	GSE6857	P<1E-10	1,2	hsa-miR-217	GSE6857	1,00E-10	1,14
hsa-miR-192	TCGA	1,88E-03	0,78	hsa-miR-217	TCGA	7,18E-04	12,1
hsa-miR-194	GSE6857	1,90E-09	1,37	hsa-miR-218	GSE6857	P<1E-10	1,22
hsa-miR-194	GSE10694	3,19E-03	0,8	hsa-miR-22	GSE10694	1,16E-06	0,67
hsa-miR-195	GSE10694	1,67E-06	0,62	hsa-miR-22	TCGA	4,85E-06	0,73
hsa-miR-195	TCGA	1,99E-05	0,61	hsa-miR-221	GSE10694	P<1E-10	5,14
hsa-miR-199a	GSE6857	P<1E-10	0,73	hsa-miR-221	GSE6857	P<1E-10	3,01
hsa-miR-199a	TCGA	1,40E-09	0,36	hsa-miR-221	TCGA	4,47E-08	2,27
hsa-miR-199a	GSE10694	6,52E-06	0,61	hsa-miR-222	GSE10694	P<1E-10	4,95
hsa-miR-199b	TCGA	2,40E-09	0,37	hsa-miR-222	GSE6857	P<1E-10	2,24
hsa-miR-199b	GSE6857	3,20E-09	0,87	hsa-miR-222	TCGA	5,56E-07	2,06
hsa-miR-19b	TCGA	3,87E-02	0,84	hsa-miR-223	TCGA	3,93E-07	0,47
hsa-miR-200c	GSE6857	7,00E-10	1,14	hsa-miR-223	GSE6857	2,87E-04	0,91
hsa-miR-200c	TCGA	2,85E-02	4,59	hsa-miR-23a	GSE10694	2,09E-02	1,16
hsa-miR-203	GSE6857	6,80E-07	1,22	hsa-miR-23b	TCGA	2,02E-03	0,8
hsa-miR-203a	TCGA	2,18E-02	0,84	hsa-miR-24	GSE10694	1,49E-02	1,1
hsa-miR-204	GSE6857	2,00E-10	1,26	hsa-miR-25	GSE10694	P<1E-10	2,99
hsa-miR-20a	GSE10694	1,41E-06	1,7	hsa-miR-25	TCGA	4,22E-07	1,63
hsa-miR-20b	GSE10694	2,08E-03	1,76	hsa-miR-26a	TCGA	1,15E-05	0,74
hsa-miR-21	TCGA	P<1E-10	2,53	hsa-miR-26a	GSE6857	7,49E-05	0,84
hsa-miR-21	GSE6857	P<1E-10	3,36	hsa-miR-26a	GSE10694	4,70E-03	0,89

hsa-miR-27b	TCGA	1,02E-06	0,66	hsa-miR-34a	GSE10694	4,15E-05	1,76
hsa-miR-27b	GSE10694	1,35E-02	0,88	hsa-miR-34b	GSE6857	6,63E-04	1,1
hsa-miR-29a	GSE10694	2,15E-04	0,81	hsa-miR-3607	TCGA	6,63E-04	0,48
hsa-miR-29b	GSE6857	6,36E-05	1,18	hsa-miR-3677	TCGA	5,20E-09	3,92
hsa-miR-29b	GSE10694	8,67E-03	0,78	hsa-miR-370	GSE6857	P<1E-10	1,19
hsa-miR-29b	TCGA	2,56E-02	0,84	hsa-miR-370	TCGA	3,29E-04	0,7
hsa-miR-29c	GSE10694	3,64E-07	0,44	hsa-miR-372	GSE6857	P<1E-10	1,48
hsa-miR-29c	TCGA	5,87E-05	0,63	hsa-miR-375	TCGA	2,79E-03	1,16
hsa-miR-301a	TCGA	2,62E-02	1,35	hsa-miR-381	TCGA	7,15E-05	0,61
hsa-miR-301b	TCGA	7,12E-04	2,8	hsa-miR-421	TCGA	1,55E-06	2,51
hsa-miR-302d	GSE6857	P<1E-10	1,56	hsa-miR-429	TCGA	2,50E-04	1
hsa-miR-30a	TCGA	7,31E-06	0,75	hsa-miR-451	GSE10694	7,55E-07	0,58
hsa-miR-30b	TCGA	1,60E-05	0,66	hsa-miR-455	TCGA	2,33E-05	0,65
hsa-miR-30c	TCGA	2,34E-06	0,65	hsa-miR-496	TCGA	1,31E-02	0,97
hsa-miR-30c	GSE10694	1,01E-03	0,8	hsa-miR-503	TCGA	1,27E-04	0,59
hsa-miR-30d	TCGA	2,46E-05	1,63	hsa-miR-550a	TCGA	1,66E-02	1,46
hsa-miR-30d	GSE10694	1,07E-04	1,37	hsa-miR-584	TCGA	4,82E-03	1,24
hsa-miR-325	GSE6857	9,10E-09	0,75	hsa-miR-675	TCGA	1,50E-06	0,68
hsa-miR-326	TCGA	7,48E-04	0,77	hsa-miR-744	TCGA	3,33E-02	0,86
hsa-miR-330	GSE6857	P<1E-10	1,26	hsa-miR-9	TCGA	9,55E-05	16,99
hsa-miR-330	TCGA	1,32E-06	1,97	hsa-miR-940	TCGA	3,98E-03	2,08
hsa-miR-33a	TCGA	4,71E-02	0,89	hsa-miR-96	TCGA	1,23E-07	10,85
hsa-miR-33b	TCGA	1,99E-05	0,48	hsa-miR-99a	GSE10694	3,00E-10	0,49
hsa-miR-34a	GSE6857	P<1E-10	1,75	hsa-miR-99a	TCGA	4,45E-08	0,49
hsa-miR-34a	TCGA	1,94E-06	2,06				

Kiválasztottam azokat a mikro-RNS-eket ( $n = 52$ ), amelyek kifejeződése legalább egy vizsgálati csoport adatai alapján szignifikáns eltérést mutatott a tumoros és a normál májszöveti mintapárok között. Ezek közül a legszignifikánsabb öt mikro-RNS a hsa-miR-199a, hsa-miR-34a, hsa-miR-106b, hsa-miR-222 és a hsa-miR-221. A hsa-miR-199a a tumoros szövetben alacsonyabb, míg a hsa-miR-34a, hsa-miR-106b, hsa-miR-222 és a hsa-miR-221 pedig magasabb kifejeződést mutat mind a három adatbázis esetén (**24. ábra**).



**24. ábra.** A legnagyobb expressziós eltérést mutató mikro-RNS-ek tumoros és normál májszövet között.



#### **5.3.2.4 Prognosztikus mikro-RNS-ek korrelációja**

Spearman-féle rang korrelációval összehasonlítottam az univariáns túlélés elemzés során a teljes túléléssel összefüggő mikro-RNS-ek expressziós értékeit, a TCGA és a GSE31384 betegcsoportok esetén. A TCGA adatok esetén az 55 prognosztikus mikro-RNS-ből a legnagyobb korrelációt a hsa-miR-182 – hsa-miR-183, a hsa-miR-221 – hsa-miR-222 és a hsa-miR-200c – hsa-miR-141 mikro-RNS-ek esetén kaptam, ahol a korrelációs koefficiens ( $r$ ) 0,98, 0,91 és 0,90 volt. A GSE31384 vizsgálat során a 84 prognosztikus mikro-RNS-ből, a hsa-miR-769-5p – hsa-miR-876-5p ( $r = 0.73$ ) mikro-RNS-ek között azonosítottam a legnagyobb korrelációt.

### **5.3.3 Génexpresszió alapú biomarkerek vizsgálata máj hepatocelluláris karcinómában**

#### **5.3.3.1 Prognosztikus biomarkerek irodalmi keresése a máj hepatocelluláris karcinómában**

Irodalmi keresést végeztünk az NCBI PubMed adatbázisban, ahol célunk volt túléléssel összefüggő, génexpresszió alapú biomarkereket azonosítani máj hepatocelluláris karcinómában az 1998 és 2017 közötti időszakban. A keresés során a „hepatocellular”, „carcinoma”, „overall”, „survival”, „biomarker” és „mRNA” keresőszavakat alkalmaztuk.

A keresés során 533 közleményt azonosítottunk, amelyből 506 angol nyelvű és nem „review” közlemény volt. Ebből összesen 355 olyan közleményt volt, amelyekben az mRNS vagy fehérje kifejeződés és a túlélés közötti kapcsolatot vizsgálták. Egyes génekhez több közlemény is tartozott, így az ugyanazon géneket leíró cikkek kizárásával, 318 biomarkert azonosítottunk, amelyekből 194 gén magas kifejeződése rosszabb túléléssel társult. Mindössze 17 közlemény alapult kaukázusi betegpopuláción, amelyekben 21 biomarkert azonosítottak, az összes többi közlemény esetén a vizsgált betegek, Ázsiából – Kína, Észak/Dél-Korea, Taiwan és Japán – származtak.

A legtöbb vizsgálatban a fehérje kifejeződés és a túlélés közötti kapcsolatot vizsgálták. A túléléselemzések során használt adatok 57 %-a fehérje expresszió, ami IHC vizsgálatokból származik, 38 %-a mRNS expresszió és a maradék 5 % pedig nem saját (már korábban közzétett) adatokat használt.

### 5.3.3.2 Prognosztikus biomarkerek vizsgálata az ázsiai és a kaukázusi betegpopulációkban

Az RNS-szekvenálásból és a microarray vizsgálatokból származó általunk létrehozott génexpressziós adatbázis használatával túlélés elemzést végeztünk az ázsiai betegpopuláción belül. A 318 korábban publikált prognosztikus markerből 180 gén kifejeződése esetén tapasztaltunk szignifikáns összefüggést a teljes túléléssel. Ebből a többszörös hiptézistesztesztelés során – 0,1 %-os FDR küszöböt alkalmazva – mindössze 40 szignifikáns biomarkert kaptunk. Nyolc gén esetén a magas génexpresszió jobb túléléssel, míg 32 gén esetén a magasabb génkifejeződés rövidebb túléléssel társult (**13. táblázat**).

**13. táblázat.** Biomarkerek hatása a teljes túlélésre. \* A rossz prognózis alacsony vagy magas génkifejeződés esetén.

Gének	Teljes túlélés HR (95% CI), P	* Kifejeződés alacsony–magas		HR = (95% CI), P	
			ANGPT2	HR = 4.25 (2.32–7.76) P = $3.3 \times 10^{-7}$	magas
			EZH2	HR = 5.51 (2.64–11.51) P = $3.5 \times 10^{-7}$	magas
BIRC5	HR = 7.64 (3.66–15.94) P = $1.9 \times 10^{-10}$	magas	STMN1	HR = 4.97 (2.5–9.88) P = $4.2 \times 10^{-7}$	magas
CDC20	HR = 5.53 (2.95–10.38) P = $2.5 \times 10^{-9}$	magas	CDC25A	HR = 4.22 (2.28–7.8) P = $6.7 \times 10^{-7}$	magas
PLK1	HR = 6.02 (3.08–11.78) P = $3 \times 10^{-9}$	magas	SLC22A1	HR = 0.24 (0.13–0.45) P = $8.4 \times 10^{-7}$	alacsony
ALDH2	HR = 0.19 (0.1–0.35) P = $4.2 \times 10^{-9}$	alacsony	AURKA	HR = 4.5 (2.33–8.66) P = $9.6 \times 10^{-7}$	magas
CCNB1	HR = 7.09 (3.29–15.29) P = $6.3 \times 10^{-9}$	magas	SEC62	HR = 0.25 (0.14–0.46) P = $1.1 \times 10^{-6}$	alacsony
FOXK2	HR = 6.83 (3.16–14.73) P = $1.5 \times 10^{-8}$	magas	FOXM1	HR = 4.5 (2.31–8.78) P = $1.4 \times 10^{-6}$	magas
KIF18A	HR = 4.9 (2.66–9.01) P = $1.9 \times 10^{-8}$	magas	ADH4	HR = 0.26 (0.14–0.47) P = $1.8 \times 10^{-6}$	alacsony
BUB1B	HR = 4.85 (2.64–8.92) P = $2.4 \times 10^{-8}$	magas	MKI67	HR = 3.89 (2.13–7.12) P = $2.2 \times 10^{-6}$	magas
MELK	HR = 5.64 (2.84–11.22) P = $3.1 \times 10^{-8}$	magas	MAD2L1	HR = 9.92 (3.06–32.14) P = $2.6 \times 10^{-6}$	magas
KIAA1524	HR = 4.71 (2.57–8.63) P = $4.0 \times 10^{-8}$	magas	CDK4	HR = 6.8 (2.68–17.28) P = $3.1 \times 10^{-6}$	magas
PTTG1	HR = 7.67 (3.24–18.2) P = $5.2 \times 10^{-8}$	magas	CKAP2	HR = 3.64 (2.01–6.61) P = $5.5 \times 10^{-6}$	magas
CCNF	HR = 4.72 (2.54–8.76) P = $7.1 \times 10^{-8}$	magas	IQGAP2	HR = 0.2 (0.09–0.44) P = $7.3 \times 10^{-6}$	alacsony
PKM2	HR = 4.48 (2.45–8.18) P = $1 \times 10^{-7}$	magas	SOCS2	HR = 0.26 (0.13–0.49) P = $7.7 \times 10^{-6}$	alacsony
DEPDC1	HR = 4.74 (2.47–9.1) P = $2.6 \times 10^{-7}$	magas	E2F1	HR = 8.66 (2.68–28.02) P = $1.4 \times 10^{-5}$	magas
STMN1 and SPP1	HR = 4.31 (2.34–7.92) P = $3.2 \times 10^{-7}$	magas	RRM2	HR = 3.7 (1.96–6.98) P = $1.6 \times 10^{-5}$	magas

HMGA1	HR = 3.38 (1.86–6.13) P = $2.1 \times 10^{-5}$	magas
E2F3	HR = 4.1 (2.02–8.31) P = $2.3 \times 10^{-5}$	magas
KIT	HR = 3.34 (1.84–6.05) P = $2.5 \times 10^{-5}$	magas
CKS1B	HR = 4.06 (2–8.23) P = $2.6 \times 10^{-5}$	magas
FAM83D	HR = 3.58 (1.89–6.78) P = $2.8 \times 10^{-5}$	magas
CENPH	HR = 4.05 (1.99–8.21) P = $2.8 \times 10^{-5}$	magas
PTEN	HR = 0.3 (0.16–0.55) P = $3.1 \times 10^{-5}$	alacsony
HLX	HR = 0.26 (0.13–0.52) P = $3.4 \times 10^{-5}$	alacsony

Az ázsiai betegcsoportra jellemző prognosztikus biomarkerek génontológiai elemzését a Database for Annotation, Visualization and Integrated Discovery (DAVID) program segítségével végeztük (184). Az elemzés során számos olyan biomarkert azonosítottunk, amelyek a sejtciklus és a sejtosztódásban szabályozásával vannak kapcsolatban, többek között részt vesznek a testvér kromatidák összetartásában, a mitotikus sejtmagosztódásban és a ciklinek degradációjában (**14. táblázat**).

**14. táblázat.** Az ázsiai betegcsoportra jellemző gének génontológiai elemzése a DAVID program segítségével.

Funkció	P-érték	Gének	FDR
Sejtosztódás	1,5E-10	CCNB1, FAM83D, CKS1B, MAD2L1, CCNF, BUB1B, BIRC5, AURKA, CDC20, PTTG1, CDK4, CDC25A	2,1E-07
Anafázist támogató complex függő ciklin lebontás	1,9E-08	CCNB1, MAD2L1, PLK1, BUB1B, AURKA, CDC20, PTTG1	2,6E-05
Mitotikus sejtmag osztódás	5,4E-08	FAM83D, PLK1, CCNF, BUB1B, BIRC5, AURKA, CDC20, PTTG1, CDC25A	7,7E-05
Testvér kromatidák összekapcsolódása	9,2E-08	MAD2L1, PLK1, KIF18A, BUB1B, BIRC5, CDC20, CENPH	1,3E-04
Sejtciklus G2/M fázis átmenet	5,0E-07	CCNB1, PLK1, FOXM1, BIRC5, AURKA, CDC25A, MELK	7,1E-04
Fehérje ubikvitináció irányított fehérjelebontás	9,5E-07	CCNB1, MAD2L1, PLK1, BUB1B, AURKA, CDC20, PTTG1	1,4E-03
Sejtciklus szabályozása	7,3E-06	CCNB1, PLK1, FOXM1, CCNF, PTEN, CDC25A	1,0E-02
Sejtosztódás	1,3E-05	FAM83D, CKS1B, MKI67, PLK1, BUB1B, PTEN, CDC25A, MELK	1,8E-02
Ubikvitin fehérje ligáz aktivitásának pozitív szabályozása a mitosis során	2,3E-05	CCNB1, MAD2L1, PLK1, BUB1B, CDC20	3,3E-02

Az ázsiai betegcsoporthoz hasonlóan megvizsgáltuk az összefüggést a 318 korábban publikált prognosztikus biomarker génkifejeződése és a teljes túlélés között a kaukázusi betegpopuláción belül is. Összesen 128 gén esetében tapasztaltunk szignifikáns összefüggést a génkifejeződés és a teljes túlélés között.

Összesen 21 prognosztikus biomarker volt, amit az irodalmi keresés során a kaukázusi populációban írtak le, ebből a mi elemzésünkben 10 biomarker volt szignifikáns: THOC5 ( $P = 0,00018$ ), PLAT ( $P = 0,0012$ ), HMGA1 ( $P = 0,0025$ ), PRKDC ( $P = 0,0096$ ), SLC22A1 ( $P = 0,013$ ), CDC20 ( $P = 0,016$ ), CD274-CXCL12 ( $P = 0,019$ ), CXCL12 ( $P = 0,027$ ), RRM2 ( $P = 0,045$ ) és a ELAVL1 ( $P = 0,049$ ). Nem találtunk olyan gént, amelyik 0,1%-os FDR küszöb esetén szignifikáns maradt volna, azonban az FDR küszöböt 5 %-ra emelve, a SOCS2 gén alacsony kifejeződése ( $HR = 0,37$ ;  $P = 1,3E-05$ ), míg a NAT10 gén magas kifejeződése ( $HR = 3,11$ ;  $P = 4,4E-05$ ) szignifikáns összefüggést mutatott a teljes túléléssel.

A vizsgált 318 biomarkerből, 82 biomarker kifejeződése, az ázsiai és a kaukázusi betegcsoportban szignifikáns összefüggést mutatott a teljes túléléssel. Az irodalmi keresés során ezen gének közül 72 gént eredetileg az ázsiai, 10 gént pedig a kaukázusi betegcsoportban írtak le.

A két betegpopuláció közötti közös biomarkereket tovább vizsgáltuk a Protein ANalysis THrough Evolutionary Relationships (PANTHER) génosztályozó segítségével (185). A program a vizsgált géneket az alapján csoportosítja, hogy azok milyen molekuláris útvonalban vesznek részt. Ennek alapján az ázsiai betegcsoportra 132, míg a kaukázusi betegcsoportra 67 útvonal találatot kaptunk. Az ázsiai populációban a gének nagy része az fibroblaszt növekedési faktor (FGF), inzulin növekedési faktor (IGF) jelátvitelhez, hipoxiára adott válaszhoz és a kemokin és citokin jelátviteli útvonalakhoz kapcsolódtak. A kaukázusi betegcsoportban a Wnt-jelátviteli útvonalhoz, Alzheimerrel kapcsolt útvonalakhoz, valamint a trombocita eredetű növekedési faktor (PDGF), cadherin, JAK/STAT és a gonadotropin-releasing hormon receptorral (GnRHR) kapcsolt jelátviteli útvonalakban részt vevő gének voltak többségben.

További elemzés során megvizsgáltuk a 318 biomarker kifejeződése és a teljes túlélés közötti kapcsolatot a teljes mRNS expressziós adatbázis felhasználásával, azaz az elemzést az etnikai besorolástól függetlenül végeztük el. Összesen 178 gén esetén tapasztaltunk szignifikáns összefüggést a génexpresszió és a teljes túlélés között. A

többszörös hipotézistesztelés során 0,1%-os FDR küszöböt használva 40 biomarker mutatott szignifikáns összefüggést a teljes túléléssel. Ezekből, 17 gén magas kifejeződése jobb, míg 23 gén magas expressziója rosszabb túléléssel korrelált (**15. táblázat**).

**15. táblázat.** A teljes túléléssel összefüggő gének listája a teljes betegcsoportban. \* A rossz prognózis alacsony vagy magas génkifejeződés esetén.

Gének	Teljes túlélés (HR, P)	* Kifejeződés			
			LIN28A	HR=0,46 P=5,2E-6	alacsony
SOCS2	HR=0,35 P=1,8E-9	alacsony	EZH2	HR=2,23 P=6,8E-6	magas
SLC5A8	HR=0,36 P=2,7E-9	alacsony	KIAA1524	HR=2,19 P=9E-6	magas
IL2	HR=0,38 P=1,1E-8	alacsony	PIWIL1	HR=0,46 P=9,7E-6	alacsony
SOX1	HR=0,38 P=1,3E-8	alacsony	PTTG1	HR=2,14 P=1,1E-5	magas
HOTAIR	HR=0,38 P=1,6E-8	alacsony	CCNF	HR=2,15 P=1,2E-5	magas
FOXD3	HR=0,39 P=5,1E-8	alacsony	MAD2L1	HR=2,25 P=1,7E-5	magas
PNLIPRP3	HR=0,4 P=5,3E-8	alacsony	KDR	HR=0,45 P=2E-5	alacsony
NKX2-8	HR=0,41 P=2,2E-7	alacsony	KIF18A	HR=2,09 P=2,2E-5	magas
CDC20	HR=2,49 P=5,1E-7	magas	TREM1	HR=2,13 P=2,4E-5	magas
SLC7A11	HR=2,41 P=5,3E-7	magas	IQGAP2	HR=0,48 P=2,5E-5	alacsony
BIRC5	HR=2,34 P=7,4E-7	magas	WASF2	HR=2,1 P=2,5E-5	magas
SLC22A1	HR=0,43 P=9,2E-7	alacsony	DEPDC1	HR=2,08 P=2,6E-5	magas
STMN1-SPP1	HR=2,45 P=1,2E-6	magas	HMGA1	HR=2,08 P=2,6E-5	magas
ALDH2	HR=0,42 P=1,3E-6	alacsony	THOC5	HR=2,07 P=2,9E-5	magas
STMN1	HR=2,24 P=2,8E-6	magas	CDK4	HR=2,15 P=3E-5	magas
PLK1	HR=2,23 P=3,2E-6	magas	CCNB1	HR=2,34 P=3,4E-5	magas
RASSF10	HR=0,45 P=3,3E-6	alacsony	FBLN5	HR=0,47 P=3,6E-5	alacsony
SPP1	HR=2,27 P=3,5E-6	magas	MELK	HR=2,22 P=3,7E-5	magas
SLC2A1	HR=2,25 P=3,7E-6	magas	PKM2	HR=2,07 P=3,9E-5	magas
ADH4	HR=0,37 P=4,3E-6	alacsony			

### 5.3.3.3 Multivariáns túlélés elemzés

A multivariáns túlélés elemzés során, a betegekhez tartozó összes klinikai adatot felhasználtuk. Elsőként az ázsiai betegcsoportban végeztük el az elemzést. Megvizsgáltuk a korábban azonosított 39 prognosztikus biomarker génexpressziójának, a nemnek és a tumor stádiumának együttes hatását a teljes túlélésre 0,1%-os FDR küszöb alkalmazásával. A vizsgálat során a tumor stádiuma ( $P = 0,0018$ ), a CENPH ( $P = 0,0038$ ) és a CDK4 ( $P = 0,038$ ) gének kifejeződése szignifikáns összefüggést mutattak a teljes túléléssel. A CENPH és a CDK4 gének magas kifejeződése rosszabb túléléssel társult.

A kaukázusi betegcsoportban a multivariáns túlélés elemzés, a klinikai változók mellett, a SOCS2 és a NAT10 gének (korábbi elemzés során a kaukázusi betegcsoportban azonosított prognosztikus gének) expresszióját tartalmazta. A vizsgálat során 5%-os FDR küszöb mellett a SOCS2 ( $P = 0,046$ ) és a NAT10 ( $P = 0,031$ ) gének expressziója függött össze a teljes túléléssel, azonban a tumor stádiuma és a nem esetén nem tapasztaltunk szignifikáns összefüggést.

A teljes adatbázisban is elvégeztük a multivariáns túlélés elemzést. A modell a teljes adatbázisban azonosított 40 prognosztikus biomarker kifejeződését, a tumor stádiumát és a betegek nemét tartalmazta. A teljes túléléssel szignifikáns összefüggést mutatott a tumor stádiuma ( $P = 2,06E-05$ ), az SPP1 ( $P = 0,0002$ ), a PKM2 ( $P = 0,0008$ ), az EZH2 ( $P = 0,0017$ ), a KIF18A ( $P = 0,0061$ ), a DEPDC1 ( $P = 0,0092$ ), a CCNF ( $P = 0,019$ ), az SLC2A1 ( $P = 0,036$ ) és a WASF2 ( $P = 0,048$ ) gének kifejeződése. Mind a nyolc gén esetén a magasabb génexpresszió rosszabb túléléssel társult.

### 5.3.3.4 A vaszkuláris invázióval összefüggő gének

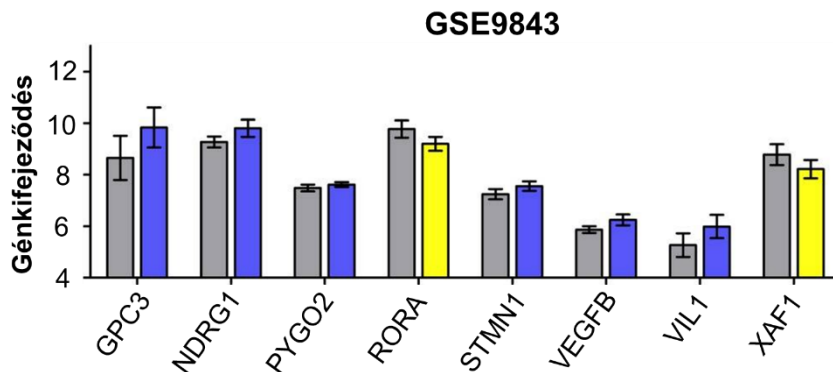
A korábbi – „hepatocellular”, „carcinoma”, „overall”, „survival”, „biomarker” és „mRNA” – kulcsszavas keresést a „vascular invasion” és a „PVT” kifejezésekkel bővítettük, és a célunk volt olyan biomarkerek azonosítása máj hepatocelluláris karcinómában, amelyek nem csak, hogy prognosztikusak, de a vaszkuláris invázióval is összefüggést mutatnak. Célunk volt továbbá, hogy a saját adatbázisunk felhasználásával Mann-Whitney teszt alkalmazásával igazoljuk ezen biomarkerek vaszkuláris invázióval való összefüggését máj hepatocelluláris karcinómában. Összesen nyolc közlemény felelt

meg a keresési feltételeinknek, amelyben 52 vaszkuláris invázióval összefüggő egyedi biomarkert írtak le.

Mivel a TCGA program betegek esetén nem állt rendelkezésre a vaszkuláris invázióval kapcsolatos adat, így az elemzésben csak a GSE9843 és a GSE20017 vizsgálat microarray génexpressziós adatait használtuk fel.

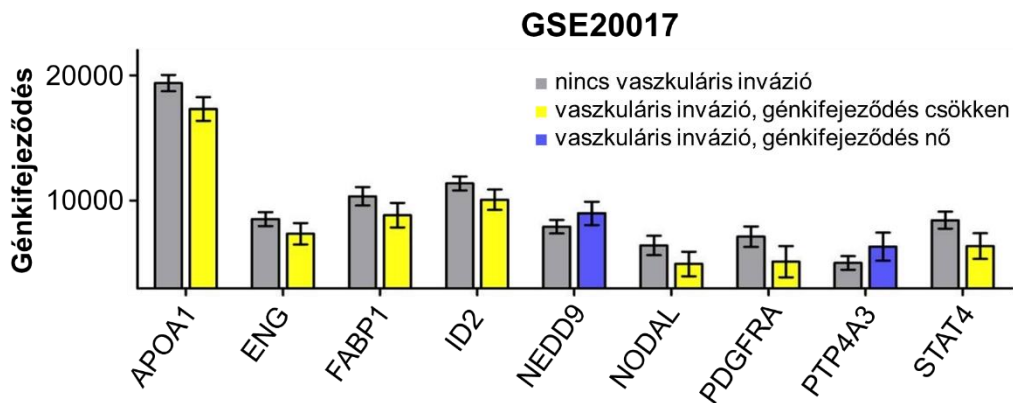
A GSE9843 vizsgálatban 79 betegnek volt meg az adata a vaszkuláris invázióra nézve, amelyek közül 45 páciens pozitív volt. Az 52 vaszkuláris invázióval összefüggő gén kifejeződését összehasonlítottuk a vaszkuláris inváziót mutató és nem mutató betegcsoportok között. Szignifikáns összefüggést kaptunk a GPC3, KIF18A, NDRG1, PYGO2, RORA, STMN1, VEGFB, VIL1 és a XAF1 gének esetén. A GPC3, KIF18A, NDRG1, PYGO2, STMN1, VEGFB és a VIL1 gének magasabb, míg a RORA és az XAF1 gének alacsonyabb kifejeződést mutattak a vaszkuláris inváziót mutató betegcsoportban (25. ábra A).

A.



B.





**25. ábra.** A vaszkuláris invázió hatása a génkifejeződésre a GSE9853 (A) és a GSE20017 (B) betegcsoportok alapján.

A GSE20017 microarray vizsgálat esetén a 135 betegből 40 beteg volt pozitív a vaszkuláris invázióra nézve. A két betegcsoport között 10 gén kifejeződése különbözött szignifikánsan: APOA1, ENG, FABP1, ID2, KIF18A, NEDD9, NODAL, PDGFRA, PTP4A3 és a STAT4. Az APOA1, ENG, FABP1, ID2, NODAL, PDGFRA és a STAT4 gének kifejeződése csökkent, míg a KIF18A, NEDD9 és a PTP4A3 gének expressziója nőtt a vaszkuláris inváziót mutató betegek körében (25. ábra B).

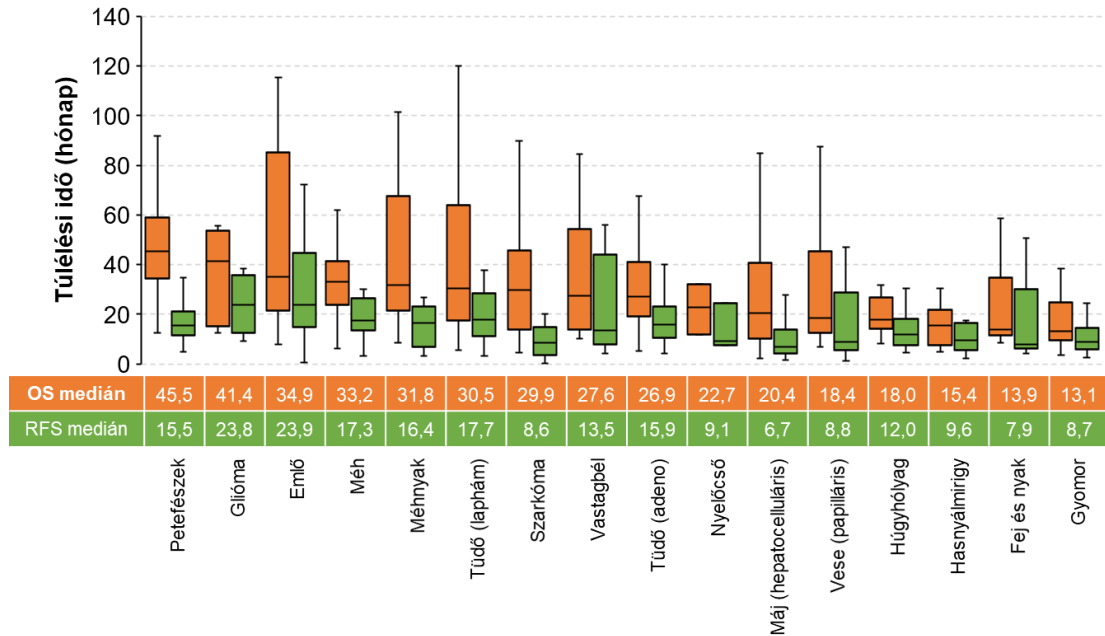
Fontos kiemelni, hogy a KIF18A volt az egyetlen gén, amely mind a két tanulmányban szignifikáns eltérést mutatott a két betegcsoport között, és szignifikánsan magasabb kifejeződést mutatott vaszkuláris invázióra pozitív betegcsoportban.

## 5.4 Daganat asszociált gének prognosztikus hatásának vizsgálata különböző tumortípusokban

### 5.4.1 RNS szekvenálási adatbázis felépítése

Az adatbázis összesen 26 különböző daganattípus, mintegy 9663 beteg RNS szekvenálási adatát tartalmazta, túlélési adattal együtt. A legtöbb tumortípus esetén mind teljes túlélési és kiújulás mentes túlélési adat is egyaránt elérhető volt. Ez alól kivételt jelentett az akut mieloid leukémia, a glioblasztóma, a melanóma és a thymóma tumortípusok, amelyek esetén csak kiújulás mentes túlélési adat volt elérhető. A medián túlélési idő az összes daganattípust véve, a teljes túlélés esetén 24,3 hónap, a kiújulás mentes túlélés esetén pedig 23,8 hónap.

A **26. ábra** mutatja a különböző daganattípusok teljes, valamint a kiújulás mentes túlélési idejének a megoszlását azon betegeknél, ahol volt túlélési esemény.



**26. ábra.** A teljes és a kiújulás mentes túlélési idő megoszlása különböző daganattípusok esetén.

A túlélési adatok mellett számos klinikai adattal rendelkeztek a betegek. Összesen 6301 beteg esetén volt ismert a daganat stádiuma, 4126 betegnek a daganat grádusa, 9720 páciensnek a neme, valamint 9471 betegnek az etnikai besorolása. A daganat stádiuma alapján a fej-nyak daganatok esetén van a legtöbb 4-es stádiumú beteg (a betegek több, mint 50%-a), míg a heredaganatok vezetnek a 0 vagy 1-es stádiumú betegek (~ 80%) tekintetében. A hólyagdaganatok esetén a betegek 95%-a magas tumor grádussal bír, valamint a legtöbb alacsony grádusú páciens a fej-nyak daganatok esetén fordul elő. A páciensek nem és etnikai hovatartozása tekintetében elmondható, hogy a TCGA program daganatos betegei között nagyobb számban vannak férfiak, akik a kaukázusi populációhoz tartoznak.

### 5.4.2 Daganat asszociált gének hatása a teljes túlélésre

Cox regresszió során megvizsgáltam a nyolc fő daganatjellemező kialakításában részt vevő 671 daganat asszociált gén (6), (180) kifejeződése és a teljes túlélés közötti kapcsolatot 26 különböző daganattípus RNS szekvenálási adatát felhasználva.

Ezt követően meghatároztuk a túlélés elemzés során szignifikáns gének százalékos mennyiségét minden daganattípus és daganatjellemező esetén. Hierarchikus klaszterezés során megvizsgáltam az egyes daganattípusok és az egyes daganattényzőt alkotó géncsoportok összefüggését. Megfigyeltem, hogy a daganatok áttétképzésében, genom instabilitásában, a folyamatos sejtosztódás fenntartásában és a metabolikus folyamatok átprogramozásában részt vevő gének különálló csoportokat, úgynevezett klasztereket alakítanak ki. A világos sejtes veserák, a glióma, a melanóma, a thymoma és a máj hepatocelluláris karcinóma daganattípusok tartalmazták a legnagyobb számú teljes túléléssel összefüggő daganat asszociált gént.

### 5.4.3 Transzkriptomikus mintázatok prognosztikus hatása

Minden tumortípus esetén meghatároztuk a daganatjellemező asszociált génextpressziós mintázatot, majd daganattípusonként vizsgáltuk az összefüggését a teljes túléléssel. A nyolc daganattényezőhöz kapcsolt transzkriptomikus mintázat közül a sejtosztódás fenntartása, genom instabilitás, energiametabolizmus átalakítása, metasztázis és invázió serkentése, valamint a sejthalál elkerülése mutatott szignifikáns összefüggést a teljes túléléssel legalább öt daganattípus esetén (**16. táblázat**). A glióma esetén, a nyolc daganattényezőhöz kapcsolt transzkriptomikus mintázatból kiemelném a DNS replikáció fenntartását, amely a legszignifikánsabb összefüggést mutatta a teljes túléléssel (HR = 3,16;  $P = 2,2E-11$ ).

A tüdő laphámkarcinoma, a méh, petefészek, szarkóma, a húgyhólyag és a nyelvcsőrák csak egyetlen szignifikáns génextpressziós mintázatot tartalmazott (**16. táblázat**).

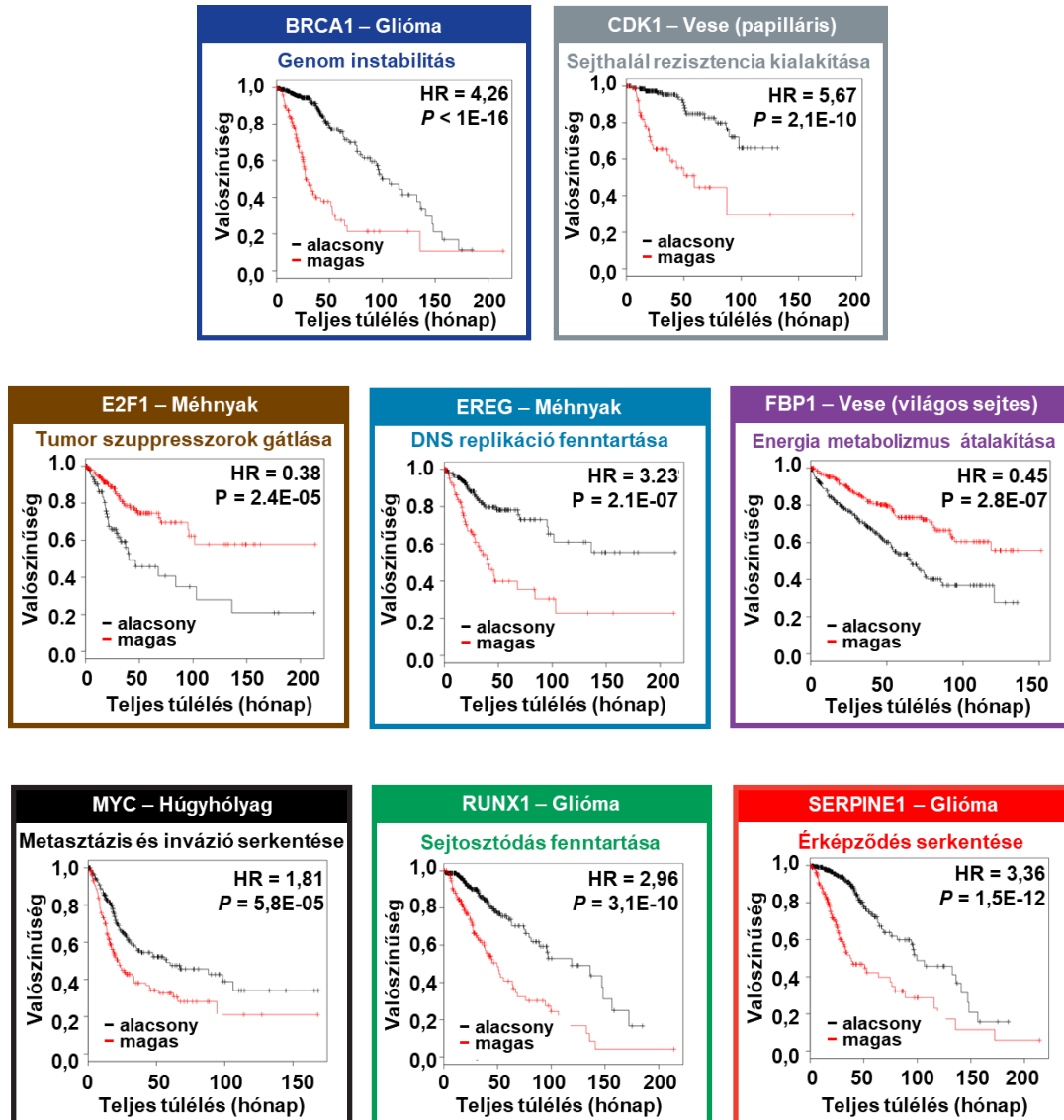
**16. táblázat.** A daganatjellemező asszociált transzkriptomikus prognosztikus hatása különböző daganattípusban.

	Sejtosztódás fenntartása	Tumor szuppr. gátlása	Érképződés s erketése	Genom instabilitás	Energia metabolizmus átalakítása	Metasztázis és invázió serketése	Sejthalál rezisztencia kialakítása	DNS replikáció fenntartása
Glióma	•	•	•	•	•	•		•
Emlő	•		•		•		•	
Tüdő (laphám)	•							
Méh	•							
Petefészek	•							
Hasnyálmirigy	•					•	•	
Máj (hepatocelluláris)	•			•		•	•	•
Akut mieloid leukémia		•			•	•		
Vese (világos sejtes)		•		•	•		•	
Thymóma			•	•				
Melanóma			•		•		•	
Méhnyak				•		•		
Vese (papilláris)				•	•			
Tüdő (adeno)				•	•	•		•
Szarkóma				•				
Fej és nyak					•	•		
Húgyhólyag						•		
Nyelőcső						•		
Pajzsmirigy							•	•

#### 5.4.4 Prognosztikus gének eloszlása különböző daganattípusokban

Összesen 39 gén volt, amelyek kifejeződése legalább 10 daganattípusban mutatott szignifikáns összefüggést a teljes túléléssel. Továbbá, meghatároztuk a nyolc daganatjellemezőt képviselő géncsoportokban, melyek a legszignifikánsabb gének (27.

*ábra*). A legszignifikánsabb gének közül a BRCA1, CDK1, EREG, MYC, RUNX1 és a SERPINE1 gének magas génkifejeződése rosszabb túléléssel társult, míg az E2F1 és az FBP1 gének magas kifejeződése esetén a betegek hosszabb ideig éltek (27. *ábra*).



27. *ábra*. A teljes túléléssel legszignifikánsabb összefüggést mutató gének a különböző daganatjellemzők esetén.

## 6 MEGBESZÉLÉS

A PhD munkám során összesen négy kérdésre kerestem a választ. A kutatásom első részében a szomatikus KRAS mutációk, valamint a KRAS génben található kópiaszám eltérések (szekvencia beépülés és kiesés) génkifejeződésre gyakorolt hatását, valamint ezen transzkriptomikus változásoknak a túléléssel való összefüggését vizsgáltam nem-kissejtes tüdőrákban.

A munkám második részben, az akut mieloid leukémiában kiemelkedő diagnosztikus szereppel bíró szomatikus NPM1 mutációk transzkriptomikus hatását vizsgáltam a túlélésre különböző, egymástól független génkifejeződési adatbázisok bevonásával. Továbbá az eredményeink igazolását *in vitro* vizsgálat során klinikai minták felhasználásával végeztük.

A kutatás harmadik részében korábban publikált mRNS és mikro-RNS alapú biomarkerek prognosztikus hatékonyságát vizsgáltuk máj hepatocelluláris karcinómában több génextpressziós adatbázis felhasználásával.

Végül egy több daganattípust átfogó elemzés során daganatjellelmezőkhöz kapcsolt gének kifejeződésének a túlélésre gyakorolt hatását, valamint az egyes daganatjellelmezőnként meghatározott, transzkriptomikus mintázat prognózisra gyakorolt hatását vizsgáltam különböző daganattípusokban.

### 6.1 KRAS mutáció függő génextpressziós változások vizsgálata nem-kissejtes tüdőrákban

A KRAS gént érintő szomatikus mutációknak a szerepe a túlélésében, valamint a terápiás válasz kialakításában jelenleg még egy meglehetősen kérdéses terület (74).

A vizsgálat során az volt a feltételezésünk, hogy a különböző szintű adatok összekapcsolásával, a KRAS mutációk által befolyásolt génkifejeződési változásoknak prognosztikus jelentősége van a nem-kissejtes tüdőrákban.

Ezen feltételezés igazolására a vizsgálat első lépésében átnéztem az irodalmat 2011 és 2016 között, a KRAS mutációk tüdőrákos betegek túlélésére gyakorolt hatásával kapcsolatban. Az irodalmi keresés során összesen 138 cikk esetében vizsgálták a KRAS mutációk prognosztikus hatását, és ebből a tanulmányok mindössze 10%-a esetén volt a KRAS mutációknak prognosztikus hatása tüdőrákban.

A KRAS mutációk génkifejeződésre gyakorolt hatásának a vizsgálata olyan nem-kissejtes tüdőrákos betegeken történt, amelyek RNS szekvenálási és mutációs adattal is rendelkeztek. A KRAS mutációk alapján a betegeket KRAS mutáns és KRAS vad típusú csoportokra bontottam, majd a két csoport között eltérő kifejeződésű géneket azonosítottam. A legszignifikánsabb öt gén a FOXRED2, KRAS, TOP1, PEX3 és az ABL2 gének voltak. Ezen gének átlagos kifejeződését véve meghatároztam a KRAS mutációval kapcsolt transzkriptomikus mintázatot, amelynek prognosztikus hatását az microarray adatokon tesztelve, erős szignifikáns összefüggést kaptam a teljes túléléssel.

Ezen gének kapcsán általánosságban elmondható, hogy szerepük van különböző daganat kialakulással kapcsolatos folyamatokban. Ezek közül kiemelendő a TOP1 gén, amely részt vesz a DNS replikáció és a genom instabilitás szabályozásában. A TOP1 génre FDA által jóváhagyott célzott terápiás készítmény az irinotecan, amelyet metasztatikus, valamint előrehaladott stádiumú nem-kissejtes tüdőrák esetén alkalmaznak (186). Az eredmények alapján az emelkedett TOP1 kifejeződés, a KRAS mutáns nem-kissejtes tüdőrák esetén egy hatékony prediktív marker lehet. Szintén fontos kiemelni az ABL2 gént, amely részt vesz a sejtosztódás, valamint az invázió és metasztázis folyamatában (187). Egy korábbi tanulmányban az ABL2 génhez kapcsolódó amplifikációt figyeltek meg invazív tüdőrákban (188). Továbbá, az ABL géncsaládba tartozó ABL1 gén esetén az imatinib és a dasatinib nevű célzott terápiás készítmény jóváhagyott a BCR-ABL1 pozitív leukémiák esetén (189). Az eredmények alapján a TOP1 génhez hasonlóan az ABL2 gén kifejeződése is magasabb a KRAS mutáns tüdődaganatban, amely egy potenciális terápiás célpont is lehet. A nem-kissejtes tüdőrákban egy további terápiás célpont lehet a FOXRED2 gén, amely irodalmi adatok alapján is magas kifejeződést mutat számos tumortípus esetén, valamint jelentős szerepet játszik a hipoxia folyamatában és befolyásolja a hipoxia aktiválta előgyógyszerek (hypoxia-activated prodrugs) terápiás hatékonyságát (190), (191).

A módszert ellenőrizve, elvégeztem az elemzést a második, valamint a harmadik legjobb öt gén esetén is, és mindegyik esetben erős szignifikáns összefüggést kaptam a teljes túléléssel.

Továbbá, a kópiaszám eltérések tekintetében, csak a KRAS gént érintő delációkkal asszociált transzkriptomikus mintázat esetén tapasztaltam nagyfokú szignifikáns összefüggést a teljes túléléssel.

Az eredmények alapján elmondható, hogy a genotípus és a génkifejeződési változások *in silico* módon történő összekapcsolásával, a KRAS mutációk génkifejeződési változásokon keresztül kifejtett közvetett hatása a túlélésre sokkal jelentősebb, mint a mutációk közvetlen túlélésre gyakorolt hatása. A KRAS mutációkkal összefüggő felülregulált gének hatékony terápiás célpontok lehetnek a KRAS mutáns nem-kissejtes tüdőrákos betegek esetén.

## **6.2 Az NPM1 mutációk transzkriptomikus lenyomata és validálása akut mieloid leukémiában**

Az NPM1 mutációk az egyik leggyakrabban előforduló mutációk közé tartoznak az akut mieloid leukémiában és fontos szerepet játszanak a daganat patogenezisében (121). A kutatásom ezen részében az NPM1 szomatikus mutációkkal összefüggő transzkriptomikus elváltozásokat azonosítottam és vizsgáltam ezek hatását a betegek túlélésére. Az elemzést három, egymástól független génkifejeződési mintahalmazon – amelyből egy RNS szekvenálási, kettő microarray alapú adat – végeztem. Az elemzések során a HOXA5, HOXA10, HOXB5, PBX3, MEIS1 és az ITM2A géneket választottam ki a klinikai mintákon történő validációra azért, mert ezen gének mutatták a legnagyobb génkifejeződési eltérést az NPM1 mutáns és a vad betegcsoport között, továbbá mert kifejeződésük a legszignifikánsabb összefüggést mutatta a teljes túléléssel. A gének közül az ITM2A génkifejeződése csökkent, míg a többi géné nőtt az NPM1 mutáns mintákban. A klinikai mintákon történő validáció során sikerült igazolnunk ezen gének génkifejeződési eltéréseit, valamint a prognosztikus tulajdonságaikat.

Az elemzés során a HOX gének mellett olyan géneket is azonosítottam az NPM1 mutáns tumorban, amelyek szorosan együttműködnek a HOX génekkel. Ezek a PBX (pre-B-cell leukemia homeobox) és a MEIS (myeloid ecotropic viral integration site homeobox) géncsalád génjei, amelyek a HOX gének kofaktoraiként működnek a sejt differenciálódás és a sejtosztódás folyamatának szabályozása során (**19. ábra**) (192). Fontos kiemelni, hogy a PBX és a MEIS gének szerepe nélkülözhetetlen a HOX gének funkciójához (193).

Az NPM1 mutációk és a HOX (homeobox) gének közötti kapcsolatot számos korábbi tanulmány is alátámasztja (194), (195). Azonban, a HOX gének fokozott kifejeződése és az NPM1 mutációk közötti kapcsolat akut mieloid leukémiában jelenleg



kevésbé ismert. Egyes feltételezések szerint az NPM1 mutációk közvetlenül képesek módosítani a HOX gének kifejeződését, más elméletek szerint pedig az NPM1 mutációk valamilyen másik úton gátolják a sejt differenciálódás folyamatát azon sejtekben, amelyekben már magas a HOX gének kifejeződése (196).

Az NPM1 mutációkkal összefüggő hat gén közül az ITM2A (integral membrane protein 2A) esetén tapasztaltam alacsonyabb kifejeződést az NPM1 mutáns tumorban, amelyet korábbi tanulmány is alátámaszt (197). Az ITM2A szerepet játszik a sejt differenciálódás és az autofágia folyamatában, azonban az akut mieloid leukémia progressziójában betöltött szerepe jelenleg ismeretlen (198).

A HOX gének és kofaktorainak kifejeződésbeli megváltozása számos hematológiai megbetegedés esetében előfordul, amelyek mintegy protoonkogén funkcióval bírnak (199). Az egyes HOX gének kifejeződésbeli megváltozása hozzájárulhat a sejt osztódás serkentéséhez és a sejthalál kikerüléséhez (200), az érképződéshez (201), az áttétképzés stimulálásához (202), továbbá a terápiás rezisztencia kialakításához is (203). Elmondható, hogy minden esetben az emelkedett HOX gén kifejeződés rosszabb klinikai kimenetellel társul (204). A HOX gének tumor szuppresszor szereppel is rendelkeznek úgy, mint például a HOXA5 gén, amely stabilizálja a TP53 tumor szuppresszor fehérjét emlőrák sejtekben (205). A HOX gének mellett a PBX és a MEIS géncsalád tagjai közül a PBX3 és a MEIS1 gének emelkedett kifejeződése is elengedhetetlen szereppel bír a HOX gének tumorigenezisben betöltött szerepében (204).

A HOX és a PBX fehérjék kiemelt terápiás célpontok lehetnek egyes daganatok kezelése során. A legkiemelkedőbb terápiás alternatíva a HOX és a PBX fehérjék közötti kölcsönhatás gátlása, amelyre a HXR9 kis molekulájú peptid gátlószer alkalmas. A HXR9, a HOX géneken található úgynevezett hexapeptid szekvenciához kötődve gátolja a PBX fehérje kötődését (204). Ezen kapcsolat hiánya a daganat növekedés gátlásához vezet. A HXR9 a daganat növekedésére gyakorolt gátló hatását számos tumortípus – úgymint tüdő (206), emlő (207), prosztata (208), melanóma (200) és akut mieloid leukémia (209) – esetén is igazolták. A HXR9 esetében szintén leírták, hogy a HOX gének mellett a protein kináz C-t is gátolja, amely során apoptózist figyeltek meg AML sejtvonalakon (210).

Az NPM1 mutációk és a génkifejeződés összekapcsolásával akut mieloid leukémiában, olyan génexpressziós változásokat azonosítottunk, amelyek összefüggést

mutatnak a daganatsejtek osztódásával, differenciálódásával és az áttétképző képességével, továbbá befolyásolják a betegek túlélését. Elmondható, hogy ezek a gének, úgymint az egyes HOX gének és a kofaktoraik potenciális terápiás célpontok lehetnek akut mieloid leukémiás betegek esetén.

### **6.3 Prognosztikus biomarkerek azonosítása máj hepatocelluláris karcinómában**

Korábbi tanulmányok számos, a túlélés előrejelzésére alkalmas mikro-RNS, valamint mRNS alapú biomarkert írtak már le hepatocelluláris karcinómában, azonban ezek nagy része egy adott vizsgálati csoporton belülre jellemzőek. A munkám során megvizsgáltam ezeknek a korábban publikált biomarkereknek a prognosztikus hatékonyságát, több egymástól független génexpressziós mintahalmaz felhasználásával.

A munka első részében az irodalmi keresés során 173 prognosztikus mikro-RNS-t azonosítottam. Ezen mikro-RNS-ek túléléssel való összefüggésének vizsgálatát egy mikro-RNS szekvenálási, valamint egy microarray alapú génexpressziós adathalmazon végeztem. A túlélés elemzések során a 173 korábban publikált prognosztikus mikro-RNS-ből mindössze 29 esetében azonosítottunk szignifikáns összefüggést a mikro-RNS kifejeződés és a teljes túlélés között mind a két mintacsoport esetén. A mikro-RNS kifejeződés tumoros és normál szövet közötti vizsgálata során 15 mikro-RNS mutatott szignifikáns eltérést, amelyek közül az öt legszignifikánsabb mikro-RNS a hsa-miR-199a, a hsa-miR-34a, hsa-miR-106b, hsa-miR-222 és a hsa-miR-221 mikro-RNS-ek voltak. Ezek közül a hsa-miR-199a kifejeződése csökkent, a hsa-miR-34a, hsa-miR-106b, hsa-miR-222 és a hsa-miR-221 mikro-RNS-ek kifejeződése pedig nőtt a tumoros szövetben. Érdeemes kiemelni a hsa-miR-199a mikro-RNS-t, amely egy korábbi qPCR vizsgálat során mind a daganatos szövetben, mind pedig a sejt kultúrában csökkent kifejeződést mutatott (211). Korábbi kutatások jelentős összefüggést írtak le a hsa-miR-34a-val kapcsolatban is, amely során ezen mikro-RNS emelkedett kifejeződése a sejtciklus megállítást és a sejthalál serkentését idézte elő hepatocelluláris karcinómában (212).

A munka második részében szintén irodalmi kereséssel mRNS alapú biomarkereket azonosítottunk, amely során 318 gén a teljes túléléssel, valamint 52 gén a vaszkuláris invázióval mutatott összefüggést. A biomarkerek tesztelését egy integrált RNS szekvenálási és microarray alapú génkifejeződési adatbázis felhasználásával

végeztük, ahol az ázsiai betegcsoporton belül 180, míg a kaukázusi betegcsoporton belül 128 biomarker bizonyult prognosztikusnak. A többszörös hipotézis tesztelés során a kaukázusi betegcsoportban nem maradt szignifikáns gén, továbbá az ázsiai betegcsoportban a BIRC5, CDC20 és a PLK1 gének mutatták a legerősebb szignifikáns összefüggést.

A prognosztikus biomarkerek közül fontos kiemelni a BIRC5 gént, amely a legszignifikánsabb összefüggést mutatta a teljes túléléssel az ázsiai betegcsoporton belül. A BIRC5 (survivin) a sejthalál gátló fehérje családnak a legkisebb tagja, szerepe van a sejtciklus mitotikus fázisának a szabályozásában, gátolja a sejthalált, valamint befolyásolja az érésködést és a sejtes stresszválasz folyamatokat (213). Számos daganattípusban a BIRC5 magas kifejeződést mutat, amely rossz túléléssel társul (214).

A biomarkerek közül a vaszkuláris invázióval csak a KIF18A kifejeződése mutatott szignifikáns összefüggést, mind az RNS szekvenálási és a microarray adatok esetén. A KIF18A génnek kiemelkedő szerepe van a sejtciklus szabályozása során a mitotikus orsó kialakításában, továbbá emlőtumorban a gén kifejeződése összefüggést mutatott a távoli áttétek kialakulásával is (215), (216).

A korábban publikált mikro-RNS és mRNS alapú biomarkerek prognosztikus értékének alacsony reprodukálhatósága jelzi egy sokkal összetettebb vizsgálati módszer kidolgozásának szükségességét a változatos etiológiájú daganatokban. Összefoglalva, a dolgozatom ezen részében egy olyan bioinformatikai elemzési módszert mutattunk be, amely már rendelkezésre álló transzkriptomikus adatok felhasználásával lehetővé teszi a potenciális biomarkerek validációját és rangsorolását.

#### **6.4 Daganat asszociált gének prognosztikus hatásának vizsgálata különböző tumortípusokban**

A daganatok kialakulásában számos gén játszik szerepet, ezeket korábbi tanulmányok nyolc nagy funkcionális csoportba, úgynevezett daganat jellemzőkbe sorolták (6). A munkám során megvizsgáltam a nyolc funkcionális csoportba tartozó, összesen 671 daganat asszociált gén prognosztikus hatását 26 különböző tumortípus mintegy 9720 mintáján.

A túléléselemzések során, a legtöbb szignifikáns gént a világos sejtes veserák, a glióma, a melanóma, a thymoma és a máj hepatocelluláris karcinóma esetében kaptam.

A túlélési elemzés eredményét hierarchikus klaszterelemzés során tovább vizsgálva megfigyeltem, hogy a gének az egyes daganatjellemzőknek megfelelően csoportosultak. Mind a nyolc fő daganatjellemző esetén kiválasztottuk azokat a géneket – BRCA1, CDK1, E2F1, EREG, FBP1, MYC, RUNX1 és SERPINE1 –, amelyeknek a kifejeződése és a teljes túlélés között a legszignifikánsabb összefüggést kaptam. Továbbá, minden daganatjellemző esetén meghatároztam a gének átlagos kifejeződéséből számolt transzkripciós mintázatokat, amelyet korreláltattam a teljes túléléssel különböző daganattípusokban. A nyolc daganatjellemzőhöz kapcsolt transzkriptomikus mintázat közül a sejtsztódás fenntartása, a genom instabilitás, az energiametabolizmus átalakítása, a metasztázis és az invázió serkentése, valamint a sejthalál elkerülése mutatott szignifikáns összefüggést a teljes túléléssel legalább öt daganattípus esetén.

Több gén együttes kifejeződésének a hatása a daganatok kialakulására, pathogenézisére és progressziójára jelenleg is a daganatkutatás egyik kiemelt területe. Egy korábbi vizsgálatban, megfigyelték, hogy számos tumortípus esetén, a daganat sejtek növekedésével és osztódásával összefüggő gének emelkedett kifejeződése, valamint a sejt differenciálódással összefüggő gének alacsony kifejeződése esetén csökkent a betegek túlélése (217). Az egyes daganattípusok áttétképző képessége kulcsfontosságú szerepet tölt be a betegek túlélésében. Ennek kapcsán számos korábbi vizsgálatban olyan génkifejeződési elváltozásokat azonosítottak, amelyek a daganat metasztázissal mutattak összefüggést és kiemelt prognosztikus, valamint prediktív értékük van egyes tumortípusok – a hasnyálmirigyrák (218), a vastagbél (219) valamint a hepatocelluláris karcinóma (220) – esetén. A génkifejeződési mintázatok közül még érdemes kiemelni a nem-kissejtes tüdőrákban azonosított génkifejeződési mintázatokat, amelyek szintén szoros összefüggést mutattak a betegek túlélésével (221) és a tumor mikrokönyezettel (222).

Fontos kiemelni, hogy az elmúlt évtizedben számos génexpresszió alapuló többgénes vizsgálatot fejlesztettek ki, amelyek hozzájárulnak a daganatok pontosabb prognosztikus és prediktív diagnózisához. Ezen vizsgálatok vagy teszteknek a célja, hogy a beteg számára leghatékonyabb terápia alkalmazásán keresztül ne csak a várható túlélés időtartamát növeljék, de a szükségtelen terápiák elhagyásán keresztül a költségeket és mellékhatásokat is csökkentsék (223). A legtöbb tesztet emlőtumorokra fejlesztették ki, ilyen például a 70 gén kifejeződésén alapuló MammaPrint, amely a betegeket alacsony

és magas daganat kiújulási kockázatú csoportba sorolja, ezzel hozzájárulva az emlőtumoros betegek hatékonyabb betegek terápiás döntéshozatalához (224). A MammaPrinten kívül számos emlőtumorról kapcsolatos génkifejeződésen alapuló teszt létezik, úgymint a 21 génes Oncotype DX (225) vagy a Prosigna Breast Cancer Prognostic Gene Signature teszt, amely 50 gén kifejeződését vizsgálja (226). Azonban fontos kiemelni, hogy ezek a tesztek egyik hátránya az, hogy a hatásosságukat független vizsgálat nem igazolta és a legtöbb teszt esetében a fejlesztés meglehetősen kisszámú, viszonylag heterogén betegcsoporton történt (227). Így az előrejelző képességük ezen betegcsoportokon kívül eső betegek körében korlátozott.

Jelenleg kevés olyan vizsgálat van, amely különböző daganattípusokban vizsgálja a fentebb említett nyolc daganattényezővel kapcsolt génkifejeződési változásokat, valamint ezek hatását a betegek túlélésére. Mivel a sejtosztódás fenntartásában és a genom instabilitásban részt vevő gének által képzett génkifejeződési mintázat mutatott a legtöbb daganattípus esetén szignifikáns összefüggést a teljes túléléssel, ezért a megbeszélés során ezekre térek ki részletesebben.

Az onkogének számos kiemelt szereppel rendelkeznek a tumorigenezis során, úgymint a sejtosztódás, sejt differenciálódás és a daganatsejtek túlélésének a szabályozásában és fenntartásában. Sajnos, kevés olyan tanulmány létezik, amelyben az onkogének által alkotott génkifejeződési mintázat és a teljes túlélés közötti összefüggést vizsgálják. Ezen vizsgálatok közül az első tanulmányban, MCF-7 emlőrák sejtvonalban olyan MAP kináz jelátviteli útvonallal összefüggő génkifejeződési mintázatot azonosítottak – tartalmazva a HER2, az EGFR, a RAF és a MEK géneket –, amely jelentős összefüggést mutatott a sejtnövekedéssel, a sejtek túlélésével, valamint a metabolizmussal (228). Hasonló eredményeket kaptak az ER-negatív betegmintákon végzett vizsgálatok során is (228). A második vizsgálatban Creighton és munkatársai a prosztatákban magas kifejeződést mutató géneket – c-MYC, AKT, c-SRC, HER2 – azonosított, amelyek összefüggést mutattak az onkogénekkal kapcsolt génkifejeződési mintázat felülreguláltságával (229). Azonban tanulmányban nem vizsgálták az onkogén génkifejeződési mintázat és a túlélés közötti összefüggést. Egy harmadik vizsgálat során pedig megfigyelték, hogy a c-MYC, RAS, E2F3 és a CTNNB1 gének által képzett génkifejeződési mintázat változása prediktív értékű a RAS gátlószerekkel történő terápiára nézve. Továbbá számos daganattípus esetén – többek között petefészek, tüdő és

emlődaganatok esetén –, ezen gének magasabb kifejeződése rosszabb túléléssel társult (230). Elmondható, hogy ezen adatokkal a mi eredményeink is korrelálnak, ahol a 29 gén által képzett onkogénikus génkifejeződési mintázat magasabb kifejeződése, rosszabb túléléssel társult számos daganattípus, mint például a petefészek, a glióma, a hasnyálmirigy és a májdaganatok esetén.

A genom instabilitás az egyik alapvető tényező a daganatok fennmaradása során. Számos gén esetén igazolták, hogy részt vesznek a daganatok genomiális instabilitásának a fenntartásában. Korábbi tanulmányokban vizsgálták a genomiális instabilitással összefüggő génkifejeződési mintázatok összefüggését a túléléssel és egyes klinikai-patológia paraméterrel. Néhány daganattípusban – úgymint az emlő, a tüdő, a glióma, a medulloblasztóma, a mezothelióma és a limfóma – a genomiális instabilitással összefüggő génkifejeződési mintázat magas expressziója rosszabb túléléssel korrelált (231). Továbbá, Habermann és munkatársai olyan 12 génből álló aneuploidiával összefüggő génkifejeződési mintázatot azonosított, amely a várható túlélésnek a független előrejelzője emlőtumoros betegekben (232). Az elemzésünk során a 150 genom instabilitással összefüggő gén által képzett génkifejeződési mintázat nyolc daganattípus esetén mutatott szignifikáns összefüggést a túléléssel. Ezek közül a magas kifejeződés rosszabb túléléssel társult a glióma, a máj hepatocelluláris karcinóma, a papilláris veserák, a szarkóma, valamint a tüdő adenokarcinómában. A világossejtes veserák és a thymóma esetén a fordított összefüggést tapasztaltunk, ahol a mintázat magas kifejeződése a jobb túléléssel társult.

Az eredmények alapján elmondható, hogy lehetőség van minden egyes daganattípus esetén a gyógyszerfejlesztés szempontjából legrelevánsabb, a daganat kialakulásában részt vevő tényezők rangsorolására és kiválasztására, amely elősegíti a leghatékonyabb terápiás célpont kiválasztását.

## 7 KÖVETKEZTETÉSEK

### **7.1. A KRAS mutáció függő génexpressziós változások vizsgálata során a nem-kissejtes tüdőrákban a következő következtetéseket vontam le:**

- 7.1.1. A KRAS mutációknak nincsen közvetlenül hatása a túlélésre a nem-kissejtes tüdőrákban.
- 7.1.2. A KRAS mutációkkal és a KRAS gént érintő deléciókkal összefüggő génkifejeződési mintázatnak jelentős prognosztikus ereje van.
- 7.1.3. A KRAS mutációkkal összefüggő génexpressziós változások, a TOP1, az ABL2 és a FOXRED2 géneken keresztül, közvetlen farmakológiai relevanciával rendelkeznek.

### **7.2. Az NPM1 mutációk transzkriptomikus lenyomatának vizsgálata esetén az alábbi következtetéseket vontam le:**

- 7.2.1. Az NPM1 mutációk hatására nőtt a HOX gének (HOXA5, HOXA10, HOXB5) és a kofaktoraik (PBX3, MEIS1) kifejeződése akut mieloid leukémiában.
- 7.2.2. A HOX gének és a kofaktoraik kifejeződése szignifikáns összefüggést mutatott a betegek túlélésével.
- 7.2.3. A HOX génekből és a kofaktoraiból álló funkcionális komplex hatékony terápia-célpont lehet az akut mieloid leukémia terápiájában.

### **7.3. A máj hepatocelluláris karcinómában történő prognosztikus biomarkerek vizsgálata során az alábbi következtetéseket vontam le:**

- 7.3.1. A korábban publikált 173 prognosztikus mikro-RNS-ből mindössze 29 esetében azonosítottam szignifikáns összefüggést a kifejeződés és a teljes túlélés között, független mikro-RNS kifejeződési adatok felhasználásával.
- 7.3.2. Összesen 113 mikro-RNS esetében azonosítottam expressziós eltérést a normál és a tumoros hepatocelluláris karcinóma szövet között.
- 7.3.3. A mRNS alapú biomarkerek csaknem kevesebb, mint felének maradt meg a prognosztikus tulajdonsága a túlélés elemzés során hepatocelluláris karcinómában.

7.3.4. A KIF18A volt az egyetlen gén, amely a vaszkuláris invázió hatékony előrejelzőjének bizonyul különböző májtumoros betegcsoportok vizsgálata alapján.

**7.4. A daganat asszociált gének prognosztikus hatásának vizsgálata során különböző daganattípusban az alábbi következtetéseket vontam le:**

7.4.1. A daganatjellemzők által képzett transzkriptomikus mintázatok daganatspecifikus összefüggést mutatnak a teljes túléléssel.

7.4.2. Egyes hallmark gének kifejeződése szignifikáns, de eltérő összefüggést mutat a túléléssel több daganattípus esetén.

7.4.3. Az eredmények hozzájárulnak az egyes daganattípusok esetén a legoptimálisabb terápiás célpont kiválasztásához.



## 8 ÖSSZEFOGLALÁS

A munkám első részében a szomatikus KRAS mutációk által befolyásolt transzkriptomikus változások túlélésre gyakorolt hatását vizsgáltam nem-kissejtes tüdőrákban. Mutációs és génkifejeződési adatokat tartalmazó integrált adatbázis létrehozását követően olyan géneket azonosítottam, amelyek szignifikáns kifejeződésbeli különbségeket mutattak a KRAS mutáns és a KRAS vad típusú betegcsoportok között. Ezen gének által meghatározott transzkriptomikus mintázat erős szignifikáns összefüggést mutatott a teljes túléléssel.

Munkám második részében megvizsgáltam az NPM1 mutációk génexpresszióra és a betegek túlélésre gyakorolt hatását akut mieloid leukémiában. NPM1 mutációkkal összefüggő génkifejeződési változásokat azonosítottam három egymástól független betegcsoporton. A HOXA5, HOXA10, HOXB5, PBX3, MEIS1 és az ITM2A gének mutatták a legszignifikánsabb eltérést az NPM1 mutáns és a vad típusú betegcsoport között, amelyet a klinikai mintákon is sikerült igazolnom.

Dolgozatom harmadik részében korábban publikált mikro-RNS és mRNS alapú biomarkerek prognosztikus hatását vizsgáltuk különböző expressziós adatbázisok felhasználásával hepatocelluláris karcinómában. A 173 korábban publikált prognosztikus mikro-RNS-ek kevesebb, mint a fele mutatott szignifikáns összefüggést a teljes túléléssel. Szintén irodalmi kereséssel 318 mRNS alapú prognosztikus biomarkert azonosítottunk, amelyből 180 az ázsiai betegcsoporton belül, míg 128 a kaukázusi betegcsoporton belül bizonyult prognosztikusnak.

A munkám utolsó szakaszában 26 különböző daganattípus RNS szekvenálási adatát felhasználva megvizsgáltam az irodalomban ismeretes 673 daganat asszociált gén kifejeződése és a teljes túlélés között az összefüggést. Daganattípusonként meghatároztam a gének átlagos kifejeződéséből számolt transzkripciós mintázat és a teljes túlélés közötti összefüggést. A nyolc daganatjellemzőhöz kapcsolt transzkriptomikus mintázat közül a sejtosztódás fenntartása, genom instabilitás, energiametabolizmus átalakítása, áttétképzés és invázió serkentése, valamint a sejthalál elkerülése mutatott szignifikáns összefüggést a teljes túléléssel legalább öt daganattípus esetén.

## 9 SUMMARY

In the first part of my work, I investigated the effect of transcriptomic changes influenced by somatic KRAS mutations on survival in non-small cell lung cancer. An integrated database was created containing RNA sequencing and mutation data. Differentially expressed genes were identified between the KRAS mutant and KRAS wild-type patient groups. These genes formed the KRAS transcriptomic fingerprint that associated with KRAS mutations and overall survival.

In the next part of the dissertation, I examined the effect of NPM1 mutations on gene expression and patient survival in acute myeloid leukemia. Gene expression changes were identified which were associated with NPM1 mutations using three independent datasets. The HOXA5, HOXA10, HOXB5, PBX3, MEIS1, and ITM2A genes showed the most significant gene expression differences between the NPM1 mutant and wild-type patient groups, which was also confirmed in clinical specimens.

In the third phase of my work, I demonstrated the prognostic effect of previously published microRNA and mRNA-based biomarkers in hepatocellular carcinoma using different gene expression databases. Less than half of the expression of 173 prognostic microRNAs showed significant association with overall survival in hepatocellular carcinoma. According to the mRNA-based biomarkers, we identified 318 previously published prognostic genes of which 180 biomarkers was associated with survival in the Asian group and 128 biomarkers were prognostic in the Caucasian group.

In the last phase of my work, I examined the prognostic effect of tumor-associated genes using the RNA-seq data of 26 different types of tumors. I examined the association between the expression of 673 cancer hallmark genes and overall survival across all tumor types. A transcriptomic surrogate signature for each cancer hallmark was also determined; this is based on the means of the average expression of the cancer genes associated with the given hallmark. The prognostic significance of these factors was examined in different types of cancers. Among the eight main hallmark signatures, those associated with oncogene activation, genome instability, cellular energetics, invasion and metastasis and cell death resistance were significant in at least five tumor types.

**10 IRODALOMJEGYZÉK**

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 68: 394-424.
2. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. (2013) Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*, 49: 1374-1403.
3. Dusek L, Muzik J, Maluskova D, Majek O, Pavlik T, Koptikova J, Melichar B, Buchler T, Finek J, Cibula D, Babjuk M, Svoboda M, Vyzula R, Ryska A, Ryska M, Petera J, Abrahamova J. (2014) Cancer incidence and mortality in the Czech Republic. *Klin Onkol*, 27: 406-423.
4. Menyhart O, Fekete JT, Gyorffy B. (2018) Demographic shift disproportionately increases cancer burden in an aging nation: current and expected incidence and mortality in Hungary up to 2030. *Clin Epidemiol*, 10: 1093-1108.
5. Hanahan D, Weinberg RA. (2000) The hallmarks of cancer. *Cell*, 100: 57-70.
6. Hanahan D, Weinberg RA. (2011) Hallmarks of cancer: the next generation. *Cell*, 144: 646-674.
7. Davies MA, Samuels Y. (2010) Analysis of the genome to personalize therapy for melanoma. *Oncogene*, 29: 5545-5555.
8. Jiang BH, Liu LZ. (2009) PI3K/PTEN signaling in angiogenesis and tumorigenesis. *Adv Cancer Res*, 102: 19-65.
9. Sudarsanam S, Johnson DE. (2010) Functional consequences of mTOR inhibition. *Curr Opin Drug Discov Devel*, 13: 31-40.
10. Sherr CJ, McCormick F. (2002) The RB and p53 pathways in cancer. *Cancer Cell*, 2: 103-112.
11. Curto M, Cole BK, Lallemand D, Liu CH, McClatchey AI. (2007) Contact-dependent inhibition of EGFR signaling by Nf2/Merlin. *J Cell Biol*, 177: 893-903.
12. Shaw RJ. (2009) Tumor suppression by LKB1: SIK-ness prevents metastasis. *Sci Signal*, 2: pe55.

13. Ikushima H, Miyazono K. (2010) TGFbeta signalling: a complex web in cancer progression. *Nat Rev Cancer*, 10: 415-424.
14. Lowe SW, Cepero E, Evan G. (2004) Intrinsic tumour suppression. *Nature*, 432: 307-315.
15. Adams JM, Cory S. (2007) The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene*, 26: 1324-1337.
16. White E, DiPaola RS. (2009) The double-edged sword of autophagy modulation in cancer. *Clin Cancer Res*, 15: 5308-5316.
17. Blasco MA. (2005) Telomeres and human disease: ageing, cancer and beyond. *Nat Rev Genet*, 6: 611-622.
18. Cong Y, Shay JW. (2008) Actions of human telomerase beyond telomeres. *Cell Res*, 18: 725-732.
19. Masutomi K, Possemato R, Wong JM, Currier JL, Tothova Z, Manola JB, Ganesan S, Lansdorp PM, Collins K, Hahn WC. (2005) The telomerase reverse transcriptase regulates chromatin state and DNA damage responses. *Proc Natl Acad Sci U S A*, 102: 8222-8227.
20. Maida Y, Yasukawa M, Furuuchi M, Lassmann T, Possemato R, Okamoto N, Kasim V, Hayashizaki Y, Hahn WC, Masutomi K. (2009) An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. *Nature*, 461: 230-235.
21. Hanahan D, Folkman J. (1996) Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell*, 86: 353-364.
22. Olive KP, Jacobetz MA, Davidson CJ, Gopinathan A, McIntyre D, Honess D, Madhu B, Goldgraben MA, Caldwell ME, Allard D, Frese KK, Denicola G, Feig C, Combs C, Winter SP, Ireland-Zecchini H, Reichelt S, Howat WJ, Chang A, Dhara M, Wang L, Ruckert F, Grutzmann R, Pilarsky C, Izeradjene K, Hingorani SR, Huang P, Davies SE, Plunkett W, Egorin M, Hruban RH, Whitebread N, McGovern K, Adams J, Iacobuzio-Donahue C, Griffiths J, Tuveson DA. (2009) Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science*, 324: 1457-1461.
23. Zee YK, O'Connor JP, Parker GJ, Jackson A, Clamp AR, Taylor MB, Clarke NW, Jayson GC. (2010) Imaging angiogenesis of genitourinary tumors. *Nat Rev Urol*, 7: 69-82.

24. Turner HE, Harris AL, Melmed S, Wass JA. (2003) Angiogenesis in endocrine tumors. *Endocr Rev*, 24: 600-632.
25. Baeriswyl V, Christofori G. (2009) The angiogenic switch in carcinogenesis. *Semin Cancer Biol*, 19: 329-337.
26. Ferrara N. (2009) Vascular endothelial growth factor. *Arterioscler Thromb Vasc Biol*, 29: 789-791.
27. Kessenbrock K, Plaks V, Werb Z. (2010) Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell*, 141: 52-67.
28. Kazerounian S, Yee KO, Lawler J. (2008) Thrombospondins in cancer. *Cell Mol Life Sci*, 65: 700-712.
29. Fidler IJ. (2003) The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer*, 3: 453-458.
30. Cavallaro U, Christofori G. (2004) Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer*, 4: 118-132.
31. Micalizzi DS, Farabaugh SM, Ford HL. (2010) Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. *J Mammary Gland Biol Neoplasia*, 15: 117-134.
32. Warburg O. (1956) On the origin of cancer cells. *Science*, 123: 309-314.
33. Jones RG, Thompson CB. (2009) Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes Dev*, 23: 537-548.
34. Yen KE, Bittinger MA, Su SM, Fantin VR. (2010) Cancer-associated IDH mutations: biomarker and therapeutic opportunities. *Oncogene*, 29: 6409-6417.
35. Reitman ZJ, Yan H. (2010) Isocitrate dehydrogenase 1 and 2 mutations in cancer: alterations at a crossroads of cellular metabolism. *J Natl Cancer Inst*, 102: 932-941.
36. Davis A, Gao R, Navin N. (2017) Tumor evolution: Linear, branching, neutral or punctuated? *Biochim Biophys Acta Rev Cancer*, 1867: 151-161.
37. Kinzler KW, Vogelstein B. (1997) Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature*, 386: 761, 763.
38. Negrini S, Gorgoulis VG, Halazonetis TD. (2010) Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol*, 11: 220-228.

39. Dvorak HF. (1986) Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. *N Engl J Med*, 315: 1650-1659.
40. Grivennikov SI, Greten FR, Karin M. (2010) Immunity, inflammation, and cancer. *Cell*, 140: 883-899.
41. Teng MW, Swann JB, Koebel CM, Schreiber RD, Smyth MJ. (2008) Immune-mediated dormancy: an equilibrium with cancer. *J Leukoc Biol*, 84: 988-993.
42. Pages F, Galon J, Dieu-Nosjean MC, Tartour E, Sautes-Fridman C, Fridman WH. (2010) Immune infiltration in human tumors: a prognostic factor that should not be ignored. *Oncogene*, 29: 1093-1102.
43. Vajdic CM, van Leeuwen MT. (2009) Cancer incidence and risk factors after solid organ transplantation. *Int J Cancer*, 125: 1747-1754.
44. Yang L, Pang Y, Moses HL. (2010) TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol*, 31: 220-227.
45. Mougiakakos D, Choudhury A, Lladser A, Kiessling R, Johansson CC. (2010) Regulatory T cells in cancer. *Adv Cancer Res*, 107: 57-117.
46. Ostrand-Rosenberg S, Sinha P. (2009) Myeloid-derived suppressor cells: linking inflammation and cancer. *J Immunol*, 182: 4499-4506.
47. Ettinger DS, Akerley W, Borghaei H, Chang AC, Cheney RT, Chirieac LR, D'Amico TA, Demmy TL, Ganti AK, Govindan R, Grannis FW, Jr., Horn L, Jahan TM, Jahanzeb M, Kessinger A, Komaki R, Kong FM, Kris MG, Krug LM, Lennes IT, Loo BW, Jr., Martins R, O'Malley J, Osarogiagbon RU, Otterson GA, Patel JD, Pinder-Schenck MC, Pisters KM, Reckamp K, Riely GJ, Rohren E, Swanson SJ, Wood DE, Yang SC, Hughes M, Gregory KM, Nccn. (2012) Non-small cell lung cancer. *J Natl Compr Canc Netw*, 10: 1236-1271.
48. Govindan R, Page N, Morgensztern D, Read W, Tierney R, Vlahiotis A, Spitznagel EL, Piccirillo J. (2006) Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic, and end results database. *J Clin Oncol*, 24: 4539-4544.
49. Murray N, Turrisi AT, 3rd. (2006) A review of first-line treatment for small-cell lung cancer. *J Thorac Oncol*, 1: 270-278.

50. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, Rivers CS, Foo CK, Bhatt D, Stinson J, Gnad F, Haverty PM, Gentleman R, Chaudhuri S, Janakiraman V, Jaiswal BS, Parikh C, Yuan W, Zhang Z, Koeppen H, Wu TD, Stern HM, Yauch RL, Huffman KE, Paskulin DD, Illei PB, Varella-Garcia M, Gazdar AF, de Sauvage FJ, Bourgon R, Minna JD, Brock MV, Seshagiri S. (2012) Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat Genet*, 44: 1111-1116.
51. Peifer M, Fernandez-Cuesta L, Sos ML, George J, Seidel D, Kasper LH, Plenker D, Leenders F, Sun R, Zander T, Menon R, Koker M, Dahmen I, Muller C, Di Cerbo V, Schildhaus HU, Altmuller J, Baessmann I, Becker C, de Wilde B, Vandesompele J, Bohm D, Ansen S, Gabler F, Wilkening I, Heynck S, Heuckmann JM, Lu X, Carter SL, Cibulskis K, Banerji S, Getz G, Park KS, Rauh D, Grutter C, Fischer M, Pasqualucci L, Wright G, Wainer Z, Russell P, Petersen I, Chen Y, Stoelben E, Ludwig C, Schnabel P, Hoffmann H, Muley T, Brockmann M, Engel-Riedel W, Muscarella LA, Fazio VM, Groen H, Timens W, Sietsma H, Thunnissen E, Smit E, Heideman DA, Snijders PJ, Cappuzzo F, Ligorio C, Damiani S, Field J, Solberg S, Brustugun OT, Lund-Iversen M, Sanger J, Clement JH, Soltermann A, Moch H, Weder W, Solomon B, Soria JC, Validire P, Besse B, Brambilla E, Brambilla C, Lantuejoul S, Lorimier P, Schneider PM, Hallek M, Pao W, Meyerson M, Sage J, Shendure J, Schneider R, Buttner R, Wolf J, Nurnberg P, Perner S, Heukamp LC, Brindle PK, Haas S, Thomas RK. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet*, 44: 1104-1110.
52. George J, Lim JS, Jang SJ, Cun Y, Ozretic L, Kong G, Leenders F, Lu X, Fernandez-Cuesta L, Bosco G, Muller C, Dahmen I, Jahchan NS, Park KS, Yang D, Karnezis AN, Vaka D, Torres A, Wang MS, Korbel JO, Menon R, Chun SM, Kim D, Wilkerson M, Hayes N, Engelmann D, Putzer B, Bos M, Michels S, Vlasic I, Seidel D, Pinther B, Schaub P, Becker C, Altmuller J, Yokota J, Kohno T, Iwakawa R, Tsuta K, Noguchi M, Muley T, Hoffmann H, Schnabel PA, Petersen I, Chen Y, Soltermann A, Tischler V, Choi CM, Kim YH, Massion PP, Zou Y, Jovanovic D, Kontic M, Wright GM, Russell PA, Solomon B, Koch I, Lindner M,

- Muscarella LA, la Torre A, Field JK, Jakopovic M, Knezevic J, Castanos-Velez E, Roz L, Pastorino U, Brustugun OT, Lund-Iversen M, Thunnissen E, Kohler J, Schuler M, Botling J, Sandelin M, Sanchez-Cespedes M, Salvesen HB, Achter V, Lang U, Bogus M, Schneider PM, Zander T, Ansen S, Hallek M, Wolf J, Vingron M, Yatabe Y, Travis WD, Nurnberg P, Reinhardt C, Perner S, Heukamp L, Buttner R, Haas SA, Brambilla E, Peifer M, Sage J, Thomas RK. (2015) Comprehensive genomic profiles of small cell lung cancer. *Nature*, 524: 47-53.
53. Crabtree JS, Singleton CS, Miele L. (2016) Notch Signaling in Neuroendocrine Tumors. *Front Oncol*, 6: 94.
54. Taniwaki M, Daigo Y, Ishikawa N, Takano A, Tsunoda T, Yasui W, Inai K, Kohno N, Nakamura Y. (2006) Gene expression profiles of small-cell lung cancers: molecular signatures of lung cancer. *Int J Oncol*, 29: 567-575.
55. Zappa C, Mousa SA. (2016) Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res*, 5: 288-300.
56. Kenfield SA, Wei EK, Stampfer MJ, Rosner BA, Colditz GA. (2008) Comparison of aspects of smoking among the four histological types of lung cancer. *Tob Control*, 17: 198-204.
57. Muscat JE, Stellman SD, Zhang ZF, Neugut AI, Wynder EL. (1997) Cigarette smoking and large cell carcinoma of the lung. *Cancer Epidemiol Biomarkers Prev*, 6: 477-480.
58. Boolell V, Alamgeer M, Watkins DN, Ganju V. (2015) The Evolution of Therapies in Non-Small Cell Lung Cancer. *Cancers (Basel)*, 7: 1815-1846.
59. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, Meyerson M. (2004) EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science*, 304: 1497-1500.
60. Wieduwilt MJ, Moasser MM. (2008) The epidermal growth factor receptor family: biology driving targeted therapeutics. *Cell Mol Life Sci*, 65: 1566-1584.
61. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, Chewaskulyong B, Jiang H, Duffield EL, Watkins CL, Armour AA, Fukuoka M. (2009) Gefitinib



- or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med*, 361: 947-957.
62. Zhou C, Wu YL, Chen G, Feng J, Liu XQ, Wang C, Zhang S, Wang J, Zhou S, Ren S, Lu S, Zhang L, Hu C, Hu C, Luo Y, Chen L, Ye M, Huang J, Zhi X, Zhang Y, Xiu Q, Ma J, Zhang L, You C. (2011) Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol*, 12: 735-742.
  63. Belani CP. (2010) The role of irreversible EGFR inhibitors in the treatment of non-small cell lung cancer: overcoming resistance to reversible EGFR inhibitors. *Cancer Invest*, 28: 413-423.
  64. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, Maki RG, Ou SH, Dezube BJ, Janne PA, Costa DB, Varella-Garcia M, Kim WH, Lynch TJ, Fidias P, Stubbs H, Engelman JA, Sequist LV, Tan W, Gandhi L, Mino-Kenudson M, Wei GC, Shreeve SM, Ratain MJ, Settleman J, Christensen JG, Haber DA, Wilner K, Salgia R, Shapiro GI, Clark JW, Iafrate AJ. (2010) Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med*, 363: 1693-1703.
  65. Takahashi T, Sonobe M, Kobayashi M, Yoshizawa A, Menju T, Nakayama E, Mino N, Iwakiri S, Sato K, Miyahara R, Okubo K, Manabe T, Date H. (2010) Clinicopathologic features of non-small-cell lung cancer with EML4-ALK fusion gene. *Ann Surg Oncol*, 17: 889-897.
  66. Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS, Solomon B, Stubbs H, Admane S, McDermott U, Settleman J, Kobayashi S, Mark EJ, Rodig SJ, Chirieac LR, Kwak EL, Lynch TJ, Iafrate AJ. (2009) Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J Clin Oncol*, 27: 4247-4253.
  67. Paik PK, Arcila ME, Fara M, Sima CS, Miller VA, Kris MG, Ladanyi M, Riely GJ. (2011) Clinical characteristics of patients with lung adenocarcinomas harboring BRAF mutations. *J Clin Oncol*, 29: 2046-2051.
  68. Gyorffy B, Schafer R. (2010) Biomarkers downstream of RAS: a search for robust transcriptional targets. *Curr Cancer Drug Targets*, 10: 858-868.

69. Karnoub AE, Weinberg RA. (2008) Ras oncogenes: split personalities. *Nat Rev Mol Cell Biol*, 9: 517-531.
70. Vasani N, Boyer JL, Herbst RS. (2014) A RAS renaissance: emerging targeted therapies for KRAS-mutated non-small cell lung cancer. *Clin Cancer Res*, 20: 3921-3930.
71. Riely GJ, Marks J, Pao W. (2009) KRAS mutations in non-small cell lung cancer. *Proc Am Thorac Soc*, 6: 201-205.
72. Andreyev HJ, Norman AR, Cunningham D, Oates JR, Clarke PA. (1998) Kirsten ras mutations in patients with colorectal cancer: the multicenter "RASCAL" study. *J Natl Cancer Inst*, 90: 675-684.
73. Ogino S, Nosho K, Kirkner GJ, Kawasaki T, Meyerhardt JA, Loda M, Giovannucci EL, Fuchs CS. (2009) CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut*, 58: 90-96.
74. Scoccianti C, Vesin A, Martel G, Olivier M, Brambilla E, Timsit JF, Tavecchio L, Brambilla C, Field JK, Hainaut P, European Early Lung Cancer C. (2012) Prognostic value of TP53, KRAS and EGFR mutations in nonsmall cell lung cancer: the EUELC cohort. *Eur Respir J*, 40: 177-184.
75. Meng D, Yuan M, Li X, Chen L, Yang J, Zhao X, Ma W, Xin J. (2013) Prognostic value of K-RAS mutations in patients with non-small cell lung cancer: a systematic review with meta-analysis. *Lung Cancer*, 81: 1-10.
76. Shigematsu H, Gazdar AF. (2006) Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int J Cancer*, 118: 257-262.
77. Massarelli E, Varella-Garcia M, Tang X, Xavier AC, Ozburn NC, Liu DD, Bekele BN, Herbst RS, Wistuba, II. (2007) KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. *Clin Cancer Res*, 13: 2890-2896.
78. Mao C, Qiu LX, Liao RY, Du FB, Ding H, Yang WC, Li J, Chen Q. (2010) KRAS mutations and resistance to EGFR-TKIs treatment in patients with non-small cell lung cancer: a meta-analysis of 22 studies. *Lung Cancer*, 69: 272-278.
79. Davies BR, Logie A, McKay JS, Martin P, Steele S, Jenkins R, Cockerill M, Cartlidge S, Smith PD. (2007) AZD6244 (ARRY-142886), a potent inhibitor of

- mitogen-activated protein kinase/extracellular signal-regulated kinase kinase 1/2 kinases: mechanism of action in vivo, pharmacokinetic/pharmacodynamic relationship, and potential for combination in preclinical models. *Mol Cancer Ther*, 6: 2209-2219.
80. Janne PA, Shaw AT, Pereira JR, Jeannin G, Vansteenkiste J, Barrios C, Franke FA, Grinsted L, Zazulina V, Smith P, Smith I, Crino L. (2013) Selumetinib plus docetaxel for KRAS-mutant advanced non-small-cell lung cancer: a randomised, multicentre, placebo-controlled, phase 2 study. *Lancet Oncol*, 14: 38-47.
  81. Perz JF, Armstrong GL, Farrington LA, Hutin YJ, Bell BP. (2006) The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol*, 45: 529-538.
  82. Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, Washington KM, Carneiro F, Cree IA, Board WHOCOTE. (2020) The 2019 WHO classification of tumours of the digestive system. *Histopathology*, 76: 182-188.
  83. Yang JD, Roberts LR. (2010) Hepatocellular carcinoma: A global view. *Nat Rev Gastroenterol Hepatol*, 7: 448-458.
  84. Sia D, Villanueva A, Friedman SL, Llovet JM. (2017) Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. *Gastroenterology*, 152: 745-761.
  85. Schulze K, Nault JC, Villanueva A. (2016) Genetic profiling of hepatocellular carcinoma using next-generation sequencing. *J Hepatol*, 65: 1031-1042.
  86. Llovet JM, Villanueva A, Lachenmayer A, Finn RS. (2015) Advances in targeted therapies for hepatocellular carcinoma in the genomic era. *Nat Rev Clin Oncol*, 12: 408-424.
  87. Zucman-Rossi J, Villanueva A, Nault JC, Llovet JM. (2015) Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology*, 149: 1226-1239 e1224.
  88. Schulze K, Imbeaud S, Letouze E, Alexandrov LB, Calderaro J, Rebouissou S, Couchy G, Meiller C, Shinde J, Soysouvanh F, Calatayud AL, Pinyol R, Pelletier L, Balabaud C, Laurent A, Blanc JF, Mazzaferro V, Calvo F, Villanueva A, Nault JC, Bioulac-Sage P, Stratton MR, Llovet JM, Zucman-Rossi J. (2015) Exome

- sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet*, 47: 505-511.
89. Totoki Y, Tatsuno K, Covington KR, Ueda H, Creighton CJ, Kato M, Tsuji S, Donehower LA, Slagle BL, Nakamura H, Yamamoto S, Shinbrot E, Hama N, Lehmkuhl M, Hosoda F, Arai Y, Walker K, Dahdouli M, Gotoh K, Nagae G, Gingras MC, Muzny DM, Ojima H, Shimada K, Midorikawa Y, Goss JA, Cotton R, Hayashi A, Shibahara J, Ishikawa S, Guiteau J, Tanaka M, Urushidate T, Ohashi S, Okada N, Doddapaneni H, Wang M, Zhu Y, Dinh H, Okusaka T, Kokudo N, Kosuge T, Takayama T, Fukayama M, Gibbs RA, Wheeler DA, Aburatani H, Shibata T. (2014) Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nat Genet*, 46: 1267-1273.
  90. Huynh H, Hao HX, Chan SL, Chen D, Ong R, Soo KC, Pochanard P, Yang D, Ruddy D, Liu M, Derti A, Balak MN, Palmer MR, Wang Y, Lee BH, Sellami D, Zhu AX, Schlegel R, Huang A. (2015) Loss of Tuberous Sclerosis Complex 2 (TSC2) Is Frequent in Hepatocellular Carcinoma and Predicts Response to mTORC1 Inhibitor Everolimus. *Mol Cancer Ther*, 14: 1224-1235.
  91. Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y, Lee NP, Lee WH, Ariyaratne PN, Tennakoon C, Mulawadi FH, Wong KF, Liu AM, Poon RT, Fan ST, Chan KL, Gong Z, Hu Y, Lin Z, Wang G, Zhang Q, Barber TD, Chou WC, Aggarwal A, Hao K, Zhou W, Zhang C, Hardwick J, Buser C, Xu J, Kan Z, Dai H, Mao M, Reinhard C, Wang J, Luk JM. (2012) Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet*, 44: 765-769.
  92. Nault JC, Datta S, Imbeaud S, Franconi A, Mallet M, Couchy G, Letouze E, Pilati C, Verret B, Blanc JF, Balabaud C, Calderaro J, Laurent A, Letexier M, Bioulac-Sage P, Calvo F, Zucman-Rossi J. (2015) Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat Genet*, 47: 1187-1193.
  93. Fu LL, Wen X, Bao JK, Liu B. (2012) MicroRNA-modulated autophagic signaling networks in cancer. *Int J Biochem Cell Biol*, 44: 733-736.
  94. Iorio MV, Croce CM. (2012) MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol Med*, 4: 143-159.

95. Vosa U, Vooder T, Kolde R, Fischer K, Valk K, Tonisson N, Roosipuu R, Vilo J, Metspalu A, Annilo T. (2011) Identification of miR-374a as a prognostic marker for survival in patients with early-stage nonsmall cell lung cancer. *Genes Chromosomes Cancer*, 50: 812-822.
96. Akcakaya P, Ekelund S, Kolosenko I, Caramuta S, Ozata DM, Xie H, Lindfors U, Olivecrona H, Lui WO. (2011) miR-185 and miR-133b deregulation is associated with overall survival and metastasis in colorectal cancer. *Int J Oncol*, 39: 311-318.
97. Adams BD, Wali VB, Cheng CJ, Inukai S, Booth CJ, Agarwal S, Rimm DL, Gyorffy B, Santarpia L, Pusztai L, Saltzman WM, Slack FJ. (2016) miR-34a Silences c-SRC to Attenuate Tumor Growth in Triple-Negative Breast Cancer. *Cancer Res*, 76: 927-939.
98. Ouyang H, Gore J, Deitz S, Korc M. (2014) microRNA-10b enhances pancreatic cancer cell invasion by suppressing TIP30 expression and promoting EGF and TGF-beta actions. *Oncogene*, 33: 4664-4674.
99. Wong CM, Wong CC, Lee JM, Fan DN, Au SL, Ng IO. (2012) Sequential alterations of microRNA expression in hepatocellular carcinoma development and venous metastasis. *Hepatology*, 55: 1453-1461.
100. Wei R, Huang GL, Zhang MY, Li BK, Zhang HZ, Shi M, Chen XQ, Huang L, Zhou QM, Jia WH, Zheng XF, Yuan YF, Wang HY. (2013) Clinical significance and prognostic value of microRNA expression signatures in hepatocellular carcinoma. *Clin Cancer Res*, 19: 4780-4791.
101. Zucman-Rossi J. (2010) Molecular classification of hepatocellular carcinoma. *Dig Liver Dis*, 42 Suppl 3: S235-241.
102. Fornari F, Ferracin M, Trere D, Milazzo M, Marinelli S, Galassi M, Venerandi L, Pollutri D, Patrizi C, Borghi A, Foschi FG, Stefanini GF, Negrini M, Bolondi L, Gramantieri L. (2015) Circulating microRNAs, miR-939, miR-595, miR-519d and miR-494, Identify Cirrhotic Patients with HCC. *PLoS One*, 10: e0141448.
103. Minguez B, Hoshida Y, Villanueva A, Toffanin S, Cabellos L, Thung S, Mandeli J, Sia D, April C, Fan JB, Lachenmayer A, Savic R, Roayaie S, Mazzaferro V, Bruix J, Schwartz M, Friedman SL, Llovet JM. (2011) Gene-expression signature of vascular invasion in hepatocellular carcinoma. *J Hepatol*, 55: 1325-1331.

104. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, de Oliveira AC, Santoro A, Raoul JL, Forner A, Schwartz M, Porta C, Zeuzem S, Bolondi L, Greten TF, Galle PR, Seitz JF, Borbath I, Haussinger D, Giannaris T, Shan M, Moscovici M, Voliotis D, Bruix J, Group SIS. (2008) Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*, 359: 378-390.
105. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, Luo R, Feng J, Ye S, Yang TS, Xu J, Sun Y, Liang H, Liu J, Wang J, Tak WY, Pan H, Burock K, Zou J, Voliotis D, Guan Z. (2009) Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. *Lancet Oncol*, 10: 25-34.
106. Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, Baron A, Park JW, Han G, Jassem J, Blanc JF, Vogel A, Komov D, Evans TRJ, Lopez C, Dutcus C, Guo M, Saito K, Kraljevic S, Tamai T, Ren M, Cheng AL. (2018) Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet*, 391: 1163-1173.
107. Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, Pracht M, Yokosuka O, Rosmorduc O, Breder V, Gerolami R, Masi G, Ross PJ, Song T, Bronowicki JP, Ollivier-Hourmand I, Kudo M, Cheng AL, Llovet JM, Finn RS, LeBerre MA, Baumhauer A, Meinhardt G, Han G, Investigators R. (2017) Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*, 389: 56-66.
108. Bai S, Nasser MW, Wang B, Hsu SH, Datta J, Kutay H, Yadav A, Nuovo G, Kumar P, Ghoshal K. (2009) MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib. *J Biol Chem*, 284: 32015-32027.
109. Yang F, Li QJ, Gong ZB, Zhou L, You N, Wang S, Li XL, Li JJ, An JZ, Wang DS, He Y, Dou KF. (2014) MicroRNA-34a targets Bcl-2 and sensitizes human hepatocellular carcinoma cells to sorafenib treatment. *Technol Cancer Res Treat*, 13: 77-86.

110. Krutzfeldt J, Kuwajima S, Braich R, Rajeev KG, Pena J, Tuschl T, Manoharan M, Stoffel M. (2007) Specificity, duplex degradation and subcellular localization of antagomirs. *Nucleic Acids Res*, 35: 2885-2892.
111. Tomimaru Y, Eguchi H, Nagano H, Wada H, Tomokuni A, Kobayashi S, Marubashi S, Takeda Y, Tanemura M, Umeshita K, Doki Y, Mori M. (2010) MicroRNA-21 induces resistance to the anti-tumour effect of interferon-alpha/5-fluorouracil in hepatocellular carcinoma cells. *Br J Cancer*, 103: 1617-1626.
112. Callegari E, Elamin BK, Giannone F, Milazzo M, Altavilla G, Fornari F, Giacomelli L, D'Abundo L, Ferracin M, Bassi C, Zagatti B, Corra F, Miotto E, Lupini L, Bolondi L, Gramantieri L, Croce CM, Sabbioni S, Negrini M. (2012) Liver tumorigenicity promoted by microRNA-221 in a mouse transgenic model. *Hepatology*, 56: 1025-1033.
113. Tomokuni A, Eguchi H, Tomimaru Y, Wada H, Kawamoto K, Kobayashi S, Marubashi S, Tanemura M, Nagano H, Mori M, Doki Y. (2011) miR-146a suppresses the sensitivity to interferon-alpha in hepatocellular carcinoma cells. *Biochem Biophys Res Commun*, 414: 675-680.
114. Ji J, Yu L, Yu Z, Forgues M, Uenishi T, Kubo S, Wakasa K, Zhou J, Fan J, Tang ZY, Fu S, Zhu H, Jin JG, Sun HC, Wang XW. (2013) Development of a miR-26 companion diagnostic test for adjuvant interferon-alpha therapy in hepatocellular carcinoma. *Int J Biol Sci*, 9: 303-312.
115. Iwatsuki S, Dvorchik I, Marsh JW, Madariaga JR, Carr B, Fung JJ, Starzl TE. (2000) Liver transplantation for hepatocellular carcinoma: a proposal of a prognostic scoring system. *J Am Coll Surg*, 191: 389-394.
116. Zavaglia C, De Carlis L, Alberti AB, Minola E, Belli LS, Slim AO, Airoidi A, Giacomoni A, Rondinara G, Tinelli C, Forti D, Pinzello G. (2005) Predictors of long-term survival after liver transplantation for hepatocellular carcinoma. *Am J Gastroenterol*, 100: 2708-2716.
117. Yuki K, Hirohashi S, Sakamoto M, Kanai T, Shimosato Y. (1990) Growth and spread of hepatocellular carcinoma. A review of 240 consecutive autopsy cases. *Cancer*, 66: 2174-2179.
118. Cancer Genome Atlas Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, Robertson A, Hoadley K, Triche TJ, Jr., Laird PW, Baty JD, Fulton

- LL, Fulton R, Heath SE, Kalicki-Veizer J, Kandoth C, Klco JM, Koboldt DC, Kanchi KL, Kulkarni S, Lamprecht TL, Larson DE, Lin L, Lu C, McLellan MD, McMichael JF, Payton J, Schmidt H, Spencer DH, Tomasson MH, Wallis JW, Wartman LD, Watson MA, Welch J, Wendl MC, Ally A, Balasundaram M, Birol I, Butterfield Y, Chiu R, Chu A, Chuah E, Chun HJ, Corbett R, Dhalla N, Guin R, He A, Hirst C, Hirst M, Holt RA, Jones S, Karsan A, Lee D, Li HI, Marra MA, Mayo M, Moore RA, Mungall K, Parker J, Pleasance E, Plettner P, Schein J, Stoll D, Swanson L, Tam A, Thiessen N, Varhol R, Wye N, Zhao Y, Gabriel S, Getz G, Sougnez C, Zou L, Leiserson MD, Vandin F, Wu HT, Applebaum F, Baylin SB, Akbani R, Broom BM, Chen K, Motter TC, Nguyen K, Weinstein JN, Zhang N, Ferguson ML, Adams C, Black A, Bowen J, Gastier-Foster J, Grossman T, Lichtenberg T, Wise L, Davidsen T, Demchok JA, Shaw KR, Sheth M, Sofia HJ, Yang L, Downing JR, Eley G. (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*, 368: 2059-2074.
119. Shlush LI, Zandi S, Mitchell A, Chen WC, Brandwein JM, Gupta V, Kennedy JA, Schimmer AD, Schuh AC, Yee KW, McLeod JL, Doedens M, Medeiros JJ, Marke R, Kim HJ, Lee K, McPherson JD, Hudson TJ, Consortium HP-LGP, Brown AM, Yousif F, Trinh QM, Stein LD, Minden MD, Wang JC, Dick JE. (2014) Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*, 506: 328-333.
120. Rombouts WJ, Blokland I, Lowenberg B, Ploemacher RE. (2000) Biological characteristics and prognosis of adult acute myeloid leukemia with internal tandem duplications in the *Flt3* gene. *Leukemia*, 14: 675-683.
121. Falini B, Mecucci C, Tiacci E, Alcalay M, Rosati R, Pasqualucci L, La Starza R, Diverio D, Colombo E, Santucci A, Bigerna B, Pacini R, Pucciarini A, Liso A, Vignetti M, Fazi P, Meani N, Pettrossi V, Saglio G, Mandelli F, Lo-Coco F, Pelicci PG, Martelli MF, Party GALW. (2005) Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med*, 352: 254-266.
122. Grisendi S, Mecucci C, Falini B, Pandolfi PP. (2006) Nucleophosmin and cancer. *Nat Rev Cancer*, 6: 493-505.



123. Martelli MP, Manes N, Pettirossi V, Liso A, Pacini R, Mannucci R, Zei T, Bolli N, di Raimondo F, Specchia G, Nicoletti I, Martelli MF, Falini B. (2008) Absence of nucleophosmin leukaemic mutants in B and T cells from AML with NPM1 mutations: implications for the cell of origin of NPMc+ AML. *Leukemia*, 22: 195-198.
124. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, Potter NE, Heuser M, Thol F, Bolli N, Gundem G, Van Loo P, Martincorena I, Ganly P, Mudie L, McLaren S, O'Meara S, Raine K, Jones DR, Teague JW, Butler AP, Greaves MF, Ganser A, Dohner K, Schlenk RF, Dohner H, Campbell PJ. (2016) Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*, 374: 2209-2221.
125. Schnittger S, Schoch C, Kern W, Mecucci C, Tschulik C, Martelli MF, Haferlach T, Hiddemann W, Falini B. (2005) Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. *Blood*, 106: 3733-3739.
126. Rollig C, Thiede C, Gramatzki M, Aulitzky W, Bodenstern H, Bornhauser M, Platzbecker U, Stuhlmann R, Schuler U, Soucek S, Kramer M, Mohr B, Oelschlaegel U, Stolzel F, von Bonin M, Wermke M, Wandt H, Ehninger G, Schaich M, Study Alliance L. (2010) A novel prognostic model in elderly patients with acute myeloid leukemia: results of 909 patients entered into the prospective AML96 trial. *Blood*, 116: 971-978.
127. Paschka P, Schlenk RF, Gaidzik VI, Habdank M, Kronke J, Bullinger L, Spath D, Kayser S, Zucknick M, Gotze K, Horst HA, Germing U, Dohner H, Dohner K. (2010) IDH1 and IDH2 mutations are frequent genetic alterations in acute myeloid leukemia and confer adverse prognosis in cytogenetically normal acute myeloid leukemia with NPM1 mutation without FLT3 internal tandem duplication. *J Clin Oncol*, 28: 3636-3643.
128. Gong Q, Zhou L, Xu S, Li X, Zou Y, Chen J. (2015) High Doses of Daunorubicin during Induction Therapy of Newly Diagnosed Acute Myeloid Leukemia: A Systematic Review and Meta-Analysis of Prospective Clinical Trials. *PLoS One*, 10: e0125612.

129. Lowenberg B. (2013) Sense and nonsense of high-dose cytarabine for acute myeloid leukemia. *Blood*, 121: 26-28.
130. Zhang W, Konopleva M, Shi YX, McQueen T, Harris D, Ling X, Estrov Z, Quintas-Cardama A, Small D, Cortes J, Andreeff M. (2008) Mutant FLT3: a direct target of sorafenib in acute myelogenous leukemia. *J Natl Cancer Inst*, 100: 184-198.
131. Fischer T, Stone RM, Deangelo DJ, Galinsky I, Estey E, Lanza C, Fox E, Ehninger G, Feldman EJ, Schiller GJ, Klimek VM, Nimer SD, Gilliland DG, Dutreix C, Huntsman-Labed A, Virkus J, Giles FJ. (2010) Phase IIB trial of oral Midostaurin (PKC412), the FMS-like tyrosine kinase 3 receptor (FLT3) and multi-targeted kinase inhibitor, in patients with acute myeloid leukemia and high-risk myelodysplastic syndrome with either wild-type or mutated FLT3. *J Clin Oncol*, 28: 4339-4345.
132. Wander SA, Levis MJ, Fathi AT. (2014) The evolving role of FLT3 inhibitors in acute myeloid leukemia: quizartinib and beyond. *Ther Adv Hematol*, 5: 65-77.
133. Smith CC, Lasater EA, Lin KC, Wang Q, McCreery MQ, Stewart WK, Damon LE, Perl AE, Jeschke GR, Sugita M, Carroll M, Kogan SC, Kuriyan J, Shah NP. (2014) Crenolanib is a selective type I pan-FLT3 inhibitor. *Proc Natl Acad Sci U S A*, 111: 5319-5324.
134. Redell MS, Ruiz MJ, Alonzo TA, Gerbing RB, Tweardy DJ. (2011) Stat3 signaling in acute myeloid leukemia: ligand-dependent and -independent activation and induction of apoptosis by a novel small-molecule Stat3 inhibitor. *Blood*, 117: 5701-5709.
135. Oh DY, Lee SH, Han SW, Kim MJ, Kim TM, Kim TY, Heo DS, Yuasa M, Yanagihara Y, Bang YJ. (2015) Phase I Study of OPB-31121, an Oral STAT3 Inhibitor, in Patients with Advanced Solid Tumors. *Cancer Res Treat*, 47: 607-615.
136. Abdel-Wahab O, Levine RL. (2013) Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia. *Blood*, 121: 3563-3572.
137. Dawson MA, Kouzarides T, Huntly BJ. (2012) Targeting epigenetic readers in cancer. *N Engl J Med*, 367: 647-657.

138. Etchin J, Sanda T, Mansour MR, Kentsis A, Montero J, Le BT, Christie AL, McCauley D, Rodig SJ, Kauffman M, Shacham S, Stone R, Letai A, Kung AL, Thomas Look A. (2013) KPT-330 inhibitor of CRM1 (XPO1)-mediated nuclear export has selective anti-leukaemic activity in preclinical models of T-cell acute lymphoblastic leukaemia and acute myeloid leukaemia. *Br J Haematol*, 161: 117-127.
139. Gasiorowski RE, Clark GJ, Bradstock K, Hart DN. (2014) Antibody therapy for acute myeloid leukaemia. *Br J Haematol*, 164: 481-495.
140. Maxam AM, Gilbert W. (1977) A new method for sequencing DNA. *Proc Natl Acad Sci U S A*, 74: 560-564.
141. Sanger F, Nicklen S, Coulson AR. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74: 5463-5467.
142. Nyren P. (1987) Enzymatic method for continuous monitoring of DNA polymerase activity. *Anal Biochem*, 167: 235-238.
143. Smith LM, Fung S, Hunkapiller MW, Hunkapiller TJ, Hood LE. (1985) The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic Acids Res*, 13: 2399-2412.
144. Bakker E. (2006) Is the DNA sequence the gold standard in genetic testing? Quality of molecular genetic tests assessed. *Clin Chem*, 52: 557-558.
145. Hollstein M, Sidransky D, Vogelstein B, Harris CC. (1991) p53 mutations in human cancers. *Science*, 253: 49-53.
146. Shendure J, Ji H. (2008) Next-generation DNA sequencing. *Nat Biotechnol*, 26: 1135-1145.
147. Adessi C, Matton G, Ayala G, Turcatti G, Mermod JJ, Mayer P, Kawashima E. (2000) Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res*, 28: E87.
148. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. (2003) Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc Natl Acad Sci U S A*, 100: 8817-8822.

149. Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. (2006) BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res*, 34: e22.
150. Turcatti G, Romieu A, Fedurco M, Tairi AP. (2008) A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Res*, 36: e25.
151. Zhang M, Liu YH, Chang CS, Zhi H, Wang S, Xu W, Smith CW, Zhang HB. (2019) Quantification of gene expression while taking into account RNA alternative splicing. *Genomics*, 111: 1517-1528.
152. Rajeevan MS, Ranamukhaarachchi DG, Vernon SD, Unger ER. (2001) Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. *Methods*, 25: 443-451.
153. Tarca AL, Romero R, Draghici S. (2006) Analysis of microarray experiments of gene expression profiling. *Am J Obstet Gynecol*, 195: 373-388.
154. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M. (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, 320: 1344-1349.
155. Mullis KB, Faloona FA. (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*, 155: 335-350.
156. Morrison TB, Weis JJ, Wittwer CT. (1998) Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques*, 24: 954-958, 960, 962.
157. Hui L, DelMonte T, Ranade K. (2008) Genotyping using the TaqMan assay. *Curr Protoc Hum Genet*, Chapter 2: Unit 2 10.
158. Slonim DK, Yanai I. (2009) Getting started in gene expression microarray analysis. *PLoS Comput Biol*, 5: e1000543.
159. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*, 18: 1509-1517.
160. Boguski MS, Tolstoshev CM, Bassett DE, Jr. (1994) Gene discovery in dbEST. *Science*, 265: 1993-1994.

161. Wang Z, Gerstein M, Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, 10: 57-63.
162. Anders S, Pyl PT, Huber W. (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31: 166-169.
163. Whitley SK, Horne WT, Kolls JK. (2016) Research Techniques Made Simple: Methodology and Clinical Applications of RNA Sequencing. *J Invest Dermatol*, 136: e77-e82.
164. Anders S, Huber W. (2010) Differential expression analysis for sequence count data. *Genome Biol*, 11: R106.
165. Fekete T, Raso E, Pete I, Tegze B, Liko I, Munkacsy G, Sipos N, Rigo J, Jr., Gyorffy B. (2012) Meta-analysis of gene expression profiles associated with histological classification and survival in 829 ovarian cancer samples. *Int J Cancer*, 131: 95-105.
166. Gyorffy B, Surowiak P, Budczies J, Lanczky A. (2013) Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One*, 8: e82241.
167. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC. (2011) Jetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics*, 12: 474.
168. Verhaak RG, Wouters BJ, Erpelinck CA, Abbas S, Beverloo HB, Lugthart S, Lowenberg B, Delwel R, Valk PJ. (2009) Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. *Haematologica*, 94: 131-134.
169. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B, Delwel R. (2004) Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med*, 350: 1617-1628.
170. Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, Cornella H, Liberzon A, Kobayashi M, Kumada H, Thung SN, Bruix J, Newell P, April C, Fan JB, Roayaie S, Mazzaferro V, Schwartz ME, Llovet JM. (2011) Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. *Gastroenterology*, 140: 1501-1512 e1502.

171. Li W, Xie L, He X, Li J, Tu K, Wei L, Wu J, Guo Y, Ma X, Zhang P, Pan Z, Hu X, Zhao Y, Xie H, Jiang G, Chen T, Wang J, Zheng S, Cheng J, Wan D, Yang S, Li Y, Gu J. (2008) Diagnostic and prognostic implications of microRNAs in human hepatocellular carcinoma. *Int J Cancer*, 123: 1616-1622.
172. Budhu A, Jia HL, Forgues M, Liu CG, Goldstein D, Lam A, Zanetti KA, Ye QH, Qin LX, Croce CM, Tang ZY, Wang XW. (2008) Identification of metastasis-related microRNAs in hepatocellular carcinoma. *Hepatology*, 47: 897-907.
173. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*, 31: 213-219.
174. Pongor L, Kormos M, Hatzis C, Pusztai L, Szabo A, Gyorffy B. (2015) A genome-wide approach to link genotype to clinical outcome by utilizing next generation sequencing and gene chip data of 6,697 breast cancer patients. *Genome Med*, 7: 104.
175. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ. (2015) COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res*, 43: D805-811.
176. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*, 6: 80-92.
177. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, 43: e47.
178. Gyorffy B, Gyorffy A, Tulassay Z. (2005) [The problem of multiple testing and solutions for genome-wide studies]. *Orv Hetil*, 146: 559-563.
179. Mihaly Z, Kormos M, Lanczky A, Dank M, Budczies J, Szasz MA, Gyorffy B. (2013) A meta-analysis of gene expression-based biomarkers predicting outcome after tamoxifen treatment in breast cancer. *Breast Cancer Res Treat*, 140: 219-232.

180. Menyhart O, Harami-Papp H, Sukumar S, Schafer R, Magnani L, de Barrios O, Gyorffy B. (2016) Guidelines for the selection of functional assays to evaluate the hallmarks of cancer. *Biochim Biophys Acta*, 1866: 300-319.
181. Walter RB, Othus M, Burnett AK, Lowenberg B, Kantarjian HM, Ossenkoppele GJ, Hills RK, van Montfort KG, Ravandi F, Evans A, Pierce SR, Appelbaum FR, Estey EH. (2013) Significance of FAB subclassification of "acute myeloid leukemia, NOS" in the 2008 WHO classification: analysis of 5848 newly diagnosed patients. *Blood*, 121: 2424-2431.
182. Mann RS, Affolter M. (1998) Hox proteins meet more partners. *Curr Opin Genet Dev*, 8: 423-429.
183. Chiang DY, Villanueva A, Hoshida Y, Peix J, Newell P, Minguez B, LeBlanc AC, Donovan DJ, Thung SN, Sole M, Tovar V, Alsinet C, Ramos AH, Barretina J, Roayaie S, Schwartz M, Waxman S, Bruix J, Mazzaferro V, Ligon AH, Najfeld V, Friedman SL, Sellers WR, Meyerson M, Llovet JM. (2008) Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. *Cancer Res*, 68: 6779-6788.
184. Huang da W, Sherman BT, Lempicki RA. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4: 44-57.
185. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. (2017) PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*, 45: D183-D189.
186. Georgoulas V, Kouroussis C, Agelidou A, Boukovinas I, Palamidas P, Stavriniadis E, Polyzos A, Syrigos K, Veslemes M, Toubis M, Ardavanis A, Tselepatiotis E, Vlachonikolis I, Lung Cancer Committee of the Hellenic Oncology Research G. (2004) Irinotecan plus gemcitabine vs irinotecan for the second-line treatment of patients with advanced non-small-cell lung cancer pretreated with docetaxel and cisplatin: a multicentre, randomised, phase II study. *Br J Cancer*, 91: 482-488.
187. Colicelli J. (2010) ABL tyrosine kinases: evolution of function, regulation, and specificity. *Sci Signal*, 3: re6.

188. Sos ML, Michel K, Zander T, Weiss J, Frommolt P, Peifer M, Li D, Ullrich R, Koker M, Fischer F, Shimamura T, Rauh D, Mermel C, Fischer S, Stuckrath I, Heynck S, Beroukhim R, Lin W, Winckler W, Shah K, LaFramboise T, Moriarty WF, Hanna M, Tolosi L, Rahnenfuhrer J, Verhaak R, Chiang D, Getz G, Hellmich M, Wolf J, Girard L, Peyton M, Weir BA, Chen TH, Greulich H, Barretina J, Shapiro GI, Garraway LA, Gazdar AF, Minna JD, Meyerson M, Wong KK, Thomas RK. (2009) Predicting drug susceptibility of non-small cell lung cancers based on genetic lesions. *J Clin Invest*, 119: 1727-1740.
189. Greuber EK, Smith-Pearson P, Wang J, Pendergast AM. (2013) Role of ABL family kinases in cancer: from leukaemia to solid tumours. *Nat Rev Cancer*, 13: 559-571.
190. Hunter FW, Jaiswal JK, Hurley DG, Liyanage HD, McManaway SP, Gu Y, Richter S, Wang J, Tercel M, Print CG, Wilson WR, Pruijn FB. (2014) The flavoprotein FOXRED2 reductively activates nitro-chloromethylbenzindolines and other hypoxia-targeting prodrugs. *Biochem Pharmacol*, 89: 224-235.
191. Zardavas D, Baselga J, Piccart M. (2013) Emerging targeted agents in metastatic breast cancer. *Nat Rev Clin Oncol*, 10: 191-210.
192. Azpiazu N, Morata G. (1998) Functional and regulatory interactions between Hox and extradenticle genes. *Genes Dev*, 12: 261-273.
193. Chang CP, Shen WF, Rozenfeld S, Lawrence HJ, Largman C, Cleary ML. (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev*, 9: 663-674.
194. Drabkin HA, Parsy C, Ferguson K, Guilhot F, Lacotte L, Roy L, Zeng C, Baron A, Hunger SP, Varella-Garcia M, Gemmill R, Brizard F, Brizard A, Roche J. (2002) Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. *Leukemia*, 16: 186-195.
195. Mullighan CG, Kennedy A, Zhou X, Radtke I, Phillips LA, Shurtleff SA, Downing JR. (2007) Pediatric acute myeloid leukemia with NPM1 mutations is characterized by a gene expression profile with dysregulated HOX gene expression distinct from MLL-rearranged leukemias. *Leukemia*, 21: 2000-2009.



196. Rau R, Brown P. (2009) Nucleophosmin (NPM1) mutations in adult and childhood acute myeloid leukaemia: towards definition of a new leukaemia entity. *Hematol Oncol*, 27: 171-181.
197. Verhaak RG, Goudswaard CS, van Putten W, Bijl MA, Sanders MA, Hagens W, Uitterlinden AG, Erpelinck CA, Delwel R, Lowenberg B, Valk PJ. (2005) Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. *Blood*, 106: 3747-3754.
198. Namkoong S, Lee KI, Lee JI, Park R, Lee EJ, Jang IS, Park J. (2015) The integral membrane protein ITM2A, a transcriptional target of PKA-CREB, regulates autophagic flux via interaction with the vacuolar ATPase. *Autophagy*, 11: 756-768.
199. Alharbi RA, Pettengell R, Pandha HS, Morgan R. (2013) The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia*, 27: 1000-1008.
200. Morgan R, Pirard PM, Shears L, Sohal J, Pettengell R, Pandha HS. (2007) Antagonism of HOX/PBX dimer formation blocks the in vivo proliferation of melanoma. *Cancer Res*, 67: 5806-5813.
201. Kachgal S, Mace KA, Boudreau NJ. (2012) The dual roles of homeobox genes in vascularization and wound healing. *Cell Adh Migr*, 6: 457-470.
202. Hong CS, Jeong O, Piao Z, Guo C, Jung MR, Choi C, Park YK. (2015) HOXB5 induces invasion and migration through direct transcriptional up-regulation of beta-catenin in human gastric carcinoma. *Biochem J*, 472: 393-403.
203. Jin K, Sukumar S. (2016) HOX genes: Major actors in resistance to selective endocrine response modifiers. *Biochim Biophys Acta*, 1865: 105-110.
204. Morgan R, El-Tanani M, Hunter KD, Harrington KJ, Pandha HS. (2017) Targeting HOX/PBX dimers in cancer. *Oncotarget*, 8: 32322-32331.
205. Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, Marks J, Sukumar S. (2000) Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature*, 405: 974-978.

206. Plowright L, Harrington KJ, Pandha HS, Morgan R. (2009) HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). *Br J Cancer*, 100: 470-475.
207. Morgan R, Boxall A, Harrington KJ, Simpson GR, Gillett C, Michael A, Pandha HS. (2012) Targeting the HOX/PBX dimer in breast cancer. *Breast Cancer Res Treat*, 136: 389-398.
208. Morgan R, Boxall A, Harrington KJ, Simpson GR, Michael A, Pandha HS. (2014) Targeting HOX transcription factors in prostate cancer. *BMC Urol*, 14: 17.
209. Li Z, Zhang Z, Li Y, Arnovitz S, Chen P, Huang H, Jiang X, Hong GM, Kunjamma RB, Ren H, He C, Wang CZ, Elkahloun AG, Valk PJ, Dohner K, Neilly MB, Bullinger L, Delwel R, Lowenberg B, Liu PP, Morgan R, Rowley JD, Yuan CS, Chen J. (2013) PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood*, 121: 1422-1431.
210. Alharbi RA, Pandha HS, Simpson GR, Pettengell R, Poterlowicz K, Thompson A, Harrington K, El-Tanani M, Morgan R. (2017) Inhibition of HOX/PBX dimer formation leads to necroptosis in acute myeloid leukemia cells. *Oncotarget*, 8: 89566-89579.
211. Song J, Gao L, Yang G, Tang S, Xie H, Wang Y, Wang J, Zhang Y, Jin J, Gou Y, Yang Z, Chen Z, Wu K, Liu J, Fan D. (2014) MiR-199a regulates cell proliferation and survival by targeting FZD7. *PLoS One*, 9: e110074.
212. Sun F, Fu H, Liu Q, Tie Y, Zhu J, Xing R, Sun Z, Zheng X. (2008) Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. *FEBS Lett*, 582: 1564-1568.
213. Mita AC, Mita MM, Nawrocki ST, Giles FJ. (2008) Survivin: key regulator of mitosis and apoptosis and novel target for cancer therapeutics. *Clin Cancer Res*, 14: 5000-5005.
214. Zhang LQ, Wang J, Jiang F, Xu L, Liu FY, Yin R. (2012) Prognostic value of survivin in patients with non-small cell lung carcinoma: a systematic review with meta-analysis. *PLoS One*, 7: e34100.
215. Stumpff J, von Dassow G, Wagenbach M, Asbury C, Wordeman L. (2008) The kinesin-8 motor Kif18A suppresses kinetochore movements to control mitotic chromosome alignment. *Dev Cell*, 14: 252-262.

216. Kasahara M, Nagahara M, Nakagawa T, Ishikawa T, Sato T, Uetake H, Sugihara K. (2016) Clinicopathological relevance of kinesin family member 18A expression in invasive breast cancer. *Oncol Lett*, 12: 1909-1914.
217. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhorji G, Benfeitas R, Arif M, Liu Z, Edfors F, Sanli K, von Feilitzen K, Oksvold P, Lundberg E, Hober S, Nilsson P, Mattsson J, Schwenk JM, Brunnstrom H, Glimelius B, Sjoblom T, Edqvist PH, Djureinovic D, Micke P, Lindskog C, Mardinoglu A, Ponten F. (2017) A pathology atlas of the human cancer transcriptome. *Science*, 357.
218. Birnbaum DJ, Finetti P, Lopresti A, Gilabert M, Poizat F, Raoul JL, Delpero JR, Moutardier V, Birnbaum D, Mamessier E, Bertucci F. (2017) A 25-gene classifier predicts overall survival in resectable pancreatic cancer. *BMC Med*, 15: 170.
219. Barrier A, Lemoine A, Boelle PY, Tse C, Brault D, Chiappini F, Breittschneider J, Lacaine F, Houry S, Huguier M, Van der Laan MJ, Speed T, Debuire B, Flahault A, Dudoit S. (2005) Colon cancer prognosis prediction by gene expression profiling. *Oncogene*, 24: 6155-6164.
220. Nault JC, De Reynies A, Villanueva A, Calderaro J, Rebouissou S, Couchy G, Decaens T, Franco D, Imbeaud S, Rousseau F, Azoulay D, Saric J, Blanc JF, Balabaud C, Bioulac-Sage P, Laurent A, Laurent-Puig P, Llovet JM, Zucman-Rossi J. (2013) A hepatocellular carcinoma 5-gene score associated with survival of patients after liver resection. *Gastroenterology*, 145: 176-187.
221. Lu Y, Lemon W, Liu PY, Yi Y, Morrison C, Yang P, Sun Z, Szoke J, Gerald WL, Watson M, Govindan R, You M. (2006) A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med*, 3: e467.
222. Navab R, Strumpf D, Bandarchi B, Zhu CQ, Pintilie M, Ramnarine VR, Ibrahimov E, Radulovich N, Leung L, Barczyk M, Panchal D, To C, Yun JJ, Der S, Shepherd FA, Jurisica I, Tsao MS. (2011) Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. *Proc Natl Acad Sci U S A*, 108: 7160-7165.
223. Penzvalto Z, Mihaly Z, Gyorffy B. (2009) [Gene expression based multigene prognostic and predictive tests in breast cancer]. *Magy Onkol*, 53: 351-359.
224. 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, Witteveen A, Glas

- A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R. (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*, 347: 1999-2009.
225. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*, 351: 2817-2826.
226. Nielsen T, Wallden B, Schaper C, Ferree S, Liu S, Gao D, Barry G, Dowidar N, Maysuria M, Storhoff J. (2014) Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer*, 14: 177.
227. Gyorffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B, Pusztai L. (2015) Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res*, 17: 11.
228. Creighton CJ, Hilger AM, Murthy S, Rae JM, Chinnaiyan AM, El-Ashry D. (2006) Activation of mitogen-activated protein kinase in estrogen receptor alpha-positive breast cancer cells in vitro induces an in vivo molecular phenotype of estrogen receptor alpha-negative human breast tumors. *Cancer Res*, 66: 3903-3911.
229. Creighton CJ. (2008) Multiple oncogenic pathway signatures show coordinate expression patterns in human prostate tumors. *PLoS One*, 3: e1816.
230. 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, Dressman HK, West M, Nevins JR. (2006) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature*, 439: 353-357.
231. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. (2006) A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet*, 38: 1043-1048.
232. Habermann JK, Doering J, Hautaniemi S, Roblick UJ, Bundgen NK, Nicorici D, Kronenwett U, Rathnagiriswaran S, Mettu RK, Ma Y, Kruger S, Bruch HP, Auer G, Guo NL, Ried T. (2009) The gene expression signature of genomic instability

in breast cancer is an independent predictor of clinical outcome. *Int J Cancer*, 124: 1552-1564.

## 11 SAJÁT PUBLIKÁCIÓK JEGYZÉKE

### 11.1 A disszertációhoz kapcsolódó publikációk jegyzéke

- **Nagy Á**, Pongor LS, Szabó A, Santarpia M, Győrffy B. (2017) KRAS driven expression signature has prognostic power superior to mutation status in nonsmall cell lung cancer. *Int J Cancer*, 140: 930-937. **IF=7,36**
- **Nagy Á**, Ósz Á, Budczies J, Krizsán S, Szombath G, Demeter J, Bödör C, Győrffy B. (2019) Elevated HOX gene expression in acute myeloid leukemia is associated with NPM1 mutations and poor survival. *J Adv Res*, 20: 105-116. **IF=6,992**
- **Nagy Á**, Lánckzy A, Menyhárt O, Győrffy B. (2018) Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep*, 8: 9227. **IF=4,011**
- Menyhárt O, **Nagy Á**, Győrffy B. (2018) Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. *R Soc Open Sci*, 5: 181006. **IF=2,515**
- **Nagy Á**, Munkácsy G, Győrffy B. (2020) Pancancer survival analysis of cancer hallmark genes. Publikálás alatt.

### 11.2 A disszertációtól független publikációk jegyzéke

- **Nagy Á**, Győrffy B. (2016) Internet-based opportunities in breast cancer diagnostics and research. *Magy Onkol*, 60: 273-280.
- Lánckzy A, **Nagy Á**, Bottai G, Munkácsy G, Szabó A, Santarpia L, Győrffy B. (2016) miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat*, 160: 439-446. **IF=3,626**
- Szász AM, Lánckzy A, **Nagy Á**, Förster S, Hark K, Green JE., Boussioutas A, Busuttill R, Szabó A, Győrffy B. (2016) Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget*, 7: 49322-49333. **IF=5,168**

## 12 KÖSZÖNETNYILVÁNÍTÁS

Elsőként témavezetőmnek, Prof. Dr. Gyórfy Baláznak szeretnék köszönetet mondani, aki lehetőséget biztosított munkám elvégzéséhez, a munkámat javaslataival és tanácsaival segítette.

Szeretnék köszönetet mondani az II. Sz. Gyermekgyógyászati Klinika igazgatójának, Prof. Dr. Kovács Gábornak, valamint a klinika korábbi igazgatójának, Prof. Dr. Szabó Andrásnak, hogy kutatásaimat a II. Sz. Gyermekgyógyászati Klinikán végezhettem.

Köszönettel tartozom a munkacsoport minden tagjának, akik munkámat segítették.

Külön köszönetet szeretnék mondani Dr. Menyhart Otíliának, a máj hepatocelluláris karcinómában végzett mRNS alapú prognosztikus biomarkerek igazolásáért.

Szeretném megköszönni Dr. Demeter Juditnak és Dr. Szombath Gergelynek, az akut mieloid leukémiás minták gyűjtésében nyújtott segítségüket.

Végül hálás köszönettel tartozom családomnak, akik támogattak munkám során.