

МЕДИЦИНСКИ УНИВЕРСИТЕТ-СОФИЯ

МЕДИЦИНСКИ ФАКУЛТЕТ

КАТЕДРА „МЕДИЦИНСКА ХИМИЯ И БИОХИМИЯ”

ЦЕНТЪР ПО МОЛЕКУЛНА МЕДИЦИНА

**МОЛЕКУЛНО ПРОФИЛИРАНЕ НА
ОСНОВНИТЕ ХИСТОЛОГИЧНИ ТИПОВЕ
БЕЛОДРОБНИ КАРЦИНОМИ**

ВЕРНИКА ЙОРДАНОВА ПЕТКОВА

**ДИСЕРТАЦИОНЕН ТРУД
ЗА ПРИСЪЖДАНЕ НА ОБРАЗОВАТЕЛНА И НАУЧНА СТЕПЕН
ДОКТОР**

Научна специалност: „Молекулярна генетика”

Професионално направление: 4.3. Биологически науки

Област на висше образование:

4. Природни науки, математика и информатика

НАУЧНИ РЪКОВОДИТЕЛИ:

ПРОФ. Д-Р РАДКА ПЕТРОВА КЪНЕВА, ДБ

АКАД. ПРОФ. Д-Р ВАНЬО ИВАНОВ МИТЕВ, ДМ, ДБН

СОФИЯ, 2023

Благодаря на ръководителите ми проф. д-р Радка Петрова Кънева, дб и акад. проф. д-р Ваньо Иванов Митев, дм, дбн за предоставената възможност за обучение и изработване на дисертационния труд в Център по Молекулна Медицина, МУ-София. Благодаря за техните съвети, подкрепа и насърчаване!

Благодаря на колабораторите от Клиниката по Гръдна Хирургия на Многопрофилна Болница за Активно Лечение по Белодробни Болести (МБАЛББ) „Св. София” ЕАД за съдействието, насочването на пациенти и събиране на включения биологичен материал!

Благодаря на всички колеги от Център по Молекулна Медицина и Катедра по Медицинска Химия и Биохимия за помощта, съветите и приятелско отношение, както и за хубавите моменти!

Благодаря на семейството и най-близките ми за силното търпение, подкрепа и обич!

Съдържание

Използвани съкращения	8
ВЪВЕДЕНИЕ.....	9
ЧАСТ ЛИТЕРАТУРЕН ОБЗОР	10
1.1. Епидемиология и рискови фактори при рак на белия дроб.....	11
1.2. Класификация на белодробните тумори.	11
1.3. Клиничното протичане на белодробните тумори.....	13
1.4. Туморна хетерогенност и молекулярно-генетични характеристики на РБД.....	14
1.5. Предиктивни молекулярно-генетични биомаркери при РБД.....	17
1.5.1. <i>EGFR</i> при РБД.....	17
1.5.2. Резистентност към анти - <i>EGFR</i> тирозин киназни инхибитори (TKIs).	18
1.5.3. Мутации в <i>KRAS</i> гена.....	19
1.5.4. Мутации в <i>BRAF</i>	20
1.5.5. Мутации в <i>MET</i>	20
1.5.6. Мутации в <i>LKB1/STK11</i>	21
1.5.7. <i>ALK</i> пренареждания.....	21
1.5.8. <i>ROS1</i> пренареждания.....	23
1.5.9. <i>RET</i> пренареждания.....	24
1.5.10. <i>NTRK1</i> пренареждания.....	26
1.6. Имунотерапия при РБД.....	26
1.6.1. PD-1/PD-L1 оста.....	26
1.7. Епигенетични промени при РБД.....	29
1.7.1. Хистонови модификации, асоциирани с РБД.....	29
1.7.2. МикроРНКи.....	31
1.8. Приложение на машинното самообучение в генетиката и геномиката.....	36
ЧАСТ 2 ЦЕЛ И ЗАДАЧИ	38
2.1. ЦЕЛ.	39
2.2. ЗАДАЧИ.	39
ЧАСТ 3 МАТЕРИАЛИ И МЕТОДИ.....	40
3.1. Клиничен материал.....	41
3.2. Биологичен материал.	41
3.2.1. Клинични показатели на изследваната група.....	41
3.3. ПРЕДНАЛИТИЧНА ОБРАБОТКА НА БИОЛОГИЧНИТЕ МАТЕРИАЛИ.....	43
3.3.1. Изолиране на тотална РНК от тъкан.....	43
3.3.2. Изолиране на тотална ДНК от тъкан.....	44
3.3.3. Оценка на качеството и концентрацията на изолирана тотална РНК/ДНК.....	45
3.3.3.1. Определяне концентрация на изолирана тотална РНК/ДНК чрез NanoDrop 2000.....	45
3.3.3.2. Определяне концентрация на изолирана тотална РНК/ДНК с флуориметричен метод с помощта на Qubit v2.0.....	45
3.3.3.3. Определяне на концентрацията и RINe на изолираната РНК от тъкан чрез Agilent 2200 TapeStationSystems.....	46
3.3.4. Обратна транскрипция на изолираната РНК за синтез на кДНК.....	47
3.4. АНАЛИТИЧНА ОБРАБОТКА.....	49
3.4.1. микроРНК експресионно профилиране чрез микрочипове на туморна и нормална белодробна тъкан.....	49
3.4.2. Детекция на изследваните некодиращи РНК молекули посредством PCR в реално време.....	52
3.4.3. Детекция на иРНКи посредством PCR в реално време (real time-qPCR).....	54
3.4.4. Секвениране от ново поколение, с използването на TSACP и секвенционна платформа MiSeq (Illumina).....	55
3.5. Биоинформатичен и статистически подходи, използвани за анализ на получените резултати.....	58
3.5.1. Биоинформатичен анализ на получените резултати от секвенирането от ново поколение с MiSeq платформата.....	58
3.5.2. Обработка на резултатите от микроРНК микрочиповия анализ.....	58
3.5.2.1. Нормализация на данните.....	58
3.5.2.2. Анализ на основните компоненти – PCA (Principle Component Analysis).....	59
3.5.2.3. Volcano plot.....	59

3.5.3. Количествен анализ на резултатите от генната експресия чрез real time-qPCR.....	59
3.5.4. Статистически методи.....	60
3.5.4.1.Статистическа проверка на хипотези.....	60
3.5.4.2. χ^2 -квадрат	61
3.5.4.3. Параметрични методи	62
3.5.4.3.1.Student's T тест анализ.....	62
3.5.4.3.2.Едномерен еднофакторен дисперсионен анализ (ANOVA).....	62
3.5.4.4. Непараметрични методи.....	62
3.5.4.4.1.Тест на Ман-Уитни (Mann-Whitney).	62
3.5.4.4.2. .Тест на Крускал-Уолис (Kruskal-Wallis) за съпоставка на три или повече независими извадки.....	63
3.5.4.5.Оценка на процента на фалшиво позитивни резултати (асоциации) при множествени сравнения: тест на Benjamini-Hochberg FDR	63
3.5.4.6.Корелационен анализ.....	63
3.5.4.6.1.Коефициент на линейна корелация на Пиърсън (Pearson; r)	64
3.5.4.6.2.Рангова корелация на Спийрман (Spearman's rho).	64
3.5.4.6.3.Посторояване на корелационно - базирани мрежи.....	65
3.5.5. Прогнозиране на взаимодействията между микроРНК и целеви гени.....	66
3.5.6. Анотация на генната онтология (GO) и анализ на сигналните пътища.	66
3.5.7. Диаграма на Вен.	66
3.6. Машинно самообучение.....	67
3.6.1. Подготовка на данните и почистване на данните.....	67
3.6.2. Нормализиране на данни.	68
3.6.3. Моделиране.	68
3.6.4. Избрани ML алгоритми за класификационен анализ.....	69
3.6.5. Анализ на методи за валидиране.....	69
3.6.6. Градиентно спускане.....	70
3.6.7. Метрики за оценка на класификационни модели.	70
3.6.7.1.Матрица за оценка на класификацията (Confusion Matrix).....	70
3.6.7.2.Основни метрики за ефективност.....	71
3.6.8. Йерархичен клъстерен анализ.	73
3.6.9. Метод на k-най-близките съседи.....	74
ЧАСТ 4 РЕЗУЛТАТИ.....	75
4.1. РЕЗУЛТАТИ ОТ микроРНК МИКРОЧИПОВИЯ АНАЛИЗ.....	76
4.1.1. Измерване на концентрацията и определянето на качеството на изолираните РНК с помощта на Nano Drop и Agilent 2200 TapeStation System. Изготвяне на 50 ng/ μ l разредки за miRNA експресионното профилиране чрез микрочипов анализ.....	77
4.1.2. Подбор на метод за нормализация на резултатите от микрочиповия анализ.	79
4.1.3. Контрол на качеството на резултатите и PCA (principle component analysis - анализ на основните компоненти).	80
4.1.4. Филтриране на получените резултати.	80
4.1.5. Установяване на микроРНКи с променена експресия при АК и ПК в сравнение с прилежащите им нормални тъкани.....	82
4.1.6. Резултати от йерархичен клъстерен анализ.	91
4.1.7. Избор на микроРНКи за валидация.	91
4.2. Валидация на избраните микроРНКи чрез RT-qPCR.....	95
4.3. Нива на експресия на 14-те изследвани микроРНКи в тъканни проби от рак на белия дроб.	95
4.4. Асоциация между експресионните нива на микроРНКите и клиничнопатологичните характеристики на пациентите.....	99
4.5. Корелационен анализ между изследваните микроРНКи.	109
4.6. Анализ на гените – мишени на микроРНКите.	112
4.7. Онтологичен анализ.	112
4.8. Най-силно засегнати и важни за раковите клетки пътища повлияни от показалите диференциална експресия микроРНК-и, регулиращи участници в тях.....	116
4.8.1. EGF/EGFR сигнален път.....	116
4.8.2. Сигнален път при НДРБД.....	116
4.8.3. VEGFA-VEGFR2 сигнален път.	117

4.8.4. Сигнални пътища свързани с ДНК поправката.	117
4.8.5. TGF-бета сигнален път.	118
4.8.6. Hippo-Merlin сигнален път.	119
4.8.7. BDNF сигнален път.	120
4.8.8. Wnt сигнален път.	120
4.8.9. Клетъчно стареене и автофагия.	121
4.8.10. Хомеостаза на медта.	121
4.8.11. Онкостатин М сигнален път.	122
4.9. Избор на мРНКи, регулирани от показалите диференциална експресия микроРНК-и в изследваните АК и ПК и имащи важно клинично значение при РБД.	123
4.10. РЕЗУЛТАТИ ОТ ЕКСПРЕСИОННИЯ АНАЛИЗ на <i>PD-L1</i> , <i>EGFR</i> , <i>STAT3</i> , <i>PTEN</i> , <i>MINA53</i> и <i>mTOR</i> при НДРБД.	125
4.10.1. Корелационен анализ между изследваните иРНКи и микроРНКи.	127
4.10.2. Асоциация между експресионните нива на иРНКите и клиничнопатологичните характеристики на пациентите.	129
4.11. Резултати от контролираното самообучение (Supervised Learning) за разграничаване на двата подтипа НДБРД.	137
4.11.1. Избор на алгоритъм за класификационен анализ.	137
4.11.1.1. Трениране и тестване на модели.	137
4.11.1.2. Logistic regression.	137
4.11.1.2.1. Конфигуриране на алгоритъма.	137
4.11.1.2.2. Резултати след извършено трениране на модела.	137
4.11.1.3. Decision tree.	138
4.11.1.3.1. Конфигуриране на алгоритъма.	138
4.11.1.3.2. Резултатите след трениране на модела.	139
4.11.1.4. Random forest (RF).	139
4.11.1.4.1. Конфигуриране на алгоритъма.	140
4.11.1.4.2. Резултатите след трениране на модела.	140
4.11.1.5. Метод на опорните вектори (Support Vector Machines; SVM).	140
4.11.1.5.1. Конфигуриране на алгоритъма.	141
4.11.1.5.2. Резултатите след трениране на модела.	141
4.11.2. Оценка на моделите.	141
4.12. Резултати от неконтролирано самообучение (Unsupervised Learning) за разграничаване на морфологичните подтипове при АК и ПК.	143
4.12.1. Йерархичен клъстерен анализ.	143
4.12.2. Метод на k-най-близките съседи.	146
4.13. Резултати от проведеното таргетно секвениране от ново поколение.	147
4.14. Връзка между нивата на експресия на микроРНКите и иРНКите и наличието на мутации.	150
ЧАСТ 5 ОБСЪЖДАНЕ НА ПОЛУЧЕНИТЕ РЕЗУЛТАТИ.	158
5.1. Обсъждане на резултатите от микроРНК експресионното профилиране чрез микрочипове в туморни и нормални белодробни тъкани.	159
5.1.1. Обсъждане на резултатите от анализа на основните компоненти (PCA) и клъстерния анализ.	162
5.1.2. Валидация на избраните микроРНКи чрез RT-qPCR.	164
5.1.3. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ ОТ RT-qPCR за микроРНКите и асоциация с криникопатологичните характеристики.	165
5.1.3.1. miR-4689.	165
5.1.3.2. miR-6785.	165
5.1.3.3. miR-101-3p.	165
5.1.3.4. miR-140-3p.	167
5.1.3.5. miR-144-3p.	168
5.1.3.6. miR-744-3p.	169
5.1.3.7. miR-95-3p.	169
5.1.3.8. miR-650.	170
5.1.3.9. miR-7-5p.	171
5.1.3.10. miR-21-3p.	172
5.1.3.11. miR-375.	173

5.1.3.12.miR-139-5p	173
5.1.3.13.miR-195-5p и miR-195-3p	174
5.2. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ от корелационно-базираните мрежи и йерархичния клъстерен анализ.....	176
5.3. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ проведения онтологичен анализ и анализ на сигналните пътища.....	177
5.3.1. Поддържане на клетъчната пролиферация.....	178
5.3.2. Потискане на супресори на растежа.....	180
5.3.3. Инвазия и метастазирание.....	180
5.3.4. Геномна нестабилност.....	184
5.3.5. Потискане на клетъчната смърт.....	185
5.3.6. Индуциране/достъп до кръвоносни съдове.....	186
5.3.7. Препрограмиране на клетъчния метаболизъм.....	187
5.3.8. Избягване на имунния отговор.....	190
5.3.9. Немутационни промени водещи до трансформация на клетките.....	191
5.4. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ ОТ RT-qPCR за мРНКите и асоциация с клинично-патологичните характеристики.....	193
5.4.1. Обсъждане на резултати получени при анализа на експресията на <i>PD-L1</i> при АК и ПК.....	194
5.4.2. Обсъждане на резултати получени при анализа на експресията на <i>STAT3</i> при АК и ПК.....	195
5.4.3. Обсъждане на резултати получени при анализа на експресията на <i>PTEN</i> при АК и ПК.....	196
5.4.4. Обсъждане на резултати получени при анализа на експресията на <i>mTOR</i> при АК и ПК.....	196
5.4.5. Обсъждане на резултати получени при анализа на експресията на <i>EGFR</i> при АК и ПК.....	197
5.4.6. Обсъждане на резултати получени при анализа на експресията на <i>MINA53</i> при АК и ПК.....	198
5.5. Обсъждане на резултатите от проведеното машинно самообучение.....	198
5.6. Обсъждане на резултатите от проведеното секвениране от следващо поколение.....	200
5.6.1. Мутации в <i>TP53</i> гена.....	202
5.6.2. Мутации в <i>KRAS</i> гена.....	205
5.6.3. Мутации в <i>STK11</i> гена.....	206
6. ИЗВОДИ:	208
7. ПРИНОСИ:	211
8. ПУБЛИКАЦИИ, УЧАСТИЯ В НАУЧНИ ПРОЯВИ И ПРОЕКТИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД.....	212
8.1. Публикации.....	212
8.2. Участия в научни прояви:.....	212
8.3. Участие в проекти:.....	214
9. БИБЛИОГРАФИЯ	215

Използвани съкращения

рак на белия дроб (РБД)
не-дребноклетъчен рак на белия дроб (НДРБД)
невроендокринен тумор на белия дроб (НЕТбд)
дребноклетъчен карцином (ДКБД)
едроклетъчния невроендокринен карцином (ЕКНЕК)
умерено диференцирания атипичен карциноид (АтК)
високо диференцирания типичен карциноид (ТК)
аденокарцином (АК)
плоскоклетъчен рак на белия дроб (ПК)
туморна микросреда (ТМС)
епително мезенхимен преход (ЕМП)
ДНК (Дезоксирибонуклеиновата киселина)
РНК (Рибонуклеинова киселина)
аргонавт (AGO)
микроРНК индуциран комплекс за заглушаване (miRISC)
нетранслиращи се области (UTR)
RINe (RNA integrity number)
тирозин киназни инхибитори (TKIs)
EGFR тирозин киназни инхибитори (EGFR-TKIs)
моноклонални антитела (mAbs)
Агенция за контрол на храните и лекарствата на САЩ (FDA)
Европейската агенция по лекарствата (EMA)
полимеразна верижна реакция в реално време (RT-qPCR)
микрочипов анализ (МА)
нкРНКи (non-codingRNA)
секвенирането от следващо поколение (NGS)
миссенс мутация (missense;MS)
мутация с изместване рамката на четене (frameshift; FS)
нонсенс мутация (nonsense; SG)

TruSeq Amplicon-Cancer Panel (TSACP)
биологични процеси (БП)
клетъчни компоненти (КК)
молекулни функции (МФ)
генната онтология (Gene Ontology - GO)
анализ на основните компоненти (principle component analysis; PCA)
стандартно отклонение (standard deviation;SD)
Benjamini - Hochberg FDR (BH-FDR)
корелационно - базирани мрежи (КБМ)
контролирано самообучение (Supervised learning)
неконтролирано самообучение (Unsupervised learning)
подсилено обучение (Reinforcement learning)
логистична регресия (Logistic Regression; LR)
метод на опорните вектори (Support Vector Machine; SVM)
дърво на решения (Decision Tree; DT)
случайна гора (Random Forest; RF)
L1 (Least Absolute Shrinkage and Selection Operator; LASSO)
L2 (Ridge Regularization)
истински положителни резултати (TP)
истински отрицателни (TN)
фалшиви положителни резултати (FP)
фалшиви негативи резултати (FN)
положителна прогностична стойност (PPV)
отрицателна прогностична стойност (NPV)
средна хармонична стойност (F1 Score)
ROC крива (ROC Curve)
истински положителен размер (True Positive Rate; TPR)
фалшив положителен размер (False Positive Rate; FPR)
Площта под кривата (area under the curve;AUC)
метод на k-най-близките (KNN)

ВЪВЕДЕНИЕ

Ракът на белия дроб (РБД) е една от водещите причини за смъртност вследствие на онкологично заболяване в световен мащаб. По честота той е на първо място при мъжете в Европа и на второ място в САЩ (1). В България РБД е най-честото злокачествено заболяване при мъжете и шесто по честота при жените (2). Едва 15% се диагностицират в ранен стадий с петгодишна преживяемост от 50% (2). При останалата част заболяването се открива в напреднал стадий, с наличие на лимфни и/или далечни метастази, а петгодишната преживяемост е около 17% (2).

Въпреки напредъка в науката и медицината, ракът на белия дроб е водеща причина за смъртност, вследствие на онкологично заболяване сред мъжете и жените в световен мащаб. Тютюнопушенето, начинът на живот, различни фактори на околната среда и фамилната история са основните рискови фактори, свързани със заболяването.

С развитието на новите технологии, като новогенерационното секвениране, в последните години се разкриват молекулярните механизми за възникването на различните хистологични групи белодробни тумори. Намерени са специфични соматични мутации, амплификации и транслокации, които вече навлязоха в клиничната практика като диагностични, прогностични и предиктивни биомаркери. Те позволяват по-ранното диагностициране, проследяване на пациентите и определяне отговора им към терапия. Успехът на таргетните терапии и новите имунотерапевтични подходи доведоха до смяна на парадигмата за диагноза и лечение.

Малките некодиращите РНК молекули в последните години се разглеждат с особен интерес, тъй като могат да послужат като диагностични и прогностични биомаркери. Те се характеризират със специфична експресия в туморните и нормалните тъкани и се свързват с карциногенезата. В настоящия дисертационен труд, чрез микрочиповия анализ, се проведе глобално изследване на експресията на 2548 известни микроРНК при два подтипа не-дребноклетъчен белодробен карцином. От тях бяха избрани 14 специфични микроРНК с променена експресия, които имат отношение към туморогенезата. От проведения онтологичен анализ на иРНК молекулите, мишени на ключови микроРНК при аденокарциномите и плоскоклетъчните белодробни карциноми и последващия анализ на сигналните пътища, се откриха различни засегнати ракови-специфични клетъчни процеси. Подбрани бяха ключови белтъци, участващи в тях, имащи отношение към провеждане на терапията и чиято експресия да се проследи. Чрез методите на машинното самообучение се изследва възможността чрез анализ на ко-експресията на 14 микроРНК да се разграничат, както двата подтипа не-дребноклетъчен белодробен карцином, така и различните хистологични типове аденокарциномите и плоскоклетъчните белодробни карциноми.

Този дисертационен труд разглежда различни генетични и епигенетични промени при рака на белия дроб и тяхното молекулярно и клинично значение, както и методите за тяхното изследване. Откриването и валидирането на нови биомаркери при това хетерогенно онкологично заболяване би било полезно в клиничната практика, а идентифицирането на ключовите засегнати молекулни пътища и специфични регулаторни механизми би дало насоки и за търсенето на нови по-ефикасни лекарствени средства.

ЧАСТ 1

ЛИТЕРАТУРЕН *ОБЗОР*

1.1. Епидемиология и рискови фактори при рак на белия дроб.

За развитие на рак на белия дроб важно значение имат влиянието на околната среда, начинът на живот и генетичната предразположеност. Тютюнопушенето е един от основните рискови фактори за РБД и 80-90% от диагностицираните пациенти са пушачи (2). Тютюневият дим е смес от химикали и около 60 канцерогена (3). Най-силно канцерогенни са полицикличните ароматни хидровъглероди, ароматните амини и азот-нитрозамините. При взаимодействие на инхалираните канцерогени с епителните клетки на дихателните пътища се индуцират повреди в ДНК (3). Ако не се отстранят от клетъчните репарационни системи тези повреди водят до увеличаване на мутационния товар. По този начин нараства рискът от възникване на соматични мутации в онкогени и тумор-супресорни гени (3). Най-често тези промени се получават при ковалентно свързване на метаболитно активния канцероген с ДНК и образуване на т.нар. ДНК-адукти (3). Всеки биологичен процес, който води до възникване на соматични мутации, се характеризира с определен мутационен профил. При пушачи тъканите, които директно са изложени на тютюнев дим, показват повишени нива на ДНК-адукти, от които с най-висока честота са еднонуклеотидната C>A и динуклеотидната замяна CC>AA (3). Същият мутационен профил се открива при 13,8% от непущачите, което се дължи на т.нар. пасивно пушене (3).

Останалите фактори на околната среда за възникване на РБД включват експозиция на азбест, хром, арсен, никел, кадмий, силиций, предхождащи белодробни заболявания и други. При липсата им, генетичната предразположеност също увеличава риска от възникване на РБД (2).

1.2. Класификация на белодробните тумори.

Според класификацията на СЗО, въз основа на хистология, биология и терапевтичен подход, РБД се разделя на 2 основни групи – недребноклетъчен рак на белия дроб (НДРБД), който се диагностицира в 85% от случаите и невроендокринен тумор на белия дроб (НЕТбд), в останалите 15% от случаите (2).

Групата на НЕТбд обединява туморите, произлизащи от невроендокринните прекурсори (4). Морфологията на тези тумори при наблюдение под микроскоп се отличава с малки кръгли или вретеновидни клетки (4). Групата на НЕТбд е съставена от твърде различни помежду си карциноми по отношения на молекулярно биологичното си поведение, клинични, епидемиологични, генетични характеристики, прогноза и начини на лечение (1). Тя включва ниско диференцирания дребноклетъчен карцином (ДКБД), едроклетъчния невроендокринен карцином (ЕКНЕК), умерено диференцирания атипичен карциноид (АтК) и високо диференцирания типичен карциноид (ТК) (4).

Хистопатологично, ЕКНЕК и ДКБД имат висока митотична активност, средно 70-80 митози на 2mm^2 и чести некрози, докато при карциноидите се наблюдават средно 2-10 митози на 2mm^2 , ниска степен на клетъчно делене и Ki67-пролиферативен индекс около 10% (5,6).

Много от различията между ТК и карциномите се базират на тютюнопушенето, което се асоциира с ДКБД (1). Почти всички пациенти с ДКБД са настоящи или бивши тежки пушачи. Преживяемостта на болните с тази диагноза е ниска – 12% от тях преживяват 2 години, 7% преживяват 3 години и само 5% - 5 години, като

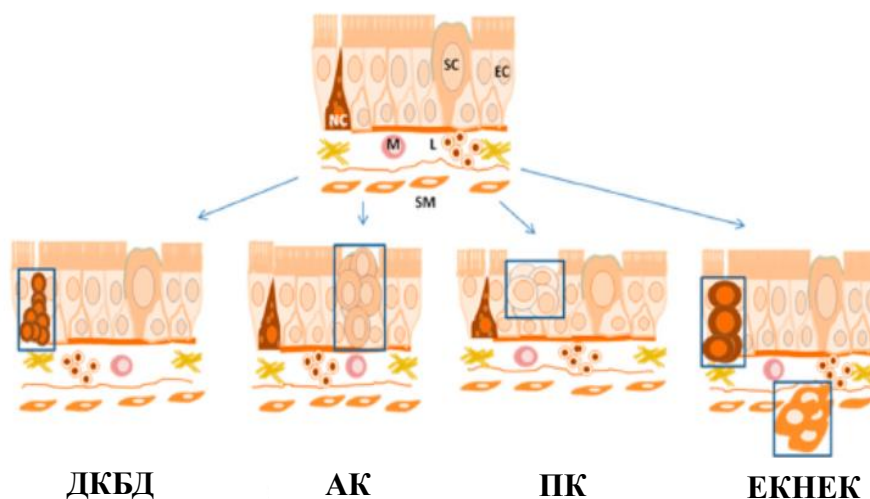
приблизително две трети от пациентите при диагностициране са с далечни метастази (7). Лечението на ДКБД най-често включва комбинирана химиотерапия и лъчетерапия (8). Туморите проявяват висока първична чувствителност към прилаганото лечение, но бързо след това се откриват рецидиви, нечувствителни към нито един вид терапия (9).

Групата на НДРБД включва 5 основни типа - аденокарцином (АК), плоскоклетъчен рак на белия дроб (ПК), едроклетъчен карцином, аденосквамозен и саркоматоиден рак (1,4). Всеки един от тях морфологично се разделя на подтипове. В България най-често диагностицираните подтипове на НДРБД са ПК – 50,2% от всички морфологично доказани случаи и АК – 21,9% от случаите (10).

ПК на белия дроб се развива от бронхиалните епителни клетки и свързва по-силно с тютюнопушене от другите типове НДРБД (2) Фигура 1.1. В повечето случаи той се локализира централно, близо до левия или десния бронх, като напоследък има данни и за неговата периферна локализация. Клетките са с полигонална или окръглена форма, на ограничени гнезда и с видими междуклетъчни мостчета. Според класификацията на Международна асоциация за проучване на белодробен карцином (IASLC), Американско торакално дружество (ATS) и Европейско респираторно дружество (ERS) (IASLC/ATS/ERS) ПК хистологично се разделя на три подтипа: ПК с кератинизация, ПК без кератинизация и базалоиден карцином (11). Наличието на кератинообразуване с характерните ракови перли, ядрения и клетъчен полиморфизъм определят три степени на диференциация – високо-, умерено- и нискодиференциран тумор (2).

АК е тумор с жлезиста диференциация или произвеждащ муцин. Води началото си от алвеоларните и бронхиалните епителни клетки Фигура 1.1. В по-голяма част от случаите се разполага периферно и туморите са малки и субплеврални, но също така могат да бъдат големи и централно разположени. Характерен е за непущачи и слабо изразени пушачи (1,4). Съгласно препоръките на IASLC/ATS/ERS АК се разделя на 5 хистологични подтипа, според вида на преобладаващата компонента: лепидични, ацинарни, папиларни, микропапиларни и солидни (11). На базата на валидирани клинични проучвания е установено, че при преобладаваща лепидична компонента е налице добра прогноза, докато преобладаващите микропапиларни и солидни архитектури се асоциират с неблагоприятна прогноза, особено при пациенти, подложени на ограничена резекция (11–13). Преобладаващите ацинарни и папиларни структури се характеризират с междинна прогностична стойност.

Едроклетъчният карцином е недиференциран НДРБД, при който липсват клетъчните характеристики на ДКРБ и се наблюдава жлезиста и плоскоклетъчна диференциация Фигура 1.1 (4). Той се локализира по-често в периферията, но се откриват и централно разположени лезии. Едроклетъчният карцином расте бързо, метастазира рано и се асоциира с тютюнопушенето (4).

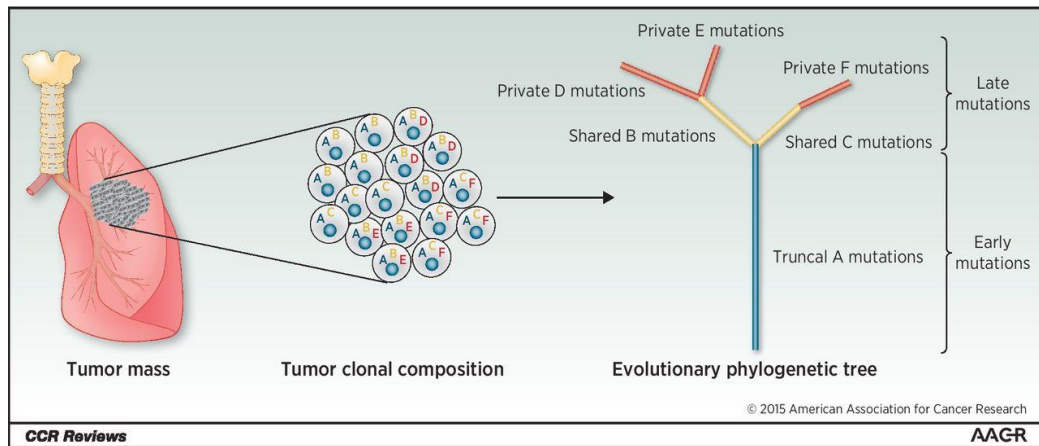


Фигура. 1.1 Различните хистологични типове рак на белия дроб. М – макрофаг; L – лимфоцит; NC – невроендокринна клетка; EC – епителна клетка; SC – секреторна клетка (4).

1.3. Клиничното протичане на белодробните тумори.

Клиничното протичане на белодробните тумори се обуславя основно от техните хистоморфологични и биологични особености. Прогнозата на пациентите се свързва с определени клинично-патологични показатели - хистологичен вид на тумора, степен на диференциация, големина на тумора (Т-статус), метастази в регионални лимфни възли (N-статус) и далечни метастази (M-статус) (14). pTNM-стадият е най-важният прогностичен фактор и служи като критерий за оценка на молекулярно биологичните характеристики на туморите (15). Посочените показатели обаче невинаги са достатъчни за прецизиране на диагнозата, определяне на биологичното поведение и отговора към терапия на РБД. Това налага необходимостта от по-задълбочени проучвания на имунохистохимични и молекулярно-генетични биомаркери с диагностична, прогностична и предиктивна стойност. Поради тази причина в последните години се наблюдава задълбочено изучаване на ролята на тумор-супресорни и онкогени в инициацията и прогресията на РБД (16).

Напредък има в лечението на НДКБК, докато при агресивния тип ДРБД нито едно от лекарствата не показва ефективност (17). На базата на наличието на „водещи“ /активиращи/ мутации са разработени или са в процес на разработка малки молекули инхибитори. Те са подходящи за лечение на определени групи от пациенти, при които се откриват специфични генетични/геномни характеристики (17,18). Към основните генетични и геномни промени, считани за „движещи мутации“ при НДКБК се отнасят мутациите в гените *EGFR* и *KRAS*, както и инверсията на хромозома 2p, водеща до фузия на гените *EML4* и *ALK* (17). Докато мутациите в *EGFR* и *ALK* се случват рано в туморната еволюция на НДКБК и присъстват във всички клетки на тумора, в отделните му субклонове могат да се открият специфични мутации, т.нар „разклоняващи“ мутации, които подпомагат туморната прогресия (19). Важни пътища, които често могат да бъдат засегнати, са RAS/RAF/MEK, фосфоинозитол-3-киназа PI3K/AKT/mTOR пътя, MET киназата, LKB1, IGF-1R и др.



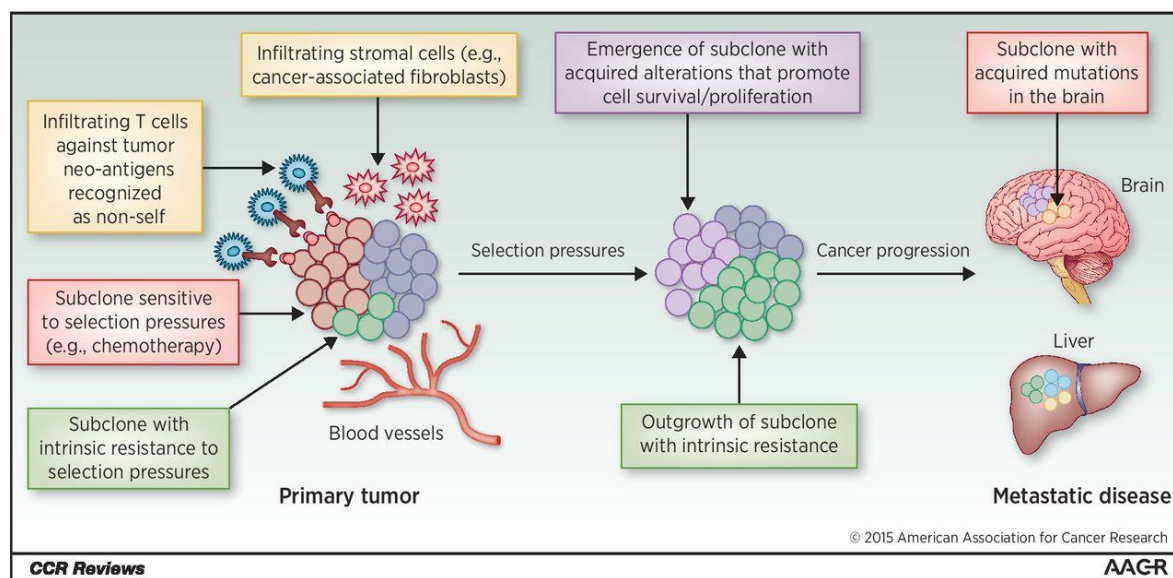
Фигура. 1. 2. Първичните тумори се състоят от различни субклонове с общи и специфични соматични промени. Промените, споделяни от всички туморни клетки, възникват в началото на туморогенезата, представен от синия ствол на филогенетичното дърво; В или С са промените, специфични за туморните клетки и присъстват в някои клетки на тумора, но не всички. Те се появяват по-късно в туморогенезата, представени като жълтите клонове на филогенетичното дървото; изолираните промени, присъстващи само в едни клетки на тумора, също се появяват по-късно в туморогенезата и са представени като червените клонове на филогенетичното дървото (20).

Новите и по-ефективни терапевтични стратегии и имунотерапията се очертават като важно научно постижение при лечението на този РБД. Въпреки това остава предизвикателството ранното им диагностициране и преодоляване на първичната/вторична лекарствена резистентност, водеща до прогресия на заболяването. Освен това при голяма част от пациентите с РБД се прилагат конвенционалните терапевтични протоколи, които имат сравнително умерен ефект (16).

1.4. Туморна хетерогенност и молекулярно-генетични характеристики на РБД.

Двата основни хистологични типа белодробни карциноми могат да бъдат разграничени на молекулярно ниво по специфични „водещи“ мутации („driver mutations“) в множество онкогени (17). Това включва еднонуклеотидни замени, амплификация на гени, хромозомни аберации и пренареждания, загуба на хетерозиготност (LOH) и различни епигенетични промени. Последните повлияват регулацията на генната експресия, чрез ДНК метилиране, деацетиране на хистони, ремоделиране на хроматина, импринтинг и специфична експресия на микроРНКи. Всички настъпили генетични изменения водят до промени в сигналните пътища и засягат процеси като клетъчния растеж и делене, диференциация, програмирана клетъчна смърт, инвазия и метастазиране (21). Така „водещите“ мутации водят до конститутивно активиране на сигнални белтъци и отключват и поддържат процеса на туморогенеза. Измененията в генома на раковата клетка задвижват туморната еволюция, подпомагат ангиогенезата и спомагат за препрограмирането на инфилтрираните от тумора стромални клетки, оформяйки благоприятна за развитие на тумора среда (22).

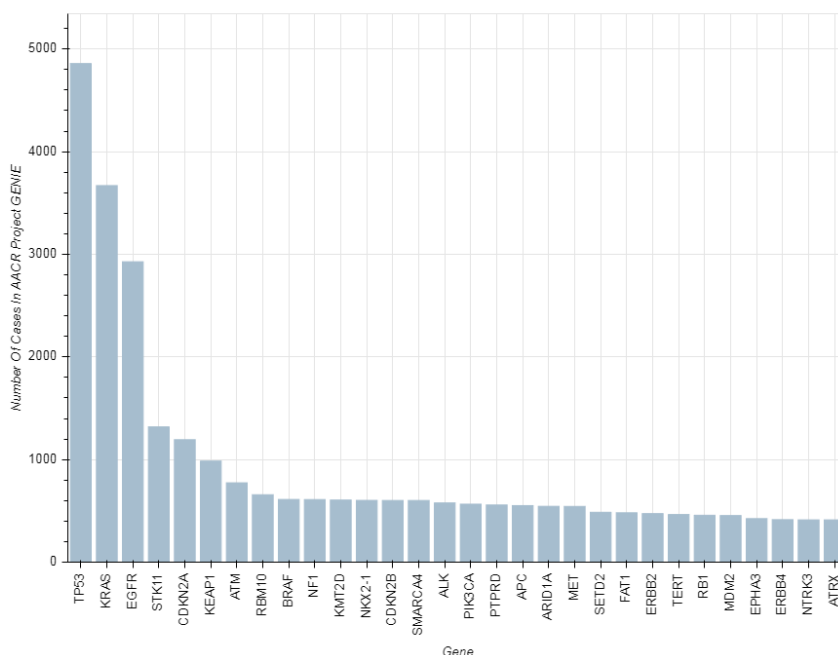
Тъй като туморите се хетерогенни, разлика в мутационния спектър може да се наблюдава както между различни популации от клетки на самия тумор, така и между първичния тумор и неговите метастази (Фигура. 1.3) (23). Вътретуморната хетерогенност се изразява в наличието на различни по генотип и фенотип клетъчни популации, които могат да се разграничават морфологично, да имат различни засегнати сигнални пътища, да притежават различен метастатичен потенциал и различна чувствителност към терапия. Предложени са две хипотези за възникване на вътретуморната хетерогенност – клонална еволюция и наличие на ракови стволови клетки (23).



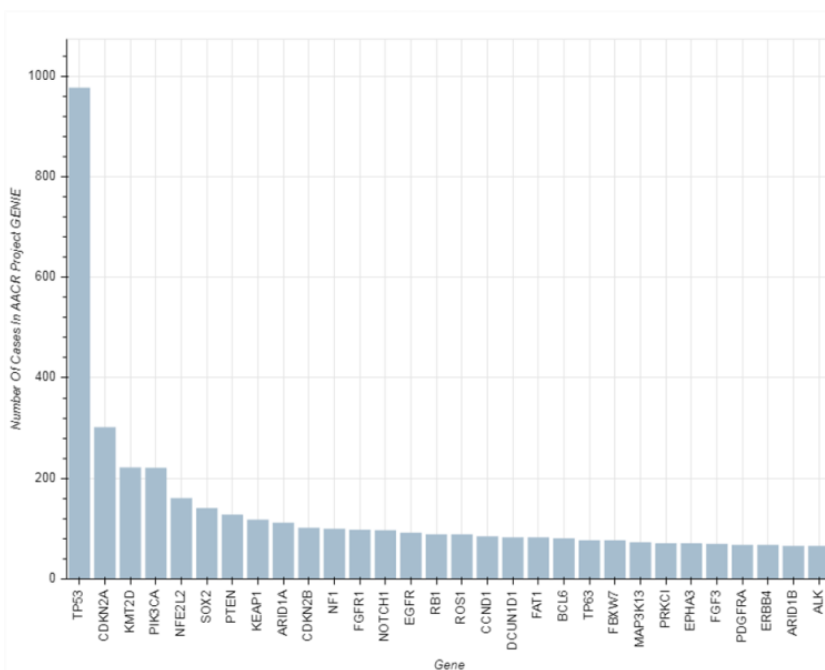
Фигура. 1.3. Вътретуморната хетерогенност и клонална еволюция. Първичните тумори, състоящи се от различни клетъчни субклонове, могат да бъдат подложени на различен селекционен натиск, например химиотерапия и различни фактори на микросреда (хипоксия или действие на тумор-инфилтрираните лимфоцити). Под влиянието на такъв селекционен натиск клетъчните субклонове с присъща резистентност (зелени клетки) могат да се делат усилено, което потенциално води до прогресия на заболяването, и/или субклоните могат да придобият соматични промени (пурпурени клетки), подпомагайки оцеляването на клетките, пролиферацията и метастатичния потенциал. Деленето на някои клетъчни субклонове (червени клетки) може да бъде ограничено от селекционния натиск и те са чувствителни към определена химиотерапия. Тумор-инфилтрираните лимфоцити могат да разпознаят неоантигените, представени на повърхността на туморните клетки като нетипични и се активират (20).

НДРБД се характеризира с висок соматичен мутационен товар, дефиниран като брой несинонимни замени в кодиращата секвенция на гените за мегабаза от генома (24). Особено висок мутационен товар (около 10-пъти повече) се наблюдава при пушачите, представляващи болшинството от пациентите. При АК се откриват средно 8.1 мутации на мегабаза, а при ПК 9.9 мутации на мегабаза (24). Най-често се откриват мутации в онкогени като TP53, Rb1, BRAF, KRAS, MET, PIK3CA, EGFR, ErbB2/HER2-neu, MDM2, MYC, NKX2-1, PDGFRA, c-KIT, CD44, Bcl-2, CCND1, CRK, p53, PTEN, Rb, LKB1, CDKN2A (p16/p14ARF), FHIT, CAV1, APC, TUSC2 (FUS1), CDH 1, CDH 13, DAPK1, GSTP1, MGMT, RARβ, RASSF1A, SEMA3B, TIMP3, EML4-ALK фузии (25). Честотата им е различна при АК, ПК и ЕКБК, както и биологията на туморите е

индивидуална при настоящи, бивши пушачи и непушачи (26). Някои мутации в онкогени са взаимно изключващи се и не могат да бъдат открити в един и същ тумор (27).



Фигура 1.4. Най-често срещани генетични изменения при аденокарцинома на белия дроб (28).



Фигура 1.5 Най-често срещани генетични изменения при плоскоклетъчен карцином на белия дроб (28).

Мутации в тумор-супресорите *TP53* и *RBI* се наблюдават често при белодробните карциноми, но все още няма разработени терапевтични подходи на засегнатите от тях сигнални пътища (29). При НДРБД рядко се откриват мутации в *RBI* и много често в *TP53* гена (30,31). *TP53* (17p13.1) е тумор супресорен ген, който има роля в поддържането на геномната стабилност, клетъчния цикъл, ДНК репарацията, апоптозата, ангиогенезата и стареенето. Над 70% от ПК и 30-40% от АК имат p53

мутации, което води до загуба на функция (32). Централно разположените клетки на белодробния епител, от които произлиза ПК са изложени на по-голяма концентрация на канцерогените, влизащи в състава на тютюневия дим. Поради тази причина се наблюдават сравнително по-често мутации при ПК в сравнение с клетките с периферно локализиран аденокарцином (33). Мутациите в TP53 се появяват рано в процеса на канцерогенезата и са свързани с лоша обща преживяемост, резистентност към терапия и чести случаи на рецидиви. При нормални условия концентрацията на p53 в клетката се поддържа ниска. При хипоксичен стрес или ДНК повреда, ядрената концентрация на p53 се увеличава. В отговор на увреждането на ДНК, p53 се регулира по MDM2-зависим начин и се активира от CDK1 и CDK2, което води до спиране на клетъчния цикъл и апоптоза. При настъпване на мутация в гена и загуба на функция, белтъка p53 не може да се свърже с промоторите на таргетните си гени и по този начин малигнените клетки не се насочват към апоптоза и се стимулира тяхната пролиферация (32). Много често при туморите, експресиращи мутантен p53 белтък се наблюдава геномна нестабилност (28).

При пациенти с мутация в TP53 гена, в комбинация с мутация в EGFR, ALK или ROS1 гените се наблюдава намаляване на общата преживяемост, в сравнение с групата, при които TP53 е див тип и се открива мутация в EGFR, ALK или ROS1 гените (34). Същото се наблюдава и при пациентите, при които се детектират едновременно мутации в KRAS или PIK3CA гените. KRAS, PIK3CA и TP53 мутациите се асоциират с по-висок метастатичен потенциал на туморите и лоша прогноза (35). Поради тази причина е необходимо провеждане на рутинно масивно паралелно секвениране не само на гените, имащи отношение към прилагането на таргетни терапии, но и на TP53 гена, имащ отношение към основните молекулни пътища, свързани с онкогенезата при рака на белия дроб. Необходими са още задълбочени изследвания, за да се определи механизмът на действие на TP53 мутационния ефект върху туморната микросреда. Същото така и за да се потвърди или отхвърли хипотезата, че мутациите в TP53 гена са подходящ маркер при рака на белия дроб, имащи отношение и към прогнозата и избора на терапия.

1.5. Предиктивни молекулярно-генетични биомаркери при РБД.

1.5.1. EGFR при РБД.

EGFR генът е локализиран на късото рамо на 7 хромозома и кодира трансмембранен рецептор с тирозин-киназна активност (36). При свързването на EGFR рецептора с лиганд се активира цитоплазмения му тирозин-киназен домен чрез димеризиране и фосфорилиране. По този начин се активират низходящите сигнално-трансдукционни пътища, водещи до неконтролируемо клетъчно делене, пролиферация, метастазиране и възпрепятстване на апоптозата (37). При 62% от НДРБД EGFR показва повишена експресия, която се асоциира с по-лоша прогноза (38).

Приблизително при 10% от НДРБД в Европа и 50% в Източна Азия се откриват мутации в EGFR. EGFR мутациите често се откриват при жени непущачи, диагностицирани с АК (38). При ПК мутациите в EGFR се срещат при около 8% от непущачите и около 2.1% от непущачите (39). Пациентите с EGFR мутирали тумори са подходящи за таргетна терапия с анти-EGFR инхибитори. Мутациите се откриват в екзони 18, 19, 20, 21 на EGFR гена, кодиращи тирозин-киназния домен и се

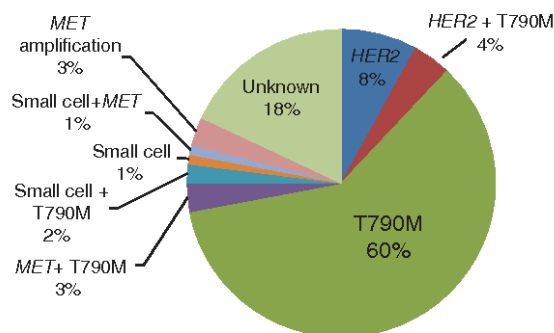
характеризират с различна честота и различна чувствителност към анти-EGFR инхибитори (36). При над 90% от АК преобладават делеции в екзон 19, водещи до изместване в рамката на четене или “hotspot” точковата мутация L858R в екзон 21. При ПК тези две аберации се откриват сравнително рядко. Активиращите мутации в тирозин-киназия домен на EGFR водят до лиганд-независимо активиране на стартираните от него сигнални пътища (36).

Определянето на мутационния статус на *EGFR* е важна съпътстваща диагностика за избора на таргетната група пациенти, подходящи за лечение с анти-EGFR тирозин киназни инхибитори (TKIs). Такива инхибитори са gefitinib (Iressa1, AstraZeneca, Wilmington, DE, USA), erlotinib (Tarceva1, F. Hoffmann-La Roche, Basel, Switzerland), afatinib – (Gilotrif, Boehringer Ingelhem, International GmbH).

За детекция на *EGFR* мутации в ДНК, изолирана от парафинови блокчета, се използват молекулярно-генетични методи, като директно секвениране, пиросеквениране, секвениране от следващо поколение (NGS) или други методи, базирани на количествен PCR (39). Хетерогеността на мутациите в *EGFR* може да повлияе резултата от генетичния тест. Поради тази причина е необходимо правилно да се избере метода на изследване и пробите да се подготвят по стандартни процедури, позволяващи адекватно извършване на последващите анализи и надеждност на получените резултати (36).

1.5.2. Резистентност към анти-EGFR тирозин киназни инхибитори (TKIs).

Въпреки първоначалния добър отговор при голяма част от пациентите с НДРБД, лекувани с анти-EGFR TKIs, на по-късен етап се развива резистентност (39). Най-честа причина за тази резистентност, в около 60% от случаите, е наличието на мутацията T790M в екзон 20 на *EGFR* гена (39). Замяната води до конформационни промени и намаляване на свързването на TKIs с рецептора. Също така се увеличават афинитета на киназия домен към АТФ и нивата на фосфорилиране, с което се намалява потенциала на TKIs (39). Поради тази причина скринингът за мутацията T790M е необходим за прогноза и предикция на отговора на пациента към терапия с анти-EGFR TKIs. Други механизми за възникване на резистентност са амплификация на *MET* (5-11% от случаите), загуба на *RB1* и *TP53*, хистологична трансформация на НДРБД към ДКБК (3-20% от случаите), амплификация на *ERBB2* (12-13% от случаите), мутации *PI3KCA* (5% от случаите), мутации в *KRAS*, *BRAF*, загуба на *PTEN*, мутации в *HER2* и *VINI* [30,31].



Фигура 1.6. Честота на различните механизми за придобиване на резистентност към тирозин-киназните инхибитори (40).

С цел преодоляване на резистентността е създадено трето поколение TKIs, специфично насочени срещу *EGFR* мутацията T790M. Osimertinib (Tagrisso, Astra Zeneca) е орален селективен и необратим анти-EGFR TKIs (17). Той се свързва ковалентно с цистеина на 797 позиция и се прилага при пациенти, развили *EGFR* T790M-резистентна мутация. Появата на мутацията C797S в 20 екзон на гена, води до невъзможност за свързване на TKIs с рецептора и също е причина за развитие на резистентност (41). Известно е, че когато C797S и T790M са в „дис-“ позиция (на един алел) е невъзможно да се преодолее резистентността. Когато двете мутации са в „транс“ позиция (на различни алели) се наблюдава положителен отговор на пациентите към комбинирана терапия с първо и трето поколение TKIs (42).

Четвърто поколение EGFR TKIs в момента се проучва за преодоляване на възникналата резистентност към osimertinib. EAI045 е първият алостеричен инхибитор, насочен срещу *EGFR* мутациите T790M и C797S (43). BLU-945 е друго четвърто поколение EGFR TKI, което преодолява резистентността, придобита от появата на трите резистентни мутации - делеция в екзон 19 или L858R и T790M и C797S (44).

Тъй като екстрацелуларният домен на EGFR е обект на прицелна терапия с моноклонални антитела (cetuximab), а интрацелуларния с тирозин киназни инхибитори (gefitinib и erlotinib) с цел да се блокира сигнално трансдукционната каскада, водеща до неконтролируема клетъчна пролиферация, количествената експресия на EGFR в белодробни тумори би могла да бъде полезна за оценка на евентуалния терапевтичен отговор при прицелна терапия (45–47).

1.5.3. Мутации в *KRAS* гена.

Вирусният онкогенен хомолог на саркома на *Kirsten* при плъх (*Kirsten rat sarcoma viral oncogene homolog*, *KRAS*) принадлежи към семейството на RAS-онкогените, заедно с *HRAS* и *NRAS* (48). Генът *KRAS* е локализиран на хромозома 12.1 и кодира GTP-свързващ белтък. *KRAS* действа низходящо от няколко тирозин киназни рецептора, вкл. *EGFR*, и е свързан с активирането на RAS/RAF/MAP-киназа (MEK/ERK) и RAS/MAPK-сигнални пътища (36). Мутации в *KRAS* се наблюдават при 25% - 35% от пациентите с НДКБК, най-често при АК (48). Известно е, че *KRAS* мутациите са по-силно асоциирани с тютюнопушене, отколкото повечето други активиращи онкогенни мутации, описани при НДКБК. При около 95% от случаите те са “миссенс” варианти в кодони 12 и 13 (36). При непущачи най-често срещаните замени са G12D и G12V, докато при пушачи най-често се открива мутацията G12C (48,49). Мутациите в *KRAS* при НДКБК, променяйки туморната микросреда се асоциират с лоша прогноза и намален отговор към химиотерапия и имунотерапия (50). Също така те са предиктивен биомаркер за резистентност към таргетна терапия с анти-EGFR тирозинкиназни инхибитори като gefitinib или erlotinib (36). При малък процент НДКБК (1-3%) се наблюдават едновременно *EGFR* и *KRAS* мутации, или *EML4-ALK* фузии и *KRAS* мутации (36).

Анализът на *KRAS* мутации може да се осъществи чрез алел-специфичен Real-Time PCR, Sanger-секвениране, пиросеквениране и NGS. Количественият RT-PCR е по-чувствителен метод и позволява откриване на мутантни алели с честота до 1% в изследваните тумори (51). Разработени са мултиплексни алел-специфични PCR-тестове за едновременно откриване и много точно определяне на мутации в *KRAS* гена,

одобрени от FDA (51). Таргетното NGS секвениране на панели от тумор-асоциирани гени и цялостното геномно/ екзомно секвениране (WholeGenome/ExomeSequencing – WGS/WES) все повече се използват в клиничната практика и позволяват количествен анализ на редки алели и едновременно откриване на целия спектър от соматични мутации и геномни пренареждания при белодробните тумори (39).

1.5.4. Мутации в *BRAF*.

Генът *BRAF* е локализиран на хромозома 7.34 и кодира серин/треонин киназа, разположена надолу по веригата на RAS/RAF/MEK/ERK сигналния път, който се активира от EGFR (36). *BRAF* е член на семейството RAF серин/треонин кинази и медира клетъчните отговори към сигналите за растеж, пролиферация и оцеляване на клетките. При НДРБД мутациите в *BRAF* се откриват в 1% до 4% от случаите, като най-често (около 50%) се наблюдава “hotspot” мутацията V600E (52). Други мутации, които се откриват са G469A/VK601E, L597R, D594G и G466V (53). Всички те са взаимноизключващи се с ”водещите мутации” в *EGFR*, *KRAS* и *ALK* (36).

BRAF мутации се откриват основно (85%) при пациентите с АК, които са настоящи или бивши пушачи (36). Наличието на *BRAF* “hotspot” мутацията V600E се асоциира с лоша прогноза и е негативен маркер за отговор към платина-базирана химиотерапия (54). Около 33% - 42% от пациентите с V600E позитивни тумори, показват положителен отговор към таргетна терапия с *BRAF*-инхибиторите vemurafenib и dabrafenib (54). В момента при РБД са в ход клинични проучвания на *BRAF* и MEK инхибитори, насочени към тумори с мутация в *BRAF* гена.

1.5.5. Мутации в *MET*.

Белтъкът c-Met е тирозин-киназен рецептор за хепатоцитния растежен фактор. Мутации в *MET* гена се откриват при около 5.18% пациентите с НДРБД (55). В повечето случаи това са възрастни пациенти с АК или ПК, характеризиращи се със специфична морфология на клетките. Амплификация или увеличена експресия на един от *MYC* протоонкогените — *MYC*, *MYCN*, или *MYCL* може да се открие при 20%–30% от ДКРБД (56).

Установено е, че мутациите, засягащи интронните граници около 14 екзон на *MET* възникват по-рано в процеса на туморогенеза, докато амплификация на гена се открива на по-късен етап и е свързана с по-агресивен ход заболяването (56). Най-честата аберация е засягане на донор-ацепторното място на 13 интрон на гена, при което сплайсингът на мРНК-а не протича правилно (56). Други мутации, които се откриват, са различни еднонуклеотидни замени в 14 екзон. Всички те водят до нарушения в структурата на белтъка, свързани с неговата резистентност към убиквитиниране и разграждане (56). По този начин MET се натрупва в клетката удължавайки действието на MAP-киназията.

Разработени са няколко MET-инхибитори, в това число малки TKIs и моноклонални антитела (mAbs) срещу MET или неговите лиганди (50,57). При тяхното прилагане се наблюдава добър първоначален отговор на пациентите и удължаване на времето до прогресия с 7.4 месеца, след което се появяват вторични мутации, водещи до резистентност (58).

1.5.6. Мутации в *LKB1/STK11*.

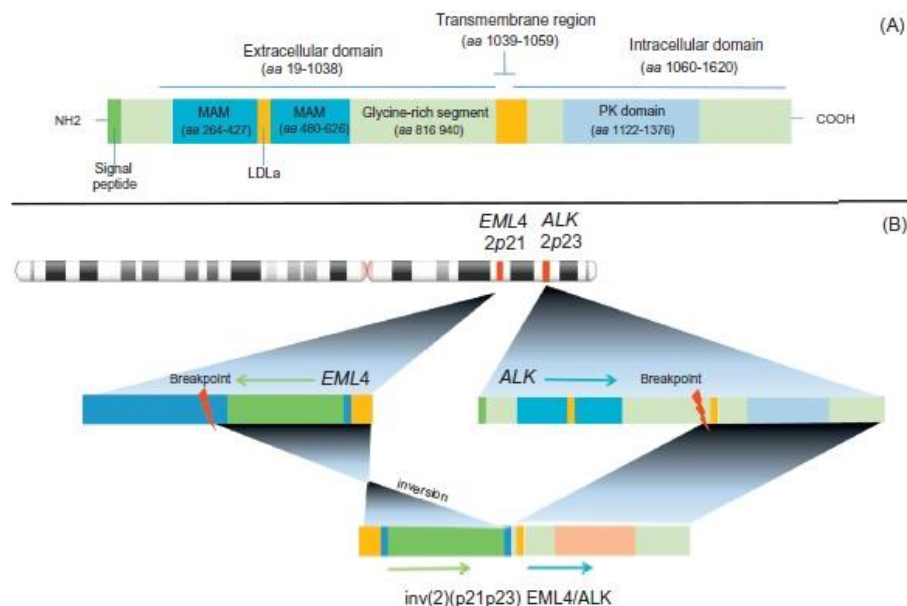
Серин/треонин киназа 11 е тумор-супресор, в който много често се откриват мутации при НДРБД (59). Физиологичната му роля е свързана с модулация на mTOR сигналния път чрез директна регулация на AMPK (59). Загубата на активност на *STK11* води до промени в диференциацията на раковата клетка, ремоделиране на екстрацелуларния матрикс, локална прогресия и метастазиране (59). В скорошно проучване Berger и колеги показват, че *STK11* „миссенс“ мутации се срещат най-често при АК (92%), като в голяма част от случаите се откриват съвместно с *KRAS* мутации. Намалената експресия на *STK11*, вследствие на епигенетично заглушаване или загуба на хетерозиготност, е негативен прогностичен фактор (59).

1.5.7. *ALK* пренареждания.

Анапластичната лимфом киназа (*ALK*, OMIM *105590) е рецептор с тирозин киназна активност, открит за пръв път през 1998 г. като резултат от транслокация при анапластичен едроклетъчен лимфом (Anaplastic Large Cell Lymphoma - ALCL) (60). Генът е локализиран на хромозома 2p23.2-p23.1 и има важна роля в развитието на ЦНС по време на ембрионалното развитие. При възрастни той се експресира слабо в ЦНС, тънкото и дебелото черво, простата, но не и в останалите тъкани, включително и в бял дроб. *ALK* рецепторът е част от голямото семейство на инсулиновите рецептори (61).

През 2007 г. за пръв път е установена специфична фузия на *ALK* с *EML4* при пациент с НДКБК (62). Двата гена са локализирани сравнително близо (на разстояние ~12 Мбр един от друг) и в резултат на хромозомна инверсия, обхващаща двата локуса, се получава фузионен онкоген *EML4-ALK*. Известни са 13 варианта и около 30 вида, при които се сливат различни екзони от 5'-края на *EML4* гена (2, 3, 6, 10, 13, 14, 15, 17, 18, 20, 21), с екзон 20 на *ALK* (61). При почти всички варианти фузионният ген съдържа киназния домен на *ALK* в 3'-края. Най-често се откриват транслокацията, включваща 13 екзон на *EML4* (E13;A20 (вариант 1)), следвана от варианта, съдържащ екзон 6a/b на *EML4* (61). Химерният белтък, продукт на тази транслокация, активира киназите STAT3 и ERK и води до растеж, пролиферация и миграция на раковите клетки (36).

EML4-ALK фузията се открива при 5-6% от НДКБК. Среща се при по-млади пациенти, като преобладават жените, непущачи/ леки пушачи, диагностицирани с АК (63)[37]. Освен *EML4* са открити и други гени, които образуват фузионни белтъци с *ALK*, като *KIF5B*, *KLC1*, *TFG*, *TPR*, *HIP1*, *STRN*, *DCTN1*, *SQSTM1*, *NPM1*, *BCL11A* и *BIRC6*. *PRKAR1A*, *SPDYA* и *NBAS* (61,64).



Фигура 1.7. Схематично представяне на ALK гена и неговия фузионен партньор EML4. (A) Белтъчните домени са оцветени, както следва: N-краен домен: екстрацелуларен регион, включващ два MAM домена (аминокиселини 264-427 и 480-626, синьо), LDLa, (аминокиселини 453-471, червено, глицин – богат домен (аминокиселини 816-940, зелено), трансмембранен домен (ТМ, аминокиселини 1039-1059, оранжево), цитоплазмен домен, съдържащ тирозин-киназия домен (аминокиселини 1122-1376; сиво). Схематичният EML4 – ALK ген е в резултат на инверсия на късото рамо на 2 хромозома (63).

Пациентите, при които се открива *EML4-ALK* фузията, показват добър отговор към лечение с ALK- тирозин киназни инхибитори. Crizotinib (Pfizer) е одобрен от US FDA през 2011 г. за лечение на ALK-позитивни НДКБК (61). Установяването на *EML4-ALK* позитивните НДКБК е от важно значение за избора на правилна терапевтична стратегия.

Установено е, че лечението с Crizotinib удължава времето до прогресия с над девет месеца, а относителната едногодишна преживяемост е приблизително 75% (61). Въпреки това при значителна част от пациентите раковите клетки придобиват резистентност към Crizotinib, което би могло да се дължи на два основни механизма – ALK-зависим или ALK-независим. При първия механизъм възникват точкови мутации (C1156Y, L1196M, F1174L) или амплификации на *ALK* гена (36,63). При втория механизъм се откриват вторични мутации в *CSMD3*, *CDKN2A*, *MAG11*, *CREBBP*, *DOT1K*, *PBX1* и *PRKDC* гените (64). За преодоляване на резистентността са разработени следващи поколения TKI – ceritinib и alectinib. Те значително по-лесно преминават през кръвно-мозъчната бариера и имат по-добро действие при вторичните лезии, засягащи ЦНС (63).

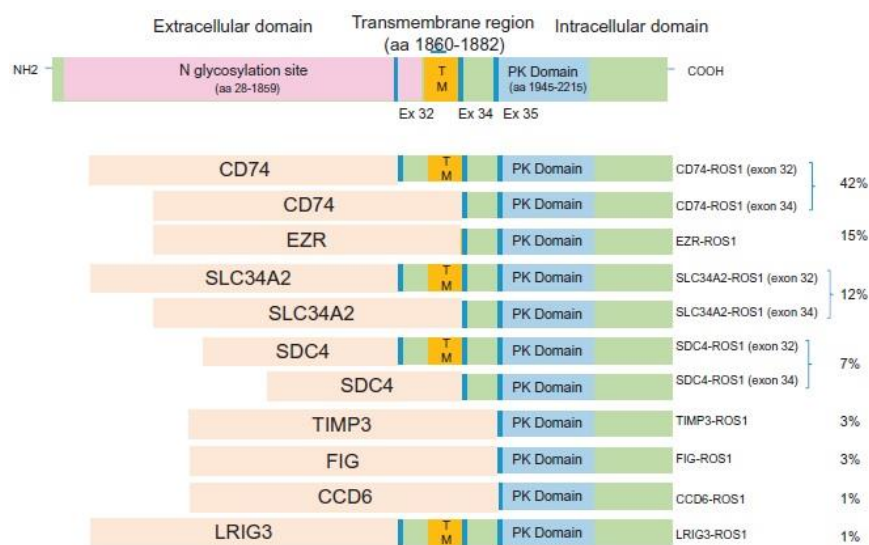
Съществуват различни методи за анализ на вариантите на *EML4-ALK* фузията, като анализ чрез имунохистохимия, флуоресцентна *in situ* хибридизация (FISH), полимеразна верижна реакция чрез обратна транскрипция (RT-PCR), от които с най-висока чувствителност е RT-PCR методът (36). Препоръчаният от FDA метод за скрининг е двойно-флуоресцентно белязан Break-apart FISH. През последните няколко години ограниченията на посочените методи се преодоляват чрез различните техники на секвенирането от следващо поколение (NGS) (65). Те позволяват идентифицирането

на транслокациите на *ALK*, както с известни, така и с нови фузионни партньори, с много висока чувствителност и специфичност (63).

1.5.8. *ROS1* пренареждания.

Генът *ROS1* (OMIM *165020) се намира в хромозома 6 (6q22.1) и кодира трансмембранен тирозин-киназен рецептор (36). Неговият екстрацелуларен N-терминален участък съдържа повече от 1800 аминокиселини, което го превръща в един от най-големите извънклетъчни домени сред всички човешки тирозин киназни рецептори (66). C-терминалната част на рецептора съдържа киназен и трансмембранен домен (67). *ROS1* взема участие в епителната диференциация по време на ембриогенезата на различни органи, но досега не са идентифицирани лиганди за този рецептор (63).

Геномни пренареждания, включващи *ROS1* се срещат в 1-2% от НДКБК (36). Тези тумори показват висока степен на диференциация и често се откриват в напреднал стадий, с нисък процент метастази в ЦНС при диагнозата (63). Пациентите с *ROS1* пренареждания са предимно от женски пол, с ранно начало на болестта, непущачи (36). Както *ALK*, така и *ROS1* транслокациите са взаимноизключващи се с другите „водещи мутации“ (68). Известни *ROS1*-фузионни партньори при НДКБК са гени като *SDC4*, *EZR*, *FIG*, *TPM3*, *LRIG3*, *KDEL2*, *CCRC6*, *MSN*, *TMEM106B*, *TPD52L1*, *CLTC*, и *LIMA1* (63). *CD74-ROS1* е най-често откривана *ROS1*-фузия. При всички известни фузионни гени киназният домен на *ROS1* е напълно запазен, а *ROS1*-мястото на фузия на мРНК-ниво винаги се извършва в 5'-края на екзони 32, 34, 35 или 36 (36). Веднъж експресиран, фузионният онкопротеин димеризира и активира тирозинфосфатазите *SHP-1* и *SHP-2*, които активират от своя страна *RAS/RAF/MAPK*, *JAK/STAT*, *VAV3/RHO* и *PLCγ* сигналните пътища. Това стимулира растежа, пролиферацията и миграцията на раковите клетки. Според вида на фузионния партньор на *ROS1* се активират различни сигнални пътища, което води до различия в характеристиките на съответните тумори (69).



Фигура 1.8. Схематично представяне на ROS1 гена и неговите фузионни партньори при АК. Белтъчните домени са оцветени, както следва: N-гликозилиран сайт (аминокиселини 28-1859, розово), трансмембранен домен (ТМ, аминокиселини 1860-1882, оранжево), тирозин-киназен домен (ПК, аминокиселини 1945-2215, сиво). Местата на транслокация в ROS1 гена за екзони 32 (e32), 34 (e34), и 35 (e35). За всеки фузионен партньор (тъмно розово) са показани честотите на срещане (63).

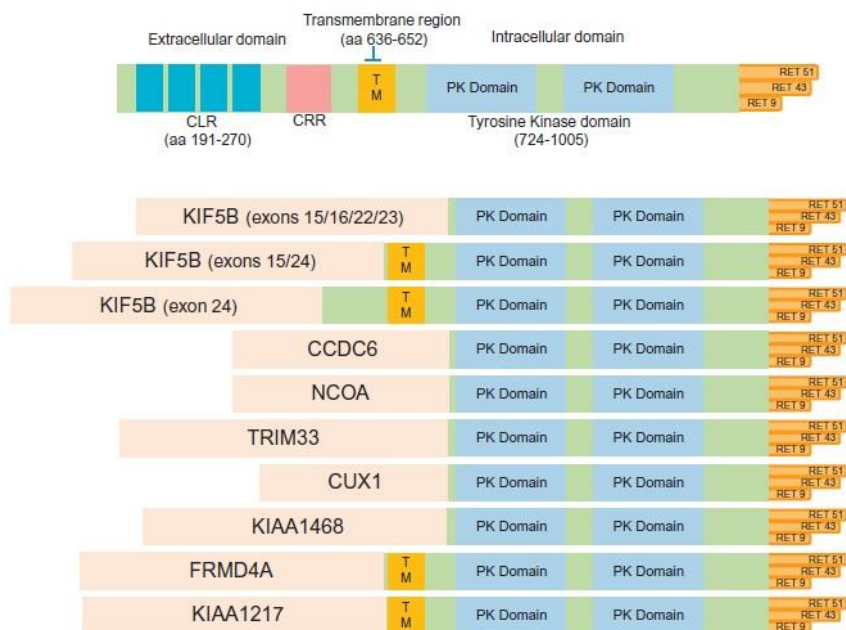
Киназният домен на *ALK* и *ROS1* споделят хомоложност в аминокиселинната си последователност, което прави възможно инхибирането на ROS1 с *ALK*-инхибитори и други TKIs (63). The National Comprehensive Cancer Network (NCCN) препоръчва всички пациенти, негативни за *EGFR*, *ALK* и *KRAS*, да се тестват за други молекулни маркери, вкл. *ROS1* пренареждания (70). Механизмите за придобиване на резистентност на раковите клетки към crizotinib включват поява на вторична мутация (G2032R, L2026M, G2032R, D2033N и S1986F), както и активиране на *KRAS* и *KIT* мутации (69,71).

Няма „златен стандарт“ за детекция на *ROS1*-генни пренареждания. Използват се имунохистохимични методи, RT-PCR, мултиплексен RT-PCR и двуцветна „break apart“ флуоресцентна *in situ* хибридизация (FISH) (36). Чрез NGS е възможно откриването на геномни пренареждания при НДКБК, включително *ROS1/ALK* (39).

1.5.9. RET пренареждания.

RET (OMIM 164761) е трансмембранен тирозин-киназен рецептор, който се експресира по повърхността на нервните и ендокринните клетки. Неговият ген е локализиран на дългото рамо на 10 хромозома и има важно участие в органогенезата на тези органи (36). След свързването с лиганда, рецепторът образува хомодимер, следва автофосфорилиране и активиране на RAS/MAP, PI3K/AKT и PLC γ сигналните пътища. По този начин се стимулира диференциацията, пролиферацията и миграцията на клетките (72).

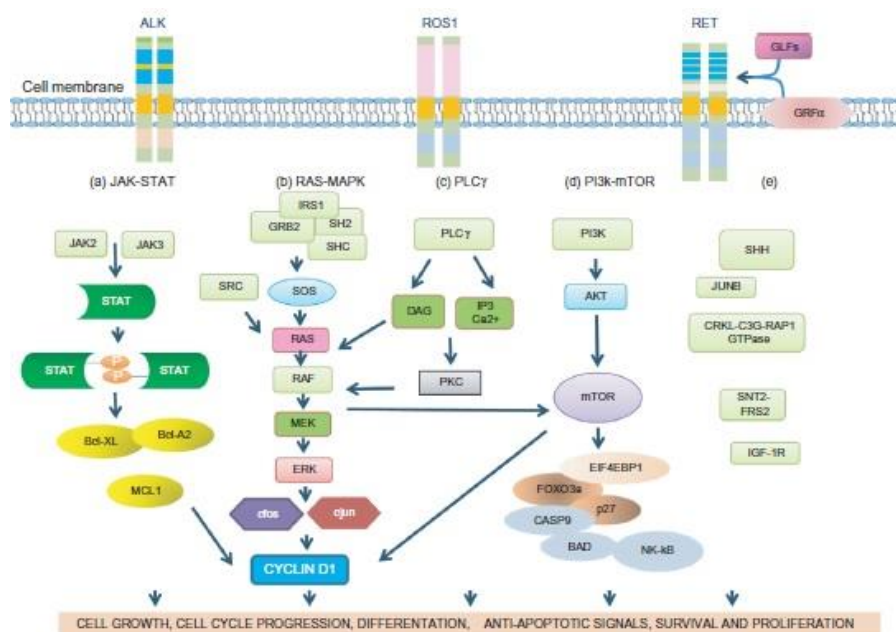
През 2012 г. за първи път е открито геномно пренареждане с участието на *RET* (73). *RET* пренарежданията се срещат при 1-2% от слабо диференцираните АК, в равно съотношение и при двата пола и са взаимноизключващи се с другите „водещи мутации“ (63). При НДРБД са идентифицирани над 12 фузионни партньора на *RET*, от които най-чести са: *KIF5B*, *CCDC6* и *NCOA4*, *TRIM33*, *EPHA5*, *PICALM* (72).



Фигура 1.9. Схематично представяне на *RET* гена и неговите фузионни партньори при АК. (А) Белгъчните домени са оцветени както следва: четири богати на cadherin повтори (CLR, аминокиселини 191-270, тъмно синьо), цистеин богат домен (CRR, тъмно червено), трансмембранен домен (ТМ, аминокиселини 636-652, оранжево), и два разделени тирозинкиназни домена (ПК, аминокиселини 724-1005, сиво). Показани са трите основни изоформи *RET*9, *RET*43, и *RET*51. Фузионните партньори са изобразени със светло розово (63).

Към настоящия момент са разработени или са в процес на клинични проучвания няколко *RET*-инхибитора - *cabozantinib*, *vandetanib*, *ponatinib*, *sunitinib* и *sorafenib* (74). Те се характеризират със сравнително нисък отговор на пациентите към лечение и висока токсичност. Поради тази причина е необходимо по-задълбочено изучаване на биологията на туморите с *RET* пренареждания и разработване на *RET*-инхибитори с по-ниска токсичност и по-добър терапевтичен ефект (74).

В практиката за детектиране на *RET* пренареждания се използва FISH (75). Въпреки това е възможно да се срещнат известни затруднения, според вида на транслокацията (63). Имунохистохимията показва по-малка чувствителност и специфичност за детекция, а RT-PCR обикновено е недостатъчен. NGS технологиите позволяват откриването на фузиите с голяма специфичност и чувствителност, също така и детекцията на нови фузионни партньори на *RET* (63).



Фигура 1.10 Основни сигнални пътища, които се активират от трите тирозин-киназни рецептора - ALK, ROS и RET.

1.5.10. *NTRK1* пренареждания.

NTRK1 е протоонкоген, локализиран на 1q21-q22 хромозома и кодира рецепторна тирозин-киназа, която участва в регулацията на клетъчния растеж и диференциация чрез активирането на RAS/RAF/MEK/ERK и PLC γ сигналните пътища (76). При РБД фузиите, с участието на *NTRK1* гена се откриват при 1-3% от АК (77). Известни фузионни партньори на *NTRK1* са гените, кодиращи RHO-взаимодействащите белтъци MPRIP и CD74 (76). Всички тези фузии водят до конститутивно активиране на киназната активност на *NTRK1* рецептора (77). Разработените *NTRK1*-инхибитори са в ранна фаза на клинични проучвания и показват обещаващи резултати при лечението на НДКБК с *NTRK1* пренареждания (36).

1.6. Имунотерапия при РБД.

1.6.1. PD-1/PD-L1 оста.

Напредъкът на молекулярно генетичните технологии в последните години осигури възможност за нови терапевтични подходи, подходящи за пациентите, носещи определени генетични изменения. Въпреки постиженията на таргетните терапии и развитието на прецизираната медицина, прогнозата остава лоша при голям брой от случаите. Успехът на имуноонкологията при лечението на меланом и бъбречен карцином предостави нова надежда за пациентите с РБД. Постигнати са значими резултати при терапия с имуномодулатори моноклонални антитела, каквито са „чекпойнт“-инхибиторите, като anti-CTLA-4 и anti-PD1.

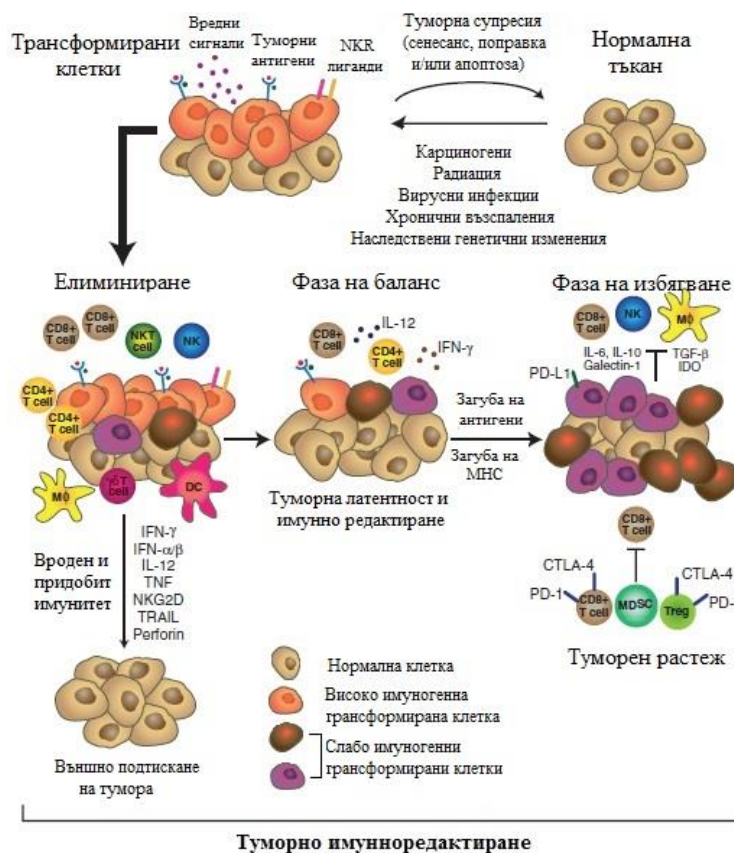
Туморното имунно редактиране представлява възможността на туморната среда да подпомогне раковите клетки да се „изплъзнат“ от имунния отговор (78). То се изразява в три фази – фаза на елиминиране, фаза на баланс и фаза на избягване (78). Във фазата на елиминиране Т-лимфоцитите разпознават и елиминират туморните клетки. Това става чрез разпознаване на специфични тумор асоциирани антигени по тяхната повърхност. Разпознаването им води до активиране на НК клетките и

макрофагите (78). Също така дендритните клетки представят неоантигените в комплекс с МНС на Т-хелперите тип 1 и на цитотоксичните Т-клетки (CTL), които елиминират раковата клетка (78).

Често туморните клетки оцеляват и преминават фазата на елиминиране (78). Във фазата на баланс те продължават да натрупват генетични изменения и увеличават мутационния си товар (78). Това им дава селективно предимство да „избягат“ от имунната система, да пролиферират неконтролируемо и да метастазират (78). Коинхибиторните сигнали на CTLA-4 и PD-1 имат много важна роля по време на фазата на елиминиране (79). Блокирането им може да обърне механизма на тумор-индуцираната супресия и да разреши ефективното действие на имунната система.

PD-1 е трансмембранен гликопротеин, принадлежащ към CD28-семейството. Експресира се по повърхността на редица клетки – Т- и В-лимфоцити, NKT-клетки, активирани моноцити и дендритни клетки (80). При физиологични условия, свързвайки се с лигандите си, той предотвратява Т-клетъчната активация. По този начин при възпаление PD-1 ограничават активността на ефекторните Т-клетки в периферните органи и тъкани и предпазва от автоимунност (80).

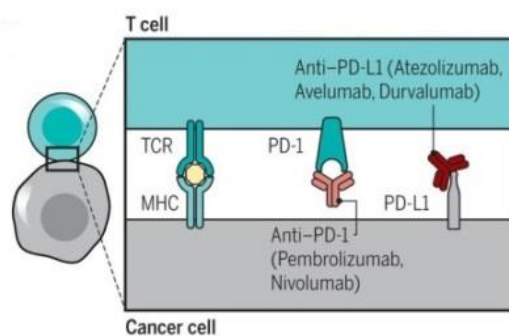
Известни са два лиганда за този рецептор - PD-L1 и PD-L2. PD-L1 се експресира при активирани Т-лимфоцити, дендритни клетки, моноцити и други миелоидни клетки (80). Извън хемопоетичната система PD-L1 се експресира при клетките на белия дроб, ендотелните клетки на кръвоносните съдове, клетки на черния дроб, кератиноцитите и при клетките, участващи в изграждането на плацентата (80). PD-L2 е характерен основно за дендритните клетки и моноцитите (81).



Фигура 1.11. Концепция за туморноимунно редактиране (78).

При 50%-60% от НДРБД туморните клетки експресират PD-L1. Още повече те могат да свръхекспресират PD-L1, за да „избягат” от имунния отговор. Свързването на PD-L1 (рецептор по повърхността на раковата клетка) с PD-1 (рецептор по тумор-инфилтрираните Т-лимфоцити) блокира Т-клетъчната пролиферация и понижава секретирането на антиапоптотични молекули и цитокини (82). Освен това повишената експресия на индолеамин 2,3-диоксигеназа 1 (IDO) допълнително стимулира „избягването” от имунната система.

Разработени са няколко антители, блокиращи взаимодействието на PD-1 и неговите лиганти - PD-L1 и PD-L2. Първият е инхибиране на PD-L1 и възпрепятстване на свързването с рецепторите му PD-1 и B7.1. При този механизъм анти – PD-L1 антителата не инхибират свързването на PD-1 с PD-L2. Вторият начин е инхибиране на PD-1, чрез анти - PD-1 антитела. Така се възпрепятства свързването на PD-1 и неговите лиганди, но не се инхибира свързването на PD-L1 и B7.1. Nivolumab и pembrolizumab са anti-PD-1 човешки IgG4, а atezolizumab се свързва с PD-1 и блокира взаимодействието с PD-L1 и PD-L2 (83).



Фигура 1.12. Инхибиране на PD-1 и PD-L1 и индуциране на противотуморен имуен отговор [78].

Агенция за контрол на храните и лекарствата на САЩ (FDA) и Европейската агенция по лекарствата (EMA) са одобрили като втора линия терапия три PD-1 IgG4 моноклонални антители - atezolizumab, nivolumab и pembrolizumab. Те могат да се прилагат на пациенти, диагностицирани с III туморен стадий и при които резекция или лъчетерапия не са приложими или НДРБД е метастазирал. През 2016 pembrolizumab е одобрен като първа линия терапия при лечение на напреднал белодробен карцином при пациенти, при които не се откриват мутации в *EGFR* и *ALK* гените и експресията на PD-L1 е $\geq 50\%$ (84).

Последни клинични проучвания показват обещаващи резултати при комбинирани терапевтични подходи - anti-PD-1/anti-PD-L1 с химиотерапия, таргетна терапия, IDO – инхибитори и антиангиогенна терапия (84). Въпреки това нито една от комбинациите не показва преимущество. При комбинирано лечение с таргетни терапии и имунотерапия се достига до ниво 3 и 4 на токсичност и това налага спиране на клиничните проучвания. Именотерапевтиците показват подобрена обща преживяемост при пациенти с РБД, при които не се наблюдава отговор след платина-базираната химиотерапия (84). Приложението на имунотерапията в онкологията показва обещаващи резултати при много пациенти, но придобитата резистентност е истинско предизвикателство.

За да се подобри ефикасността на имунотерапията трябва по-добре да се изучи ролята на имунните клетки в туморната среда. Имунохистохимията е един от основните методи за оценка на нивата на експресия на PD-L1, въпреки че не се характеризира с висока специфичност и чувствителност. При едва една трета от новодиагностицираните пациенти експресията на PD-L1 е $\geq 50\%$ (85). Експресията на PD-L1, микросателитната нестабилност и мутационния товар се използват в клиничната практика като маркери за предикция и селекция на пациенти, подходящи за лечение с „чекпойнт“-инхибитори, много от пациентите не показват добър отговор на базата на тези маркери (85). Поради тази причина се налага изучаване и задълбочено разбиране на вътреклетъчните и извънклетъчните сигнални пътища с цел намиране на по-добри биомаркери за предикция на отговора към лечение и прогноза.

1.7. Епигенетични промени при РБД.

В последните две десетилетия нарастват познанията за епигенетичната регулация при РБД. Епигенетичните промени – ДНК метилиране, експресията на нкРНКи (non-codingRNA), ремоделирането на хроматина и посттранскрипционните регулатори, са ключови изменения в карциногенезата на белодробните тумори, които водят както до значително по-агресивен фенотип, така и до резистентност към терапии. Измененията на епигенетичните модификации могат да бъдат надеждни биомаркери за стратификация на риска, ранна диагноза, прогноза, молекулярна класификация и предикция за таргетна терапия. Терапевтичният потенциал на епигенетичните таргетни лекарства в комбинация с химиотерапия, таргетна терапия и/или имунотерапия са в процес на изследване.

1.7.1. Хистонови модификации, асоциирани с РБД.

Най-високата степен на организация на ядрената ДНК е хроматинът. Негова основна единица е нуклеозомата. Тя се състои от белтъчно ядро (осем хистона), около които се усуква нишка от ДНК с дължина 147 bp (86). Хистоните участват в регулацията на транскрипцията, поправката и репликацията на ДНК (87). Структурно те притежават подвижна „опашка“, която се подлага на различни пост-транслационни модификации като метилиране, ацетилиране, фосфорилиране и др. (88). Метилирането на хистоните е свързано едновременно с транскрипционна активация и репресия. Метилирането на лизиновите остатъци в позиции 4, 36 и 79 на хистон H3 (H3K4me3, H3K36me и H3K79me) е характерно за активно транскрибиращите се гени, докато метилирането на лизиновите остатъци на 9-та и 27-ма позиция на хистон H3 (H3K9 и H3K27) и лизин 20 на хистон H4 (H4K20) е характерно за неактивния хетерохроматин (86). Ацетилирането на определени аминокиселини в хистоновите опашки намалява афинитета на хистоните към ДНК, което води до формирането на „отворена“ хроматинова конформация, позволяваща осъществяването на транскрипция.

По този начин, чрез модификация на хистоновите белтъци се повлиява свързването на различни транскрипционни фактори и белтъци, участващи в ДНК репликацията и комплексите за поправка на ДНК (89,90). Така се променя генната експресия на клетката и се подпомага туморогенезата. Йонизиращи лъчения, различни химични агенти, тежки метали, арсен, минерален прах и тютюнев дим са посочени като човешки канцерогени от група I в IARC класификацията (91). Тези фактори на

околната среда променят генетичните и епигенетични профили на нормалните белодробни клетки и водят до възпаление, фиброза и рак (92).

MINA53 е генът, експресията на който се индуцира от минерален прах. Открит е за първи път при изследване на алвеоларни макрофаги на миньори, работещи във въгледобивната индустрия (93). *MINA53* (наричан още *mdig*, *MINA53/MINA*, *NO52*, *JMJD10*, *RIOX2*) се локализира на дългото рамо на трета хромозома (3q12.1). Кодира белтък с размер 53-kDa, открит в ядрото (94). Белтъкът *MINA53* съдържа JmjC-домен, структурен мотив за хистонови H3K9me3 деметилази от JmjC-семейство и повлиява експресията на различни протоонкогени (95). Също така, *MINA53* може да действа като рибозомна оксигеназа и да хидроксилира хистидинов аминокиселинен остатък на 39 позиция на рибозомният протеин L27a. По този начин, по време на процеса на сглобяване на пре-рибозомни частици, *MINA53* участва в рибозомната биогенеза (96).

Последните проучвания съобщават за повишена експресия на *MINA53* при различни видове рак, включително рак на белия дроб, рак на гърдата, глиобластом, рак на дебелото черво, плоскоклетъчен карцином на хранопровода, лимфом, бъбречен карцином, хепатоцелуларен карцином, рак на панкреаса, множествен миелом и рак на стомаха (94,97–104). Въпреки това, точният механизъм на действие на *MINA53* все още не е добре изяснен. Доказано е, че *MINA53* може да действа като онкоген подпомагайки туморогенезата и стимулирайки пролиферацията на раковите клетки (105,106). Също така *MINA53* участва в белодробно възпаление и имунния отговор (107,108). Интересно е, че *MINA53* може да има функцията и на тумор-супресор, в зависимост от стадия и вида на рака, и да потиска клетъчния растеж и инвазия. Изследването на Thakur и съавт. при рак на гърдата показва повишена експресия на *MINA53* гена в ранните стадии на рак на гърдата и някои подтипове неинвазивен рак на гърдата, като медуларен карцином, дуктален карцином *in situ* и муцинозен карцином. Експресията на *MINA53* обаче е намалена в силно инвазивния тройно негативен карцином на гърдата и в раковите клетки, които метастазират в близките лимфните възли (109). При РБД експресията на *MINA53* е с различна прогноза, в зависимост от хистологичния подтип и туморния стадий. Нивата на *MINA53* корелират с по-лоша преживяемост при прогресия на белодробните АК, но подобна тенденция не се наблюдава при ПК (92,110). В авансирал рак на белия дроб, *MINA53* инхибира метастазирането (111).

Komiya и съавт., анализират клетъчни линии, трансфектирани с *MINA53*. Използвайки микрочипове за гена експресия, екипът открива 125 гени с повишена експресия и 129 гени с понижена експресия. Сред тях 59 гена участват в процесите на инвазия, метастазиране и транспорт, 34 гена са свързани с транскрипционна регулация, 20 гена с метаболизма на рака, 17 гени свързани с цитокини или растежни фактори и 12 гена участват в клетъчния цикъл и пролиферацията. *MINA53* регулира гени за клетъчна адхезия, метаболизъм на глюкоза, липиди и минерали (111).

Aziz и съавт., използват различен подход при рак на стомаха. Те извършват анализ публично достъпните бази от данни и функционално обогатяват генните, ко-експресирани с *MINA53*. Онтологичният анализ (GO-анализ) и анализът на засегнатите сигнални пътища (KEGG анализ) показват групиране на процесите в 12 клъстера. Това са процеси свързани с регулиране на транслацията на белтъците; регулиране на клетъчното делене; образуването на сплайсозома; зреене на мРНК; вътреклетъчен транспорт на вирусни частици и транспорт на белтъци в клетката; ЕПР и везикуло-медиран транспорт на Голджи; поправка на ДНК и регулиране на сигналната

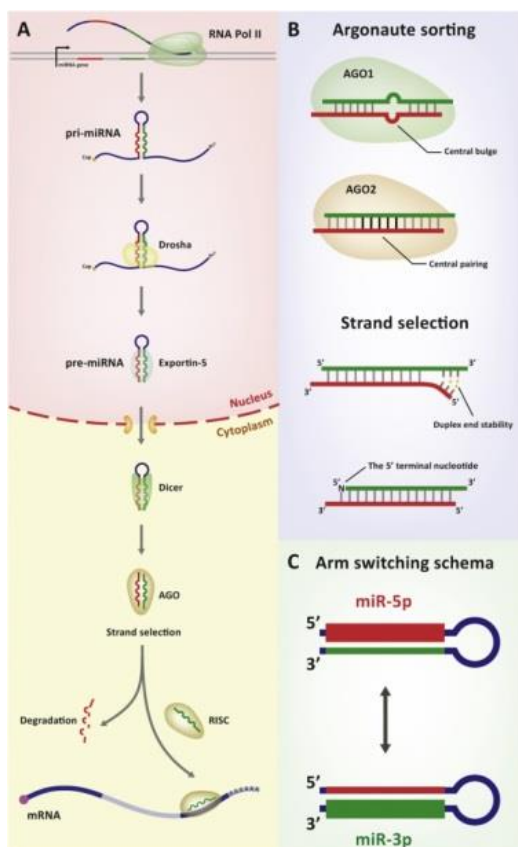
трансдукция чрез p53; T-клетъчен/ лектин тип C /Fc-епсилон рецепторен сигнален път; протеин киназна активност, клетъчен цикъл и регулиране на апоптотичните процеси; структурна асемблиране на рибозомите; анафаза-стимулиращ комплексно-зависим катаболитен процес и клетъчната адхезия (112).

Изясняването на ролята на *MINA53* гена и свързаните с нея сигнални пътища, може да бъде полезно при разработването на стратегия за предотвратяване или ограничаване на прогресията на рака на белия дроб.

1.7.2. МикроРНКи.

МикроРНКите са малки консервативни ендогенни РНК молекули с дължина 17 до 23 нуклеотида. Те участват в редица биологични пътища – апоптоза, диференциация, пролиферация, отговор към стресови фактори и др. МикроРНКите осъществяват посттранслационен контрол на генната експресия. Каноничният модел на биогенеза на тези малки некодиращи РНК молекули е строго контролиран от множество ензими, в резултат на което се синтезират три основни РНК-а продукта: първични (pri-), прекурсорни (pre-), и зрели микроРНКи (113). Зрелите микроРНКи, получени от 5' (miR-5p) или 3' рамото (miR-3p) на прекурсорните РНКи се отбелязват с наставките -5p или -3p (114). Едно от рамената е функционално активно (miR) и се инкорпорира в РНК индуцирания комплект (RISC), а другото се счита за съпътстващ продукт (miR*) и обикновено се насочва за разграждане (115). Според вида на тъканта и на клетката, в която се експресират, изборът на функционално активно рамо на прекурсорните РНКи може да е различен. Феноменът, който се наблюдава, е известен като „arm switching“ (115). Механизмите, по които се избира коя от двете вериги на гените, кодиращи микроРНКи, да е функционално активна, не са изяснени. Една от хипотезите е, че водеща е термодинамичната стабилност (116).

Известни са повече от 2500 микроРНКи и всяка една от тях регулира редица гени. Нуклеотидните последователности на 5' и 3' рамената са различни. Двете вериги могат да регулират експресията на различни таргетни молекули или да се свържат с едни и същи иРНКи, но последователностите на свързване са различни в 3'-нетранслиращия регион на иРНК-мишена (114). Заглушаването на експресията може да се осъществи или чрез репресия на транслацията, когато последователността от 2–7 нуклеотида от miRNA (т.нар. “seed region”) е перфектно комплементарна на таргетната иРНКа, или когато има непълна комплементарност между микроРНКа и иРНКа мишената (114).



Фигура 1.13. Канонична биогенеза на микроРНК и механизъм за arm switching. (А) Биогенезата на микроРНК е показана от първичния транскрипт до таргетирането на иРНК, чрез RISC-комплекса. (В) Аргонавт (AGO) сортира и разпределя двете вериги на микроРНК дуплекса. Двете рамена са оцветени в червено и зелено. (С) Механизъм за селекция на една от двете вериги. Сините линии представят hairpin – структури. Червените и зелените - miR-5p и miR-3p веригите. Удебелените линии показват експресионните нива на двете вериги (114).

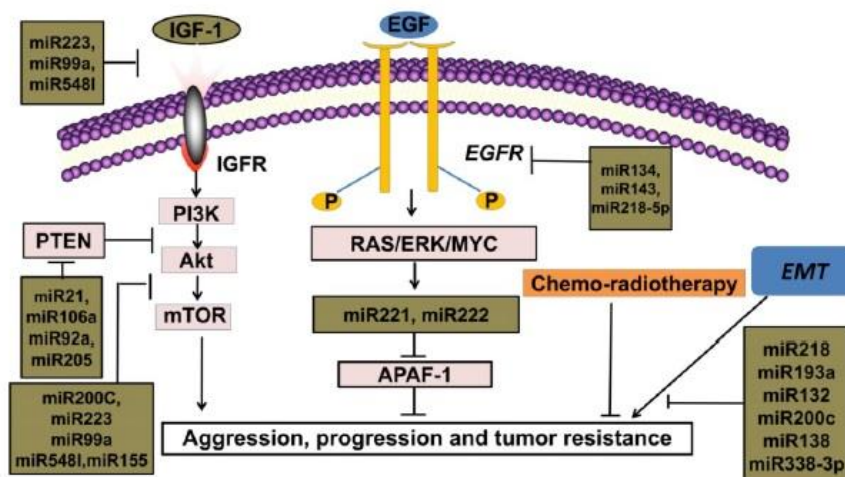
Дерегулацията на микроРНК-ите е свързана с епигенетични или генетични изменения - аберантно ДНК метилиране, амплификации, делеции или точкови мутации (117). Гените, кодиращи микроРНК често са локализирани в „чупливи региони” на хромозомите, в много малки региони на загуба на хетерозиготност и „break point” региони, което показва, че те вероятно са нов клас гени, които имат роля в туморогенезата (117). Еднонуклеотидни замени в пре-микроРНКите може да повлияят процесинга, експресията и взаимодействието на дадената микроРНКите с мРНК (118).

Профилите на микроРНКите са силно специфични за всеки клетъчен тип (119). При РБД чрез микрочипов анализ при сравняването на туморната и нормалната тъкан може ясно да се види промяната на експресията на микроРНКите (118). Много от тях са с повишена експресия в туморната тъкан спрямо нормалната тъкан и имат онкогенна функция (118).

В сравнение с традиционните радио- и химиотерапия, таргетните терапии се характеризират с по-голяма специфичност и ефективност. Известно е, че свързването на тирозин-киназните инхибитори с АТР – свързващото място на EGFR рецептора, блокира низходящите сигнални пътища. Само пациентите с активиращи EGFR мутации могат да се възползват от това лечение (119).

Скорешни изследвания показват, че много микроРНК са с променена експресия и оказват влияние върху чувствителността на раковата клетка към TKIs (120). Например една от целите на miRNA-200c е ZEB1- белтък, който инхибира PI3K/Akt сигналния път и по този начин увеличава чувствителността към лечение с gefitinib (120). Също така miRNA-200c увеличава и чувствителността към лъчелечение, като възпрепятства действието на гените, отговорни за елиминирането на свободните радикали в клетката и поправката на ДНК (121). По същия начин miRNA-143, miRNA-134, miRNA-218-5p директно се свързват с 3'UTR на иРНК, кодираща EGFR (122,123).

По този начин те инхибират миграцията и пролиферацията на туморната клетка. MiR-30b, miR-30c, miR-221 и miR-222 индуцират резистентност към gefitinib при gefitinib-чувствителни клетъчни линии с мутации в *EGFR* (124). Инхибирането им възвръща чувствителността на клетките към gefitinib. Високите нива в плазма на miR-30b и miR-30c са предиктивен маркер за намален отговор при лечение с erlotinib при пациенти с мутация в *EGFR* (124).



Фигура 1.14. МикроРНКи участващи в регулирането на таргетните терапии, химиотерапия и лъчетерапия (119).

MiRNA-223, miRNA-99a и miRNA-548 регулират експресията на инсулин-зависимите растежни фактори, фосфоинозитол-3 киназа, протеин киназа B и mTOR (119). Така те инхибират пролиферацията, миграцията и инвазията на малигнената клетка. Някои изследователски групи докладват микроРНКи, които успешно, с висока специфичност и чувствителност могат да разграничат пациенти с мозъчни метастази от тези без метастази (125).

Някои микроРНКи могат да увеличат чувствителността към платина-базираната терапия. Например мишена на miRNA-106b-5p е PKD2, а при свързването си с HMGA2 miRNA-26 забавя клетъчния растеж и увеличава отговора към терапия с платина (119).

Семейството на miRNA-17 има централна роля в регулацията на TGF β сигналния път, като по този начин е възможно да обърне резистентността към терапия с платина при пациенти с РБД (126). Генът *FZD1*, който е свързан с туморната пролиферация и резистентност към химиотерапия е мишена на miRNA-135b (127). Установено е, че miRNA-9600 участва в регулацията на експресията на STAT3 и е възможно да инхибира клетъчния растеж и да увеличи чувствителността на раковата клетка към raslitaxel (128).

Доказано е, че miR-21 е с повишена експресия и се свързва с лимфни метастази и лоша прогноза при НДРБД (122). Тя може да стимулира туморогенезата на НДРБД чрез индиректно усилване на Ras/MEK/ERK пътя и инхибиране на клетъчната апоптоза (129). Има данни, че miR-21 стимулира растежа и инвазията на NSCLC чрез директно насочване към PTEN (130). Мултивариабилан Cox регресионен анализ показва, че miRNA-21 е независим прогностичен фактор при пациентите с НДРБД. Намалената експресия на miRNA-21 в клетъчни линии инхибира пролиферацията и прави клетките чувствителни към лъчелечение (131).

MicroRNA-210 е една от миРНК-ите, чиито нива на експресия показват значителни разлики в различните етапи на процеса на диференциране на епителните

клетки (132). В някои проучвания съобщават, че експресията на miR-210 е намалена по време на епително-мезенхимния преход на туморните клетки (EMT), аберантното активиране на което води до метастази (133). В същото време е докладвано многократно, че нивото на miR-210 е повишено в плазма и серум на пациенти с белодробен карцином (131).

Ниските нива на експресия на miR-126-3p и miR-451a се свързват с по-напреднал стадий, голям диаметър на тумора и метастази в лимфни възли (134). За miR-126-3p са идентифицирани 154 таргетни гени (например *PLXNB2*), участващи в сигналните пътища на апоптоза и канцерогенеза (134). За miR-451a са идентифицирани 397 таргетни гена, участващи в 5 сигнални пътища, включително „PPAR сигнален път“ (134). Резултатите от проведените проучвания показват, че miR-126-3p и miR-451a могат да се използват като прогностичен биомаркер за тежестта на протичане на белодробния аденокарцином (134). Установено е, че miR-126-3p е сред 10-те микроРНК с повишена експресия при сравнение на серум от болни с белодробен карцином, спрямо пациенти с бенигни белодробни заболявания и здрави контроли, но не е сред най-значимите 4 (miR-23b, miR-221, miR-148b и miR-423-3p), които могат да се използват за диагностика при белодробен карцином (134).

MiR-30a е дерегулирана в няколко злокачествени типа солидни тумори, като рак на гърдата, хепатоцелуларен рак, рак на дебелото черво, назофарингеален карцином, рак на простатата, рак на ендометриума и кожен плоскоклетъчен карцином (135). Генът за hsa-miR-30a, член на miR-30 семейството, е локализиран на хромозома 6q13, където при РБД често настъпва загуба на хетерозиготност (135). Поради тази причина се наблюдава намалена експресия на тази микроРНК, която участва в контрола на множество биологични и патологични процеси, включващи растежа, диференциацията и апоптозата.

Няколко проучвания съобщават за понижената регулация на miR-30a в тъканите на РБД или в култивирани клетки (135). Експресията на miR-30a е значително понижена в тъканите на НДКБК в сравнение със съседните нормални тъкани. Намалената експресия на miR-30a е свързана с лоша прогноза при РБД (136). От друга страна miR-30a може да модулира чувствителността на НДКБК чрез намаляване на функцията на ATF1 във фосфорилирането на ATM и има потенциална терапевтична стойност (135).

Kumarswamy и съавтори доказват, че при РБД miR-30a-5p инхибира епително-мезенхимната транзиция като инхибира *Snail*, а също така участва в сигналните пътища, отговорни за чувствителността на клетките към raplataxel (137). Доказано е, че повишената експресия на miR-30a увеличава чувствителността на раковите клетки към raplataxel като намалява експресията на *BCL-2* и отключва апоптозата (138). При намалена експресия на miR-30a се увеличава експресията на *BCL-2* и резистентността на клетките към raplataxel (138). *MYBL2* има предполагаемо miR-30a прицелно място и е доказана отрицателна корелация между експресията на miR-30a и *MYBL2* при НДКБК (135).

Генът за hsa-miR-195 е локализиран на хромозома 17p13.1 (139). Тази микроРНК принадлежи към microRNA-15/16/195/424/497 семейството и има тумор-супресорна функция. Ролята на hsa-miR-195 в туморогенезата и прогресията при РБД е неизяснена. MiR-195 инхибира развитието на раковите клетки като регулира експресията на *MYB*, *CHEK1*, *HDGF*, *IGF1R*, *CCND3* и *BIRC5* (139). Намалената

експресия на miR-195 води до нарушения по оста miR-195/BIRC5 и допринася за прогресията при АК (138).

В последно време нарастват изследванията, които показват, че е възможно и двете рамена -5p и -3p на микроРНКи-те да се асоциират с RISC – комплекса (140). По този начин miR-5p/miR-3p осъществяват синергична регулация и усилват онкогенния ефект на сигналните пътища, а чрез потискане действието на тумор-супресорните гени увеличават клетъчната пролиферация, миграция и инвазия (141).

Генът за hsa-miR-140 е локализиран на хромозома 16q22.1 и има тумор-супресорна функция (142). И двете зрели miR-140-3p и miR-140-5p са с намалена експресия и се асоциират с прогресия на заболяването при различни ракови заболявания, включително рак на гърдата, остеосаркома, рак на дебелото черво. Мишени на miR-140-3p са *EGFR* и *IGF1R*, с ключова роля за клетъчната пролиферация, оцеляване и ангиогенеза (143). Друга мишена е *ERK1*, което води до намаляване на активността на c-FOS и c-JUN – отговорни за инвазията и трансформацията на раковите клетки. MiR-140-5p може да потиска *VEGFA* и *PTEN* и това води до заглушаване на PI3K/AKT сигналния път (142).

Клъстерът miR-143/145, състоящ се от две микроРНКи - hsa-miR-143, hsa-miR-145 потиска важни гени, участващи в туморогенезата, като *KRAS* и *Era*. Детектираната понижена експресия на miR-145-3p и miR-145-5p води до онкогенна активация на MTDH, медиатор на PI3K/AKT, NFκB, MAPK и Wnt – сигналните пътища (144). Експресията на hsa-miR-145 се контролира от няколко транскрипционни фактора и различни епигенетични механизми. P53 е основен клетъчен тумор-супресор и контролира експресията на няколко микроРНКи - miR-34, miR-192/215, miR-107 и miR-145. Sachdeva и съавтори показват, че p53 може да индуцира експресията на miR-145 като се свърже директно с p53 в промотора на гена и оттам p53 потиска експресията на c-Myc, един от таргетите на тази микроРНК (145). Suh и съавтори доказват, че ниската експресия на miR-145 корелира с мутационния статус на p53 (146).

Друга микроРНК с онкогенна функция е miR-197 (147). Тя играе основна роля в патогенезата на НДКБК като инхибира p53-зависимата апоптоза, чрез инхибирането на тумор-супресора FUS1 (148). MicroRNA-197 индуцира епително-мезенхимния преход, спомага инвазията и прогресията и контролира експресията на PD-L1 чрез miR-197/CKS1B/STAT3-медираната сигнална каскада (148). Намалена експресия на miR-197 е свързана с увеличена чувствителност към химиотерапия (148). MiR-197 и miRNA-184 често са с повишена експресия при пациенти с мутация в *EGFR* и наличие на мозъчни метастази. Поради тази причина тези две микроРНКи могат да послужат като биомаркер за риска от мозъчни метастази при тази група (149).

Dacic и съавтори установяват по-високата експресия на miR-25 при АК с настъпила мутация в EGRF гена (150). Експресията на miR-155 е двукратно увеличена при див тип EGRF/KRAS АК и заедно с miR-21 инхибира апоптозата и води до туморна прогресия (150). Друга микроРНК с по-висока експресия при АК е miR-495 и се асоциира с възникнали мутации в *KRAS* гена (150).

Освен в тъканите, микроРНКи могат да се детектират и в телесните течности – периферна кръв, серум, плазма, урина, цереброспинална течност и слюнка. Свободно циркулиращите нуклеинови киселини активно се освобождават от раковите клетки както самостоятелно, така и пакетирани в екзозоми и се характеризират с висока стабилност в циркулацията. Профилирането на микроРНКите е бъдеща посока за

развитие на неинвазивните биомаркери, защото подпомагат диагнозата, прогнозата и терапията при РБД (151).

Откриването на нови молекулярно генетични биомаркери при РКБ и използването им в клиничната практика дава възможност за по-ранното диагностициране, прецизиране на диагнозата, прогноза и предикция на отговора към терапия. Течните биопсии, служат като алтернативен начин за неинвазивна диагностика, но също така поради туморната хетерогенност, допълват резултатите, получени от тъканните биопсии. С развитието на технологиите се увеличава специфичността и чувствителността за лечение и мониторинг на ранния и напреднал РБД.

1.8. Приложение на машинното самообучение в генетиката и геномиката.

Исторически погледнато, областта на изкуствения интелект, в която терминът "машинно самообучение", става популярен в края на 90-те години на миналия век. Машинното самообучение се занимава с разработването и прилагането на компютърни алгоритми, които се подобряват с натрупването на опит (152). Тези алгоритми намират широко приложение в областта на генетиката и геномиката, интерпретирайки големи масиви от геномни данни и аотирайки елементи на геномната последователност. Например методите за машинно самообучение могат да се използват, за да се "научат" как да се определят местоположението на началните места на транскрипцията в ДНК последователността (153). По подобен начин могат да се използват алгоритми за идентифициране на сплайсинг местата (154), промоторните области (155), енхансери (156) или местата за свързване на нуклеозомите (157). Освен това моделите, които разпознават всеки отделен тип геномен елемент, могат да се комбинират заедно с научената логиката за относителното им местоположение, за да се изградят системи за машинно самообучение, които са в състояние да аотират гени по протежение на цели еукариотни хромозоми - включително техните нетранслирани области (UTR), интрони и екзони (158).

Освен че се научават да разпознават модели в ДНК, алгоритмите могат да се използват за анализиране на резултатите от микрочипови експресионни анализи или РНК секвениране (RNA-seq), ДНК секвениране, ChIP секвениране и др. Данните за генната експресия могат да се използват, за да се научат алгоритмите да разграничават, както различните хистологични типове НДРБД така и различните морфологичните типове АК и ПК. Също така е възможно се идентифицират потенциално важни предиктивни и прогностични биомаркери (159).

В ерата на имуноонкологията типа на НДРБД и неговият морфологичният подтип има съществена роля при определянето на оптималното лечение на пациенти с НДРБД (160,161). Например хирургичната резекция е основният начин на лечение на НДРБД, докато ДРБД обикновено се лекува с химиотерапия и облъчване дори в локализиран стадий (162). При прогресия на заболяването режимите на химиотерапия също са специфични за различните хистологични подтипове. Понастоящем механизмите, които са в основата на хистологичната клетъчна съдба, са неизвестни. Разбирането на молекулярните детерминанти на хистологията може да даде представа за различния отговори към лечение и за по-ефективното използване на хистологията за насочване на лечението на НДРБД. Въпреки че, големи проучвания, като например в TCGA, показват, че геномните характеристики до голяма степен се различават между различните хистологични подтипове на белодробния рак (163–165), геномните

изменения не винаги съвпадат с хистологичните подтипове. Таргетни геномни изменения, като мутации на *EGFR* и транслокации на *ALK/ROS1*, които са характерни за повечето пациенти с АК, са докладвани при някои пациенти с ПК (166,167). Това навежда на мисълта, че хистологията не се определя предимно от геномните характеристики. Тези данни показват, че е възможно хистологията на белодробния рак да се определя на транскриптомно ниво, въпреки че точните механизми на регулиране на генната експресия все още не са определени (168).

Голямо разнообразие от алгоритми за машинно самообучение са разработени, за да помогнат да се разберат механизмите, лежащи в основата на генната експресия. Някои модели имат за цел да предскажат експресията на даден ген само въз основа на ДНК последователността (169), докато други вземат предвид профилите на хистоновите модификации (170) или свързване на транскрипционни фактори в промотора на гена област (171). По-сложните методи, чрез обучение на невронна мрежа, се опитват съвместно да моделират експресията на всички гени в клетката (172).

Много от изброените по-горе проблеми също могат да бъдат решени с помощта на методи от областта на статистиката. Всъщност границата между машинното самообучение и статистиката в най-добрия случай е неясна (173). Някои методи са общи и за двете области и всеки от тях може да се използва за прогнозиране и изводи. Въпреки това, машинното самообучение и статистиката имат различен фокус - предсказване или извод (174). Като цяло класическите статистически методи разчитат на предположения за системите, генериращи данни. Статистиката може да осигури ясни изводи чрез напасване на определен вероятностен модел, когато са събрани достатъчно данни от добре планирани изследвания. Машинното самообучение се занимава с въпроса за създаването и прилагането на алгоритми, които се подобряват с натрупването на опит. Много от методите за машинно самообучение могат да извличат модели за разпознаване на образи, класификация и предсказване от съществуващи данни и не разчитат на строги предположения за системите, генериращи данни. Това ги прави по-ефективни в някои сложни приложения, но по-малко ефективни при създаването на ясни модели с биологично значение.

Съществуват три основни типа методи за машинно самообучение: контролирано самообучение, неконтролирано самообучение и подсилено самообучение. Алгоритмите за първия тип изучават връзката между набор от входни променливи и определена зависима променлива или етикети от обучаващите случаи и впоследствие могат да се използват за прогнозиране на резултатите от нови случаи. Такива методи са например, линейна регресия, логистична регресия, дърво на решенията, метод на опорните вектори и невронна мрежа. Алгоритмите за неконтролирано самообучение извеждат модели от данните без зависима променлива или известни етикети. Клъстерният анализ и анализът на принципните компоненти са два популярни метода за неконтролирано самообучение, които се използват за намиране на модели в данни с висока размерност. Дълбокото обучение е подвид на машинното учене, първоначално вдъхновен от неврологията, като по същество описва клас от големи невронни мрежи. Дълбокото обучение може да бъде контролирано, така и неконтролирано. То направи революция в разпознаване на изображения, има обещаващи приложения в геномиката и медицината и се прилага в много области, до голяма степен благодарение на увеличението на изчислителната мощност и на големите обеми данни (175).

ЧАСТ 2

ЦЕЛ И ЗАДАЧИ

2.1. ЦЕЛ.

Изследване ролята на микроРНК, техни мРНК-таргети и соматични мутации в гени, свързани с процесите на онкогенеза при рак на белия дроб и идентифициране на най-подходящи неинвазивни диагностични и прогностични биомаркери.

2.2. ЗАДАЧИ.

- Подбор на пациенти с рак на белия дроб.
- Събиране на свежи тъканни материали и разширение на създадената биобанка към Център по Молекулна Медицина, МУ-София.
- Изолиране на РНК и ДНК от включените биологични материали.
- Провеждане на микрочипов експресионен анализ на 24 пациента.
- Избор на панел от микроРНКи за валидиране чрез PCR в реално време, SybrGreen технологията.
- Сравнение на експресионните нива на микроРНКи спрямо наличните клиникопатологични характеристики на пациентите.
- Корелация между експресионните нива на микроРНКи и мРНКи.
- Провеждане на таргетно секвениране на соматични мутации в онкопанел.
- Статистически и биоинформатичен анализ на получените резултати.

ЧАСТ 3

МАТЕРИАЛИ

И

МЕТОДИ

3.1. Клиничен материал.

Набирането на пациентите и изследването на биологичния материал се основава на колаборацията между Център по Молекулна Медицина, Катедра Медицинска Химия и Биохимия и Клиниката по Гръдна Хирургия на Специализирана Болница за Активно Лечение по Белодробни Болести (СБАЛББ) „Св. София” към Медицински Факултет, Медицински Университет-София. Настоящото проучване включва 124 пациента, приети по повод извършване на торакотомия или белодробна резекция в различен обем - ексцизия, сегментектомия, лобектомия или пулмонектомия, по повод тумор на белия дроб. На всички включени участници им беше предоставена информация за извършените изследвания и участието им бе потвърдено чрез подписване на информирано съгласие. При всички пациенти се проведе имунохистохимично изследване на туморния материал и бяха диагностицирани като аденокарцином или плоскоклетъчен карцином белия дроб.

Изследванията, включени в дисертационния труд бяха осъществени благодарение на спечелени проекти към ФНИ, МОН и МУ-София в следния ред: проект № / Договор №ДН03/14 от 19.12.2016, Договор № Д-138/02.05.2017, Договор № Д-138/02.05.2017 и Договор № Д-№ D-70/23.04.2019.

3.2. Биологичен материал.

3.2.1. Клинични показатели на изследваната група.

В настоящото проучване са включени 124 пациента, преминали през Клиниката по Гръдна Хирургия и диагностицирани с аденокарцином и плоскоклетъчен карцином на белия дроб. По време на оперативната интервенция се отделяше от всеки пациент, материал от туморната маса и здрава белодробна тъкан. Така отделените тъкани се съхранявани в Центъра по Молекулна Медицина на -80°C до момента на изолиране на РНК и ДНК. Допълнително се отделяше материал от туморната част за имунохистологично изследване.

От събраните аденокарциноми на белия дроб и съответващите им нормални тъкани подбрахме 24 (12 тумора и 12 нормални тъкани) и от събраните плоскоклетъчни карциноми подбрахме 24 (12 тумора и 12 нормални тъкани) за микроРНК експресионното профилиране чрез микрочипов анализ. От тях изолирахме тотална РНК, част от която използвахме за микрочиповия анализ, а останалата част съхранихме в РНК биобанката в Центъра по Молекулна Медицина.

За валидиране на резултатите от микрочиповия анализ чрез PCR в реално време бяха подбрани 50 пациента с аденокарциноми на белия дроб и прилежащите им нормални тъкани и 50 пациента с плоскоклетъчни тумори и нормални тъкани. От изолираната тотална РНК бе синтезирана кДНК, част от която бе използвана за провеждане на експресионния анализ на подбрани микроРНКи. Останалата част съхранихме за научни бъдещи проучвания в сферата на рака на белия дроб. При подбраните 100 пациента с не-дребноклетъчен белодробен карцином беше проведен и PCR в реално време за анализ на експресията на подбрани мРНКи, мишени на избраните за валидация микроРНКи и също така свързани с способността на туморите

да избягват от имунната система. Подбрани 24 пациента бяха подложени на таргетното секвениране на панел от 48 гена за детекция на соматични мутации.

На Таблица 3.1. и Таблица 3.2. е представено групирането на включените и изследваните пациенти.

Таблица 3.1. Клиникопатологични характеристики на пациентите, включени с тъканен материал в микрочиповия анализ.

Характеристика	Аденокарциноми N=12	Плоскоклетъчни тумори N=12
Жени	5	4
Мъже	7	8
Средна възраст жени	62	61
Средна възраст мъже	61	68
Туморен стадий (T)		
T1	2	5
T2	8	5
T3	2	2
Лимфни метастази (N)		
N0	10	8
N1	1	2
N2	1	2
Метастази (M)		
M0	9	10
M1	3	2
Диференциация на тумора		
G1	2	3
G2	5	2
G3	5	6

Таблица 3.2. Клиникопатологични характеристики на пациентите, включени с тъканен материал в валидация на микрочиповия анализ чрез PCR в реално време.

Характеристика	Аденокарциноми N=50	Плоскоклетъчни тумори N=50
Пол		
Жени	17 (24%)	8 (16%)
Мъже	33 (66%)	42 (84%)
Възраст		
≤ 60	26 (52%)	21 (42%)
> 60	24 (48%)	29 (58%)

Туморен стадий (T)		
Ранен стадий (T1 –T2)	35 (70%)	27 (54%)
Късен стадий (T3 – T4)	15 (30%)	23 (46%)
Лимфни метастази (N)		
N0	32 (64%)	30 (60%)
N1 - N2	18 (36%)	20 (40%)
Метастази		
M0	33 (66%)	35 (70%)
M1	17 (34%)	15 (30%)
Диференциация на тумора		
G1 добре диференциран	11 (22%)	14 (15.4%)
G2 умерено диференциран	22 (44%)	18 (19.8%)
G3 слабо диференциран	17 (34%)	18 (19.8%)
Локализация на тумора		
Ляв лоб	18 (36%)	27 (54%)
Десен лоб	31 (62%)	23 (46%)

3.3. ПРЕДАНАЛИТИЧНА ОБРАБОТКА НА БИОЛОГИЧНИТЕ МАТЕРИАЛИ.

3.3.1. Изолиране на тотална РНК от тъкан.

За изолирането на тотална РНК, включваща и малки РНКи от тъкан беше използван кита miRNeasy Mini Kit (Qiagen, Хилден, Германия). На Таблица 3.4. са представени материалите и реактивите, които са включени в използвания кит.

Принцип на работа

Реагентите от miRNeasy кита комбинират фенол/гуанидин базирано лизиране на пробите и силициево базирана мембрана за изолиране на тотална РНК. QIAzol Lysis реагентът е монофазов разтвор на фенол и гуанидин тиоцианат. Той улеснява разграждането на тъканите, инхибира РНазите и също така посредством органична екстракция премахва клетъчната ДНК и белтъците от лизата. Тъканните проби бяха хомогенизирани в QIAzol Lysis реагент, чрез използването на хомогенизатор ИКА 3386000 (Cole-Parmer, Канада). След добавянето на хлороформ и последващо центрофугиране, хомогената се разделя на две основни фази: водна и органична. В горната фаза се съдържа РНК, в междинната фаза ДНК, а в долната органична фаза са белтъците и липиди.

Горната, водна фаза се отделя, и се добавя етанол, за да подпомогне свързващите условия за всички РНК молекули, по дълги от 18 нуклеотида. Пробата след това се прехвърля в RNeasy MinElute колонка, където РНК молекулите се свързват

за силициево базирана мембрана. Следват серия от промивания, с цел промиване на съдържащия се фенол и други компонент. Висококачествена РНК се елуира в малък обем вода, пречистена от РНази. На Фигура 3.1. е представена схематично протокола за изолиране на тотална РНК.

3.3.2. Изолиране на тотална ДНК от тъкан.

За изолирането на висококачествена геномна ДНК от тъкани бе използван кита QIAamp DNA Mini Kit (Qiagen, Хилден, Германия).

Принцип на работа

Чрез QIAamp DNA Mini Kit бързо и лесно се получава пречистване на тотална ДНК, подходяща за последващите аналитични методики. Тотална ДНК (вкл. геномна, вирусна и митохондриална) може да бъде изолирана от различни тъкани, съдебно-медицински проби или култивирани клетъчни култури. Количеството на изходния материал, който може да се ползва с този кит е до 25мг тъкан.

Пречистването и изолирането се основава на фенол/хлороформната екстракция или алкохолно преципитиране. Накрая ДНК е елуирана в специфичен (АЕ) буфер или вода, готова за използване в последващите аналитични методики. Изолираната ДНК е свободна от белтъци, нуклеази или други замърсители и инхибитори.

Размерът на изолираната ДНК чрез QIAamp Kit варира до 50кб като преобладават фрагментите от 20-30кб. ДНК с такава дължина денатурира напълно по време на PCR реакцията и може да бъде амплифицирана с висока ефективност.



Фигура 3.1. Схематично представяне на работния протокол на изолиране на тотална РНК, вкл. и малки РНКи от животинска тъкан.



Фигура 3.2. Визуализиране на процедурата по QIAamp DNA Mini Kit

3.3.3. Оценка на качеството и концентрацията на изолирана тотална РНК/ДНК

3.3.3.1. Определяне концентрация на изолирана тотална РНК/ДНК чрез NanoDrop 2000.

За измерване концентрацията на изолираната тотална РНК и ДНК първоначално беше използван спектрофотометричен метод. Това бе осъществено с помощта на NanoDrop 2000 (Thermo Scientific, Масачузетс, САЩ). Методът се основава на специфичното поглъщане на светлина с дължина на вълната $\lambda = 260 \text{ nm}$ от пуриновите и пиримидиновите бази на ДНК и РНК. Абсорбцията при дължина на вълната $\lambda = 280 \text{ nm}$ дава концентрацията на белтъците.

Отношението на поглъщането на пробата, измерени съответно при дължина на вълната 260 nm и 280 nm , показва чистотата на ДНК/РНК. Чистите ДНК проби дават абсорбция $260/280$ съотношение приблизително 1.8. Стойностите под 1.6 показват наличието на примеси от белтъци, а над 2 - примеси от РНК. Чистите РНК проби дават абсорбция $260/280$ съотношение в обхвата 2-2.2. Стойности под 1.8 показват примеси от белтъци. Върху съотношението на абсорбция $260/280$ влияние оказва и киселинността на пробите. Съотношението абсорбция $260/230$ дава представа за съдържанието на други контаминиращи вещества, които абсорбират при дължина на вълната 230 nm или 260 nm . Чистите проби имат съотношение на абсорбцията $260/A230$ в обхвата 2-2.2. Вещества, които поглъщат при дължина на вълната 230 nm са EDTA, въглехидратите и TRIzol (разтвор, който се използва за изолиране на РНК), гуанидинев хлорид (разтвор, който се използва за изолиране на ДНК). При контаминация с тях съотношението на абсорбцията $260/A230$ се повишава. Вещество, което поглъща при $\lambda = 260$ е гуанидин изотиоцианат, използван за изолиране на РНК. При контаминиране с него съотношението на абсорбцията $260/230$ има високи стойности.

3.3.3.2. Определяне концентрация на изолирана тотална РНК/ДНК с флуориметричен метод с помощта на Qubit v2.0

Спектрофотометричното измерване е широко използван метод за определяне на концентрацията на изолираната ДНК или РНК, но то не е достатъчно точно и надеждно. UV базираните абсорбционни измервания не са селективни и не могат да правят разграничение между ДНК, РНК, белтък, или свободни аминокиселини и свободни нуклеотиди. Така получените стойности за концентрацията могат да се повлияят от варирането на съотношението на базите в измерваните нуклеинови киселини и от присъствието на други контаминиращи вещества като соли, органични съединения или свободни нуклеотиди. В допълнение, чувствителността на спектрофотометричните измервания е ниска и е невъзможно измерването на ДНК/РНК с много ниски концентрации.

Флуоресцентно базираното определяне на концентрацията на нуклеиновите киселини е по-чувствителен и специфичен метод, тъй като прави разграничение между

ДНК и РНК и белтъци, и дава възможност да се измерват проби с много ниска концентрация. При измерване на концентрацията чрез Qubit v2.0 флуорометъра (Life Technologies, Калифорния, САЩ) се използват специфични багрила за ДНК и РНК и белтъци. Когато не са свързани към съответната си таргетна молекула (ДНК, РНК или белтък) тези багрила имат изключително ниска флуоресценция. При свързването си те започват интензивно да флуоресцират. След като се добавят към разтвор на пробите, багрилата се свързват към прицелните молекули за няколко секунди и се достига равновесие за по-малко от 2 минути.

При определено количество на багрилото, количеството на флуоресцентния сигнал от пробата е директно пропорционален на концентрацията на измерваната молекула в разтвора. След това Qubit флуорометърът отчита флуоресцентния сигнал и определя концентрацията като използва два стандарта с известна концентрация. Чрез стандарти се определя съотношението между флуоресценцията и ДНК/РНК концентрацията. След това полученото съотношение се използва за пресмятане на концентрацията на измерваните проби.

3.3.3.3. Определяне на концентрацията и RINe на изолираната РНК от тъкан чрез Agilent 2200 TapeStation Systems.

TapeStation системата, подобно на Qubit флуорометърът, използва флуоресцентни багрила специфични за молекулите, които се измерват. От една страна системата определя концентрацията на изследваната проба и от друга страна осъществява капилярна електрофореза, чрез която се определя качеството на изолираните нуклеинови киселини.

При измерване на РНК проби, се определя RIN (RNA integrity number). При капилярната електрофореза РНК молекулите се разделят и на електрофореграмите са ясни два основни пика, представляващи 18S и 28S рибозомна РНК (рРНК). В зависимост от метода на съхранение на пробите, изолирането и пречистването на РНК, то могат да се наблюдават и допълнителни малки пикове, съответстващи на 5S рРНК и други малки некодирани РНК молекули. При наличие на деградиране на РНК молекулите, постепенно пиковете при 18S и 28S рРНК намаляват и се появяват ниско молекулни пикове. Съотношението 28S/18S рРНК се използва за изчисляване на интегритета на РНК молекулите. RIN заема числова стойност от 1 до 10 и отразява степента на деградация на РНК молекулите в изследваните проби като RIN = 10 показва възможно най-добро качество РНК, а RIN = 1 показва силно разградена РНК. TapeStation системата определя RINe, което е еквивалент на RIN определен от Agilent 2100 Bioanalyzer. Разликата между RINe и RIN е в алгоритъма за пресмятането им. РНК проби със стойности на RIN=5 са подходящи за PCR в реално време анализи. Желателно е пробите за микрочипов анализ да имат стойности на RIN над 7-8.

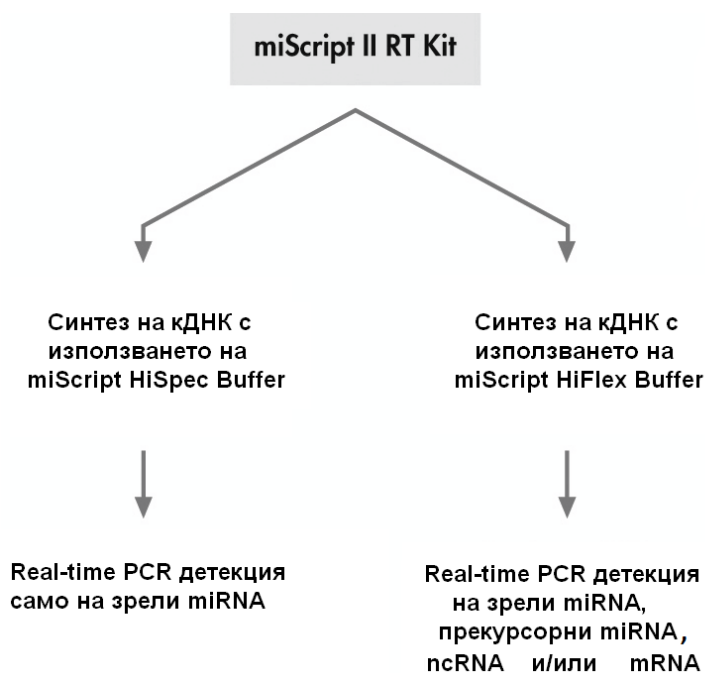


Фигура 3.3. Протокол за работа с RNA ScreenTape, TapeStationSystem.

3.3.4. Обратна транскрипция на изолираната РНК за синтез на кДНК.

За пробразуването на малки некодиращи РНКи в кДНК беше използван специфичен за превръщането, освен на иРНК, но и на малки нкРНК кит miScript II RT Kit (Qiagen, Хилден, Германия) (Таблица 3.10.).

miScript Reverse Transcriptase Mix е оптимизиран микс от поли (А) полимераза и обратна транскриптаза. 10x miScript Nucleic Mix съдържа dNTPs и oligo-dT праймери. Двата буфера, miScript HiSpec и miScript HiFlex, се използват в зависимост от целите на изследването (Фигура.3.4.). Така получената кДНК може да бъде използвана за PCR в реално време, с използването на подходящ праймер.



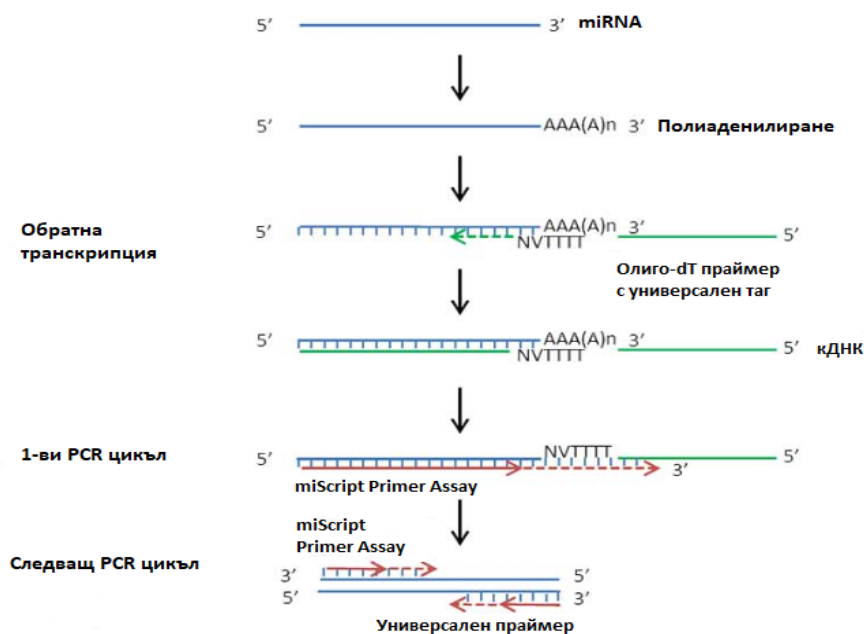
Фигура.3.4. Детекция на зрели miRNA, прекурсорни miRNA, други нкРНКи или mRNA

Принцип на работа

Зрелите микроРНКи са с дължина около 22-нуклеотидни последователности и за разлика от иРНКи, микроРНКите не са полиаделинирани. На Фигура.3.5. е представено схематично селективното превръщане на зрели микро РНКи в кДНК с помощта на miScript HiSpec буфер.

Обратната транскрипция бе осъществена с помощта на miScript II RT Kit. В кита са включени два различни буфера: HiSpec и HiFlex, чието предназначение е различно (Фигура 3.4.). В нашето проучване ние използвахме HiFlex буфер за синтез на кДНК само от зрелите микроРНКи и иРНК. Когато обратно транскриптазната реакция се извършва в присъствието на miScript HiFlex Buffer, зрелите микроРНКи, и прекурсорна микроРНКи, друга малки некодиращи РНК (snoRNAs, snRNAs) и иРНК се конвертират в кДНК. Тази кДНК може след това да се използва за количествено определяне на всеки вид РНК (използвайки подходящи праймери). Зрелите микроРНКи се полиаденилират от poly(A) полимераза и обратно се транскрибират в кДНК с помощта на олиго-dT праймери. Полиаденилирането и обратната транскрипция се осъществява паралелно, в една епруветка. Олиго-dT праймерите притежават поли-Т последователност в 3' края си и универсална последователност - таг в 5' края си, което

позволява амплификация на зрелите miRNAs в следващата стъпка, а именно real-time PCR (PCR в реално време). miScript Primer Assays използвани в комбинация с miScript SYBR Green PCR Kit спомагат за количественото определяне на зрелите микроРНКи чрез PCR в реално време. Комбинацията от полиаденилиране и добавяне на универсален праймер осигуряват специфичност на miScript Primer Assays и невъзможност за неспецифична амплификация на геномна ДНК.



Фигура.3.5. Селективно обръщане на miRNA в кДНК.

Като изходен материал за синтеза на кДНК използвахме 200ng тотална РНК. В таблица 3.3 са показани реактивите използвани за RT-PCR и техните количества.

Таблица 3.3. Компоненти на RT-PCR реакцията.

Компонент	Количество (µl) за една реакция
5x miScript HiSpec Buffer	4 µl
10x miScript Nucleics Mix	2 µl
RNase- free H ₂ O	Варира
miScript Reverse Transcriptase Mix	2 µl
РНК матрица	Варира
Краен Обем	20 µl

Условията за протичане на обратната транскрипция са представени в таблица 3.4.

Таблица 3.4. Условията за протичане на обратната транскрипция

Температур	Време (мин)
а	
37°C	60
95°C	5
4°C	изчакване/∞

3.4. АНАЛИТИЧНА ОБРАБОТКА

3.4.1. микроРНК експресионно профилиране чрез микрочипове на туморна и нормална белодробна тъкан

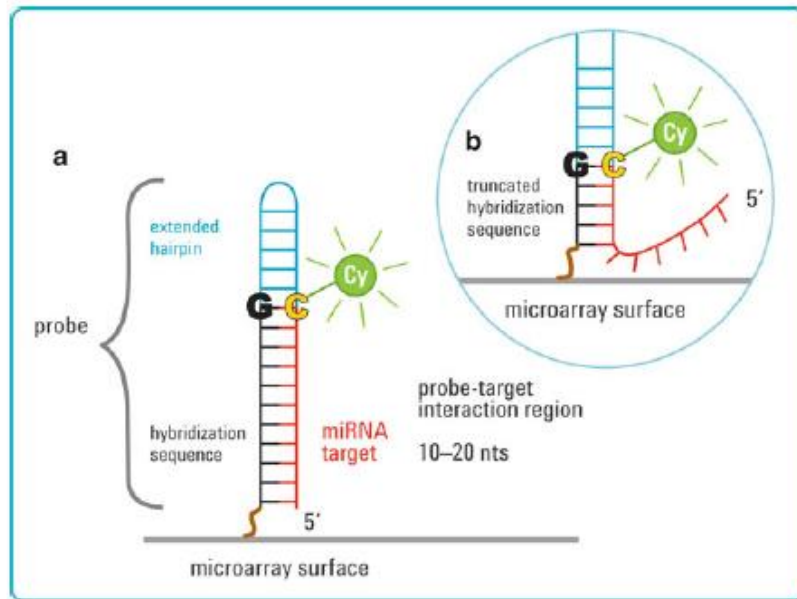
МикроРНК експресионното профилиране бе осъществено с помощта на Agilent miRNA System. Етапите на експресионно профилиране чрез микрочипове включват:

- Избор на туморна и нормална белодробна тъкан от биобанката в ЦММ
- Изолиране на РНК от белодробните тъкани
- Определяне на концентрацията и качеството на изолираните РНК молекули
- Подготовка на РНК пробите
- Белязане на РНК пробите с Cyanine3-рСр
- Пречистване на белязаните РНК проби
- Хибридизационна реакция
- Промиване на микрочиповете
- Сканиране на микрочиповете с помощта на Feature Extraction Software
- Визуализация, детекция и анализ на резултатите с помощта на GeneSpring

Принцип на работа

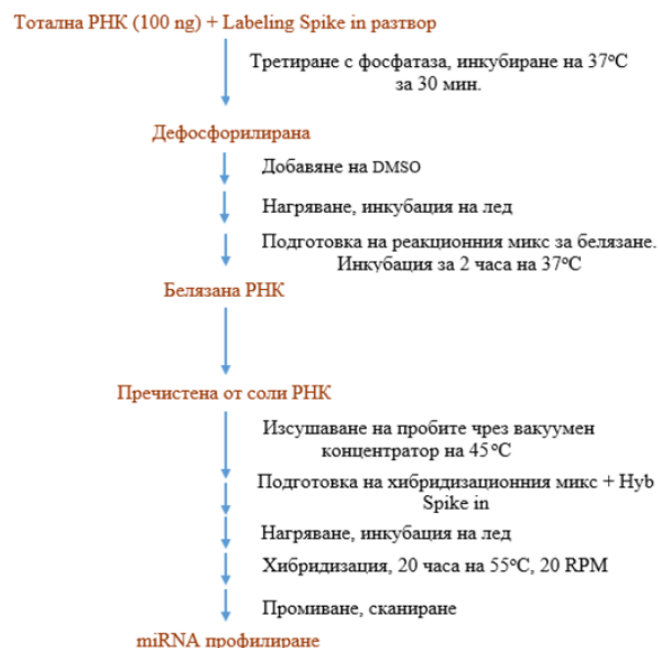
Профилирането на експресията на микроРНК-ите представлява голям интерес поради значимата роля на тези къси некодирани РНК молекули в клетъчните процеси. Експресионните профили на микроРНК-ите са асоциирани с нарушения на клетъчната регулация при различните видове рак.

Изследването на експресията на микроРНК-ите може да бъде осъществено чрез системата на Agilent за микрочипов анализ. За провеждане на анализа 100 ng тотална РНК, след нагряване в присъствието на DMSO, се бележи чрез лигиране на Cy-3 рСр към 3' края на РНК. Белязаният цитозин взаимодейства с гуанидин в 5' края на пробата (сондата) и така се придава стабилност на хибридизационния комплекс. Сондите са с дизайн, осигуряващ разграничаване според последователността, но също и според размера. По този начин зрелите микроРНК-ите могат да се детектират с високо специфичност. Пробите са с балансирана Tm и включват фуркетна последователност в 5' края на пробата (Фигура 3.6).



Фигура.3.6. Дизайн на Agilent пробите. Немодифицираната проба от микрочипа (в черно) е синтезирана последователност, която хибридира към таргетна miRNA (в червено). Пробите са закрепени към повърхността на стъклото (слайда) чрез краче (в кафяво). (a) Включването на G остатък (в черно) към 5' края на хибридиращата последователност комплементира към 3' крайния C остатък (в жълто), който се въвежда при етапа на белязането. G-C двойката в проба-таргет региона стабилизира взаимодействието. В допълнение всички сонди (проби) съдържат 5' фуркетна структура (в синьо), определяща региона на взаимодействие проба-таргет и стабилизираща свързването и повишаваща специфичността според размера. (b) Температурата на топене (T_m) на сондите е еднаква за целия микрочип. За сонди, изискващи намаляване на температурата на топене това се осъществява чрез последователно елиминиране на бази до достигане на подходяща дължина в 3' края на хибридиращата последователност така че проба - таргет взаимодействието е в рамките на 9-20 нуклеотида (176).

Процедурата за подготовката на пробите и обработка на микрочиповете е показана на Фигура.3.7.



Фигура.3.7. Работен ход за подготовка на пробите за микрочипов експресионен анализ.

За осъществяването на микрочиповия анализ следвахме протокола “miRNA Microarray System with miRNA Complete Labeling and Hyb kit”, version 3.1.1, August 2015 на фирмата производител. Използвахме Agilent G4870C Humanmi RNA Microarray Kit Release 21.0, 8x60K формат и с AMADID № 070156. На всяко от 8-те полета на този вид чипове са включени 2549 човешки микроРНКи. Беше използван Agilent microRNA Spike-In кит, състоящ се от два microRNASpike-In раствора, служещи като технологични контроли на белязането и хибридизацията.

Таблица 3.5. Етапи в микрочиповия анализ и използвани консумативи

Етап	Компонент	Количество (µl) за една реакция
Подготовка на Spike In разтворите- приготвяне на серийни разредки	Labeling Spike-In	1
	Hyb Spike-In	1
Подготовка на РНК пробите	тотална РНК (50 ng/µl)	2
	10 x Calf Intestinal Phosphatase Buffer	0.4
	Labeling Spike-In (Third dilution)	1.1
	Calf Intestinal Phosphatase	0.5
	Краен обем	4
Денатурация	100% DMSO	2.8
Белязане (Лигазна реакция)	10 x T4 RNA Ligase Buffer	1
	Цуанине3-рСр	3
	T4 RNA Ligase	0.5
	Краен обем	11.3
Пречистване на белязаните РНК-и с Micro Bio-Spin 6 Columns и Сушене	RNase-Free H ₂ O	50
Хибридизационна реакция и Хибридизация за 20 часа	RNase-Free Water	17
	Hyb Spike-In	1
	10 x GE Blocking Agent	4.5
	2X Hi-RPM Hybridization Buffer	22.5
	Краен обем	45
Микрочип	Agilent G4870A, miRBase release 16, 8x60k	-
Промиване	Agilent GE Wash Buffer 1	600 ml
	Agilent GE Wash Buffer 2	300 ml

3.4.2. Детекция на изследваните некодиращи РНК молекули посредством PCR в реално време.

Конвенционалната полимеразна верижна реакция (ПВР; англ. PCR) се състои в амплифицирането на специфичен участък ДНК в *in vitro* условия. Това се осъществява чрез олигонуклеотидни фрагменти (праймери), които на база комплементарност обграждат интересуващата ни последователност от ДНК и чрез термостабилна полимераза се размножават чрез циклични повторения на три основни стъпки - денатурация на двойната спирала, хибридизация на праймерите и синтез. При PCR в реално време (Real Time PCR; RT-qPCR) резултатите се отчитат чрез флуоресцентни сигнали. Различните разновидности на тази методика, позволяват отчитането на сигнал под формата на флуоресцентна крива след всеки цикъл на реакцията. Чрез детайлното следене на реакцията и подробен анализ на резултатите се прави количествен анализ на експресията. Съществуват редица флуорофори, които могат да интеркалират между базите на ДНК молекулата. Благодарение на планарната си структура се свързват с нея и излъчват флуоресцентен сигнал. При всяка амплификация в реално време се генерира сигнал, който след определен брой цикли се отчита като покачване на общия сигнал. В зависимост от началното количество на специфичната последователност за размножаване, цикълът, при който флуоресцентната крива пресича линията на позитивиране (threshold line) е различен. На базата на това се изчислява Ct стойност, чрез която се извършва качествен (напр. отчита се само наличие на точката на пресичане) и количествен анализ (напр. определяне на нивото на експресия спрямо вътрешна контрола за даден молекулен маркер), т.нар. RT-qPCR.

Относителната генна експресия на изследваните микроРНК-и бе определена с помощта на $2^{-\Delta\Delta Ct}$ метода описан в точка 3.4.7. За анализ на микроРНК-и в тъкан използвахме SybrGreen технологията.

Принцип на работа:

Sybr Green багрилото се свързва към двойноверижна ДНК и по този начин се засичат PCR продуктите, които се акумулират при протичането на PCR циклите. Sybr Green багрилото се свързва към всички двойноверижни ДНК молекули, включително и към неспецифичните продукти. Също така чрез този метод не могат да се осъществяват мултиплексни PCR реакции. По тази причина е много важно да се оптимизират PCR реакциите за да се получават надеждни резултати.

Принцип на действие на SybrGreen технологията:

- Sybr Green I багрилото флуоресцира, когато е интеркалирано (свързано) към двойноверижна ДНК;
- По време на PCR полимеразата амплифицира таргетната последователност, което води до създаването на PCR продукти, наречени ампликони;
- При напредването на PCR циклите се създават повече ампликони. Тъй като Sybr Green се свързва към всички двойноверижни ДНК-и това води до увеличение на флуоресцентния сигнал пропорционално на количеството PCR продукти.

При анализ на експресията на дадена микроРНК, всяка проба бе пускана в две отделни PCR реакции, всяка в три повторения. В едната PCR реакция се слагат праймерите за изследваната микроРНК, а в другата се използват праймерите за амплификация на

избраната ендогенна контрола. За нормализация бе използвана *RNU6B* като ендогенна контрола. Като калибратори бяха използвани нормалните белодробни тъкани.

Компонентите необходими за изготвянето на PCR микс са представени в Таблица 3.6.

Таблица.3.6. Компоненти за осъществяването на PCR в реално време за определяне на експресията на зрели микроРНК-и в 384-ямкови плаки

Компонент	Обем/реакция (384-ямкова плака)
2x QuantiTect SYBR Green PCR Master Mix	5мкл
10x miScript Universal Primer	1мкл
10x miScript Primer Assay	1мкл
Вода без РНази	Варира
cDNA матрица	1мкл
Краен обем	10мкл

Условията на PCR реакцията в реално време, които използвахме са обобщени в Таблица 3.7.

Таблица 3.7. Условия за протичане на PCR реакцията в реално време.

Стъпка	Време	Температура	Допълнителни коментари
Първоначална стъпка на активация на PCR	15 мин	95°C	HotStar Taq ДНК полимеразата се активира при тази стъпка
Условия на 3- стъпалния PCR			
Денатурация	15 сек	94°C	Събиране на флуоресцентните сигнали
Анилинг	30 сек	55°C	
Екстензия (Елонгация)	30 сек	70°C	
Брой цикли	40 цикъла		

Таблица 3.8. Използвани, в дисертационния труд праймери за микроРНКи и техните секвенции.

Зряла miRNAs (изоформа)	Нуклеотидна последователност (5'-3')	Каталожен номер (Qiagen)
hsa-miR-744-3p	CUGUUGCCACUAACCUCACCCU	MS00010556
hsa-miR-21-3p	CAACACCAGUCGAUGGGCUGU	MS00009086
hsa-miR-650	AGGAGGCAGCGCUCUCAGGAC	MS00005313
hsa-miR-375	UUUGUUCGUUCGGCUCGCGUGA	MS00031829
hsa-miR-140-3p	UACCACAGGGUAGAACCACGG	MS00008673
hsa-miR-144-3p	UACAGUAUAGAUGAUGUACU	MS00020328
hsa-miR-195-5p	UAGCAGCACAGAAAUUUGGC	MS00003703

hsa-miR-195-3p	CCAAUUAUUGGCUGUGCUGCUCC	MS00008953
hsa-miR-139-5p	UCUACAGUGCACGUGUCUCCAGU	MS00003493
hsa-miR-4689	UUGAGGAGACAUGGUGGGGGCC	MS00037548
hsa-miR-7-5p	UGGAAGACUAGUGAUUUUGUUGUU	MS00032116
hsa-miR-95-3p	UUCAACGGGUUUUAUUGAGCA	MS00010906
hsa-miR-101-3p	UACAGUACUGUGAUACUGAA	MS00008372
hsa-miR-6785-5p	UGGGAGGGCGUGGAUGAUGGUG	MS00046914
has-RNU6-2	ACGCAAATTCGTGAAGCGTT	MS00033740

3.4.3. Детекция на мРНКи посредством PCR в реално време (real time-qPCR).

За провеждането на PCR в реално време на мРНКи бяха използвани реактиви от QuantiTect SYBR® Green PCR Kit. Обема на използваните реактиви за подготовка на реакциите са показани на следващата таблица.3.9.

Принцип на работа:

QuantiTect SYBR Green PCR kit съдържат оптимизиран и готов за използване основен микс за високо специфично и чувствително количествено определяне на cDNA с помощта на SYBR Green I багрилото. Балансирана комбинация от K⁺ и NH⁴⁺ йони в буфера към кита подпомага свързването на праймерите и осигурява протичането на PCR реакцията с висока специфичност и чувствителност. В допълнение, HotStarTaq DNA полимераза осигурява „горещ старт“ на PCR реакцията, предотвратявайки образуването на неспецифични продукти.

Обема на използваните реактиви за подготовка на реакциите са показани на следващата таблица.3.9.

Таблица.3.9. Реактиви, необходими за една реакция RT-qPCR

Компонент	Обем/реакция (384-ямкова плака)
2x QuantiTect SYBR Green PCR Master Mix	5 мкл
10x QuantiTect Primer Assay‡	1 мкл
Вода без РНази	3 мкл
cDNA матрица	1 мкл
Краен обем	10мкл

Използваните праймери за детектиране на избраните мРНК са представени на Таблица 3.10. АСТВ (Actin) беше използван като ендогенна контрола, за изчисляване нивата на експресия на PD-L1, STAT3, PTEN, EGFR, MINA53 и mTOR.

Таблица.3.10. Използвани праймери за детекция на иРНКи и ендогенна контрола:

иРНК	Каталожен номер (Qiagen)
PD-L1	QT00082775
STAT3	QT00068754
PTEN	QT00086933
EGFR	QT00085701
MINA53	QT00075544
mTOR	QT00056133
ACTB	QT00095431

Така подготвените 384-ямкови плаки, съдържащи съответната реакция, бяха поставени в апарат Applied Biosystems 7900HT Fast Real-Time PCR System (Thermo fisher Scientific, Масачузетс, САЩ) на програма представена на Таблица.3.11.

Таблица 3.11. Условия за протичане на real time PCR

Стъпка	Време	Температура	Допълнителни коментари
Първоначална стъпка на активация на PCR	15 мин	95°C	HotStar Taq ДНК полимеразата се активира при тази стъпка
Условия на 3- стъпалния PCR			
Денатурация	15 сек	94 °C	Събиране на флуоресцентните сигнали
Анилинг	30 сек	55 °C	
Екстензия (Елонгация)	30 сек	70 °C	
Брой цикли	40 цикъла		

3.4.4. Секвениране от ново поколение, с използването на TSACP и секвенционна платформа MiSeq (Illumina).

На 12 пациента с АК и 12 пациента с ПК на белия дроб беше приложено таргетно секвениране за детекция на соматични мутации в значими тумор-асоциирани гени TruSeq Amplicon-Cancer Panel (TSACP) на платформа за следващо поколение секвениране MiSeq (Illumina, Калифорния, САЩ). При подготовката на пробите са секвениране, подготовката на библиотеките и самото им секвениране на MiSeq платформата се следваха инструкциите на производителя.

На Фигура.3.6. е представен списък на гените, които се включват в TSACP панела. Той съдържа олигонуклеотидни проби за секвенирането на мутационни горещи точки в над 35 килобазова таргетна геномна последователност. В рамките на високо мултиплексна единична реакция, 48-те гена са представени от 212 ампликона. Размерът на ампликоните е 170-190 базови двойки. TSACP панелът е много подходящ за ДНК проби, изолирани от формалин фиксирани и парафин включени проби.

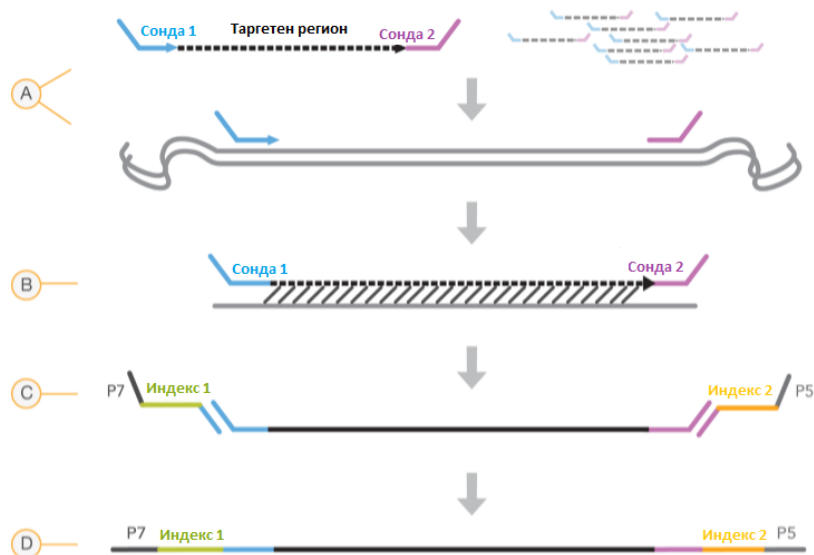
<i>ABL1</i>	<i>EGFR</i>	<i>GNAS</i>	<i>MLH1</i>	<i>RET</i>
<i>AKT1</i>	<i>ERBB2</i>	<i>HNF1A</i>	<i>MPL</i>	<i>SMAD4</i>
<i>ALK</i>	<i>ERBB4</i>	<i>HRAS</i>	<i>NOTCH1</i>	<i>SMARCB1</i>
<i>APC</i>	<i>FBXW7</i>	<i>IDH1</i>	<i>NPM1</i>	<i>SMO</i>
<i>ATM</i>	<i>FGFR1</i>	<i>JAK2</i>	<i>NRAS</i>	<i>SRC</i>
<i>BRAF</i>	<i>FGFR2</i>	<i>JAK3</i>	<i>PDGFRA</i>	<i>STK11</i>
<i>CDH1</i>	<i>FGFR3</i>	<i>KDR</i>	<i>PIK3CA</i>	<i>TP53</i>
<i>CDKN2A</i>	<i>FLT3</i>	<i>KIT</i>	<i>PTEN</i>	<i>VHL</i>
<i>CSF1R</i>	<i>GNA11</i>	<i>KRAS</i>	<i>PTPN11</i>	
<i>CTNNB1</i>	<i>GNAQ</i>	<i>MET</i>	<i>RB1</i>	

Фигура.3.8. Раково – свързани гени, включени в TSACP панела.

Принцип на работа:

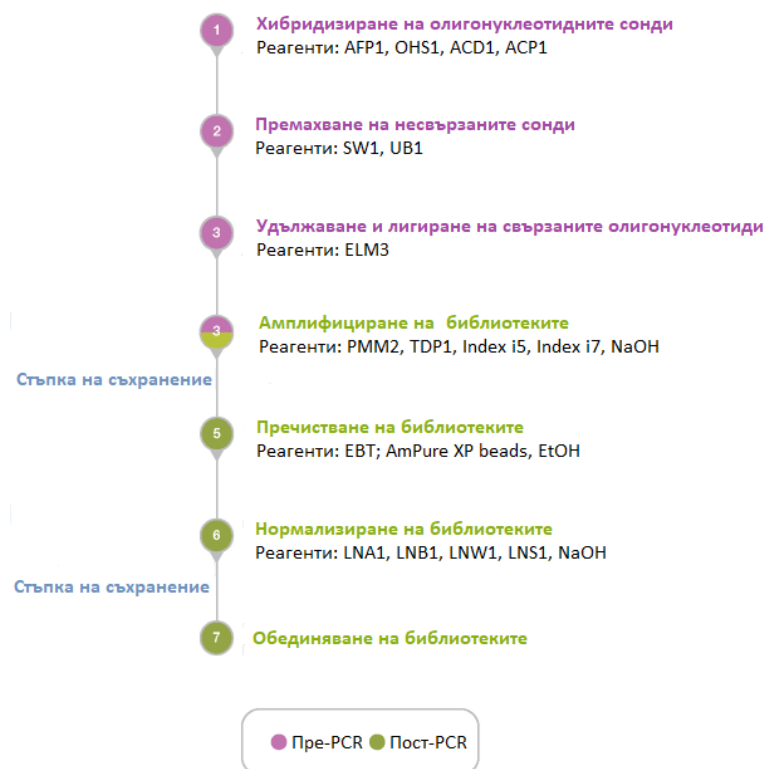
От избраните свежо замразени туморни проби от 12 АК и 12 ПК беше изолирана висококачествена геномна ДНК, посредством QIAamp DNA Mini Kit (Qiagen, Хилден, Германия). За подготовка на библиотеките е необходимо 250 ng стартово количество ДНК. Работни разредки бяха подготвени чрез спектрофотометър Nanodrop 2000 и флуорометър Qubit v2.0. След това се извърши qPCR-базиран ДНК качествен анализ. Така се оцени качеството на изолираните ДНК проби.

Протоколът на TSACP започва с хибридизация на предварително смесени, оптимизирани олигонуклеотидни проби „upstream“ и „downstream“ от таргетните региони. Всяка проба включва „capture“ последователност и адапторна последователност, които се използват в последващата реакция на намножаване. Реакция на удължаване-лигиране намножаване таргетния регион и лигира двете проби разположени „upstream“ и „downstream“ от таргетния регион. Това създава нова матрична верига и осигурява отлична специфичност. Матриците, получени при реакцията на удължаване-лигиране се амплифицират с PCR, като се инкорпорират две уникални специфични за всяка проба индекси. Крайният продукт съдържа готови за секвениране ампликони. Интегрираната нормализация с магнитни частици позволява прецизно обединяване на създадените библиотеки. Обединените библиотеки могат директно да се секвенират на MiSeq системата, без да е необходимо допълнително процесирание (Фигура.3.9.).



Фигура.3.9: (А) Хибридизация на олигонуклеотидните сонди, включени в раковия панел; (В) Екстензия и лигиране; (С) Добавяне на индекси и секвенционни адаптори посредством стъпка на PCR; (D) Крайния ампликон е готов за стъпка на секвениране посредством MiSeq платформата.

На фигура 3.10 е представен схематично протокола на работа с TruSeq Amplicon-Cancer Panel.



Фигура.3.10. Схематично представяне на протокола на работа с TruSeq Amplicon-Cancer Panel

3.5. Биоинформатичен и статистически подходи, използвани за анализ на получените резултати.

3.5.1. Биоинформатичен анализ на получените резултати от секвенирането от ново поколение с MiSeq платформата.

Получените сурови данни от таргетното секвениране с TSCAP панел (Illumina, Калифорния, САЩ) бяха анализирани със софтуерната програма VARSEQ (Golden Helix, Монтана, САЩ).

Софтуерът VARSEQ осигурява възможности за добро филтриране на данни и тяхното аотиране. С помощта на използваните филтри се стеснява списъка на варианти до тези, които представляват интерес. След вкарването на суровите данни в програмата, аотациите се прилагат автоматично въз основа на предварително конфигурирани настройки. Използват се 4 програми за предикция на новооткритите варианти - PolyPhen-2, SIFT, VEST3, PROVEAN.

3.5.2. Обработка на резултатите от микроРНК микрочиповия анализ.

След хибридизация микрочиповете бяха промити със специфични буфери и сканирани, съгласно инструкциите на фирмата производител. Данните от сканирането бяха обработени чрез Agilent Feature Extraction софтуер. От получените файлове може да се оцени качеството на анализирания резултат - дали белязването на пробите и хибридизацията на чиповете са протекли нормално и наличието на фонов шум. Ако фоновият шум е висок не могат надежно да се отчетат хибридизиралите проби. В този случай се налага повторно промиване.

След обработка с Feature Extraction софтуерът анализът на резултатите бе извършен с помощта на Gene Spring 12.6 GX-PA като бе избран 90th percentile алгоритъма за нормализация и без корекция на базовата линия.

3.5.2.1. Нормализация на данните.

Суровите данни, които се получават от микрочиповия анализ не могат да бъдат използвани директно. Необходимо е те да се нормализират и скалират. Gene spring софтуерът позволява избор между следните методи за нормализация: Scale, Percentile shift, Quantile, Normalize to control genes, Normalize to External value.

Благодарение на нормализирането и скалирането на суровите данни от всички микрочипове, се създава възможност за тяхното сравняване. Изборът на подходящ метод за нормализиране на данните от микрочиповете е много важен за отстраняване на варирането (177–179). В литературата са описани различни стратегии за анализ на микроРНК експресионни микрочипове с приложени различни методи за нормализация (180–183). От фирмата производител Agilent припоръчват суровите данни да се нормализират чрез 90th percentile (184,185). При него интензитетите на сигналите от хибридизационните проби се коригират като се взема всяка колона от спотове на чипа поотделно и се изчислява n -перцентилът на експресионните стойности на микрочипа спрямо всички сигнали (където n има стойности от 0 до 100, в случая 90). Деветдесетият перцентил е граничната стойност между горните 10% на наблюдаваните стойности на една променлива (в случая интензитета на светене на спота на микрочипа) и останалите 90%. След това се изважда получената стойност от експресионните сигнали за всяка микроРНК проба (185).

3.5.2.2. Анализ на основните компоненти – PCA (Principle Component Analysis).

Анализът на основните компоненти е математичен алгоритъм за предварителна обработка на данни. Позволява да трансформираме нормализираните данните и да редуцираме броя на измеренията на данните, като се запазва варирането. Редукцията се осъществява чрез идентифициране на нови променливи, наречени основни компоненти, при които варирането в данните е максимално и съществува някаква корелация между тях. Първичните компоненти са собствените вектори на матрицата на ковариациите на данните. Чрез използването на няколко компонента всяка проба може да бъде представена чрез относително малък брой числови стойности. PCA дава лесен поглед за данните и това подпомага тяхното групиране, тъй като пробите се визуализират и това позволява да се определят сходствата и различията между отделните проби (186).

В GeneSpring при PCA се калкулират стойности за три основни компонента и пробите могат да се нанесат визуално на тримерна декартова координатна система. Резултатите от този анализ могат да се използват за проверка на качеството на данните. На координатната система с една точка съответства на туморна тъкан, съответно от пациент с АК или ПК или нормална тъкан. Оцветяването отразява групирането на пробите. Така може да се види наличие на различия между отделните изследвани, както и представители на едната група, които са подобни на другата (и обратно). В най-добрия случай пробите от една група се групират (кълстерират) заедно и отделно от другите групи. Компонентите в PCA са представени по осите X, Y и Z и са номерирани 1, 2, 3... според тяхната намаляваща значимост.

3.5.2.3. Volcano plot.

Volcano plot е вид диаграма на разсейването, чрез която се визуализира взаимоотношението между степента на промяна в експресията на изследваните микроРНК и статистическата значимост. Диаграмата визуално представя диференциалната експресия между туморните и нормалните тъкани. По y-оста се представя отрицателен десетичен логаритъм на статистическата значимост ($-\log_{10}p$ value). МикроРНК-ите показали силна статистически значима разлика в експресията в туморните тъкани спрямо нормалните тъкани (с най-ниски p стойности) се появяват най-отгоре на графиката. На оста x се представя логаритъм от степента на промяна на експресията между двете сравнявани групи ($-\log_{10}FC$). В случая логаритмуваните стойности позволяват данните да се скалират и лесно да се визуализират. Точките на диаграмата (микроРНК-ите), които се най-отгоре на графиката и разположени далеч вляво или дясно от центъра имат най-голяма промяна в експресията и също така най-голяма статистическа значимост.

3.5.3. Количествен анализ на резултатите от генната експресия чрез real time-qPCR.

Количественият анализ се състои в измерване на количеството на РНК молекула, която е от интерес по време на циклите на амплификация. Това се осъществи с помощта на софтуер SDS v.2.2.2, който е предназначен за използване с 7900HT Fast Real Time PCR System (Applied Biosystems, Thermofisher Scientific, Масачузетс, САЩ). За оценка нивата на относителна или релативна експресия на желаната молекула

(микроРНКа или мРНК) се използва $2^{-\Delta\Delta Ct}$ метод. За всяка изследвана молекула, която представлява интерес (таргет) се изчислява:

$$\Delta Ct = Ct (\text{таргет}) - Ct (\text{ендогенна контрола})$$

Ct-стойността за ендогенна контрола се определя от софтуера за генът *RNU6* и *ACTB*. Избраните ендогенни контроли показват минимални вариации между отделните проби и в този случай са подходящи. Това се прилага с цел по-точна количествена оценка на РНК-а молекули, като така се компенсират разлики в тоталната РНК, добавена към реакцията. За всяка група се изчислява:

$$\Delta\Delta Ct = \Delta Ct (\text{експ.гр.}) - \Delta Ct (\text{контр.гр.})$$

Накрая за всеки таргет (маркер) се калкулира колко пъти е променена експресията (fold change; FC или Relative Quantification; RQ) на експерименталната в сравнение с контролната група или проба. Прилага се формулата:

$$RQ = 2^{-\Delta\Delta Ct}$$

Ако RQ е ≥ 2.000 тя се представя като повишена промяна в експресията на съответната изследвана микроРНКа или мРНКа. Ако RQ е < 0.500 се наблюдава понижена експресионна стойност. Стойности между 1.99 и 0.5 се приемат без промяна в експресионните нива.

Методът дава възможност за сравнение на няколко групи проби включени в дисертационния труд: туморни тъкани проби и съответващите им нормални тъкани проби.

3.5.4. Статистически методи.

Статистическият анализ на данните бе осъществен посредством R/R Studio или Python базирани библиотеки.

3.5.4.1. Статистическа проверка на хипотези.

Методите, позволяващи да се опише дали наблюдаваните разлики при различните изследвани маркери са значими, се наричат тестове на хипотези. Хипотезата е предположението, което правим за стойностите на изследваните маркери. Проверката на хипотезата преминава през следните основни етапи:

Дефиниране на нулевата хипотезата ($H_0: X_1 = X_2$), според която разликите между сравняваните стойности на маркерите е случайна и не е статистически значима.

Дефиниране на алтернативната (експериментална) хипотеза, тази която се противопоставя на нулевата хипотеза. Бележи се с H_1 : където $X_1 \neq X_2$ при сложна хипотеза и $X_1 < X_2$ или $X_1 > X_2$ при проста. Според нея разликата между сравняваните стойности на маркерите е статистически значима. Начина на дефиниране на алтернативната хипотеза, като сложна или проста, определя с каква критична зона се работи, съответно двустранна (2-tailed) или единична (1-tailed).

Първоначално се определя нивото на значимост – α (т.нар. Significance Level), представляваща вероятността да отхвърлим H_0 хипотезата, дори ако тя е вярна. Най-често използваното ниво на значимост е $\alpha = 0.05$ или това са 5% вероятност нулевата

хипотеза да се отхвърли, когато всъщност е вярна. След това се изчислява и вероятността (P) за получаване на наблюдаваните данни (или на данни с още по-голяма разлика между двете групи) при положение, че нулевата хипотеза е вярна и двете групи всъщност не се различават.

При $P > \alpha$ H_0 не се отхвърля, което означава, че между сравняваните параметри не съществува доказуема статистически значима разлика.

При $P < \alpha$ H_0 се отхвърля, което означава, че между сравняваните параметри съществува статистически значима разлика.

Нулевата хипотеза може да бъде проверена с различни статистически методи като изборът на тест за проверка на хипотезата зависи от:

- вида на разглежданата променлива – параметрични тестове при силни скали (при интервални или количествени променливи) и непараметрични тестове при слаби скали (при категорийни/качествени или ординални стойности на разглежданата променлива)
- според вида на разпределението на стойностите на променливата - параметричните методи имат изисквания към параметрите на разпределението. Параметричните тестове изискват нормално, равномерно, поасоново, експонциално разпределение, докато при непараметричните част от параметрите нямат значение. При тях се толерира определено отклонение в стойностите им.
- наличието на зависимост между извадките, от които се получават сравняваните параметри, т.е. дали стойностите на променливата получени от едната група (извадка) се влияят от тези на другата. Извадките са независими, ако се сравняват стойностите на признака X в две групи, които нямат връзка помежду си и не споделят външни условия, които биха повлияли на признака X в определена посока. Ако са налице фактори на средата, които влияят на променливата, влиянието им може да бъде неутрализирано чрез групиране на стойностите на променливата според стойността на променливата.
- броя на сравняваните параметри (с колко извадки се работи) - два или повече.

Тестовите, които бяха използвани за проверка на разпределението на данните бяха Kolmogorov-Smirnov и Shapiro-Wilk. Kolmogorov-Smirnov е непараметричен метод за проверка за съгласуваност между емпирично и теоретично разпределение (нормално, равномерно, поасоново, експонциално). Shapiro-Wilk тест е подходящ за проверка на разпределението при брой на проби по-малко от 50, а не е подходящ при проби с много еднакви стойности. Shapiro-Wilk е специфичен тест за нормалност, докато Kolmogorov-Smirnov е по общ и статистически по-слаб тест (което означава, че правилно отхвърля нулевата хипотеза за нормалност по-рядко). Kolmogorov-Smirnov е подходящ за проверка на разпределението на данни при по-големи извадки.

3.5.4.2. χ^2 -квадрат

χ^2 -квадрат методът е пригоден за изследване на връзки от корелационен тип. Чрез него може да се прецени, доколко предполагаемата зависимост (връзка)

съществува обективно, неслучайно и се проявява като такава сред изследвана съвкупност от единици. При него се сравняват две разпределения – емперично (действително) и теоретично и от съпоставянето на тези две разпределения се изгражда критерий, въз основа на който се проверява наличието на изследваната предполагаема връзка.

Чрез χ^2 -квадрат методът не може да се определи посоката на връзката, силата на взаимодействие между параметрите или дали тя е причинно-следствен или има друг характер. Методът е приложим само в случаите, когато общият брой на изследваните случаи е не по малко от 50 и когато теоретично очакваният брой случаи в отделните клетки на таблицата е не по-малък от 5.

3.5.4.3. Параметрични методи

3.5.4.3.1. Student's T тест анализ.

При Student's T тест се сравнява извадъчното средно за две отделни групи. Нулева хипотеза в случая е, че няма разлика между наблюдаваните средни стойности на двете сравнявани групи и всяко получено различие е случайно ($H_0: \mu = \mu_0$). T-тестът може да бъде двустранен ($H_1: \mu \neq \mu_0; t > t_{\alpha/2}$), посочвайки, че средните не са еквивалентни, или едностранен ($H_1: \mu > \mu_0; t > t_{\alpha}$), определящ дали наблюдаваната средна стойност е по-голяма или по-малка от предсказаната средна стойност. След това се изчислява тестовата статистика - t и от t-разпределението – p-стойността за конкретния тест. При стойности на $p < 0.05$ нулевата хипотеза се отхвърля. Този тест широко се използва за идентифицирането на диференциално експресирани микроРНКи и гени (187).

3.5.4.3.2. Едномерен еднофакторен дисперсионен анализ (ANOVA).

Дисперсионният анализ (ANOVA; ANalysis Of VAriance) е метод използван в статистиката за проверка на хипотези за равенство между повече от две извадъчни средни. Чрез тази проверка на хипотези може да се прецени доколко влиянието на даден фактор или група фактори е статистически значимо или не. По този начин се изследват връзки и зависимости като факторите обикновено са номинални, а значенията на резултативния признак имат числов израз. Чрез дисперсионния анализ може само да се установи дали между изследваните явления (признаци) съществува зависимост, но не може да се измери теснотата или силата на зависимостта, както и нейната посока (187).

3.5.4.4. Непараметрични методи

3.5.4.4.1. Тест на Ман-Уитни (Mann-Whitney).

Тестът на Ман-Уитни (Mann-Whitney U test) е подобен на Student's T теста. Ако при T-теста се сравняват средноаритметичните стойности за двете групи, то при теста на Ман-Уитни се сравняват техните медиани/рангове. За целта изходните стойности на числовата променлива се преобразуват в рангове и след това тестът проверява дали ранговете за двете независими групи са статистически значимо различни. Тъй като тестът на Ман-Уитни е непараметричен тест към него няма изискване за нормално разпределение на зависимата променлива и равенство на дисперсиите за двете изследвани групи. При него се проверява дали при наблюденията в едната извадка се наблюдава тенденция да имат по-големи стойности и съответно по-големи рангове в

сравнение с наблюденията в другата извадка. Нулевата хипотеза е, която се проверява е, че двете групи имат еднакво разпределение. Алтернативната хипотеза е, двете групи нямат еднакво разпределение. При този тест се цитират медианите, а не средноаритметичните стойности Mann-Whitney (187).

3.5.4.4.2. Тест на Крускал-Уолис (Kruskal-Wallis) за съпоставка на три или повече независими извадки

Тестът на Крускал-Уолис (Kruskal-Wallis) е непараметрична алтернатива на ANOVA за съпоставка на повече групи с ординални (рангови) данни. Разглежда се като разширение на теста на Ман-Уитни (Mann-Whitney U test) за три и повече независими извадки. Тестът на Крускал -Уолис, както и тестът на Ман-Уитни, използва ранговете на наблюденията като се сравняват средноаритметичните стойности на ранговете за всяка независима група. Тестът на Крускал-Уолис проверява за равенство на разпределенията между три и повече групи, а не разлика в средноаритметичните стойности за групите е както при ANOVA. При този тест се цитират медианите, а не средноаритметичните стойности.

Нулевата хипотеза, която се проверява е, че няма разлика в разпределението на независимите групите. Алтернативната хипотеза е, че се наблюдава различие в разпределението на групите. Ако тестът на Крускал -Уолис е статистически значим, $p \leq 0.05$, означава, че има различие в разпределението на групите, които са три и повече. За кои групи са статистически значимо различни, трябва да се проведат няколко допълнителни теста на Ман-Уитни, по един за всяка двойка от групите (187).

3.5.4.5. Оценка на процента на фалшиво позитивни резултати (асоциации) при множествени сравнения: тест на Benjamini–Hochberg FDR

Множественият сравнителен анализ (корекция при множествени изследвания; correction for multiple testing) е процедура, с помощта на която се коригира стойността на величината p (p – value) получена при провеждане на голям брой аналогични статистически тестове. Прилагането на тази процедура има за цел да намали вероятността от възникване на грешка от тип I (α грешка или получаване на фалшиво позитивен резултат), изразяваща се в погрешно отхвърляне на нулевата хипотеза. Benjamini–Hochberg (BH) корекцията е модифициран вариант на корекцията на Бонферони (Bonferroni correction) и за разлика от нея не е толкова строга. При BH метода първоначално p стойностите се сортират и ранкират (определят им се места в редицата). Най-малката p стойност се определя като номер 1, следващата е номер 2, а най-високата p стойност се определя като N (където N е общия брой на p стойностите). След това всяка p стойност се умножава по N и се дели на своя определен номер като така се получава коригирана p стойност.

3.5.4.6. Корелационен анализ.

Корелационният анализ е метод за изследване на зависимостта (корелацията) между дадени променливи. Прилага се за описание на силата и посоката на зависимост между променливи величини. При анализа се сравняват коефициентите на корелация между една или повече двойки променливи, за да се установят статистически взаимозависимости между тях. Силата (степената) на корелационните зависимости се измерва с помощта на коефициенти на корелация. Стойността им се движи в границите

от -1 до +1. В зависимост от измерителната скала, в която са изразени променливите, се ползват различни коефициенти на корелация.

Корелацията е индикатор за силата на връзката между две променливи. Важно е да се отбележи, че тази връзка не е задължително да бъде причинно-следствена, т.е. стойностите на едната променлива да зависят или да са следствие от стойностите на другата. Причинността може да бъде разкрита само в контекста на конкретните променливи и не се разглежда от статистиката.

3.5.4.6.1. Коефициент на линейна корелация на Пиърсън (Pearson; r)

Най-широко използваният индекс за линейна зависимост е коефициентът на корелация на Пиърсън, индекс отразяващ степента, в която стойностите на две променливи са линейно свързани (и потенциално зависими) една от друга. Прилага се, когато и двете променливи са количествени, а зависимостта им е близка до линейната. Ако този коефициент се използва при нелинейна връзка, ще се получи занижена стойност на връзката. За променливи, измерени в номинална или рангова скала корелационният коефициент се пресмята по други начини, които са производни на формулата за Пирсъновия корелационен коефициент. Корелационната скала е ординална, а не интервална.

Стандартният символ, с който се бележи коефициентът на корелация на Пиърсън е r . Абсолютната стойност на коефициентите на корелация е в границите от -1 до +1. Когато $r = 1$ или $r = -1$ има пълна линейна корелация между стойностите на двете променливи, а ако $r = 0$ няма линейна връзка между променливите. Ако коефициентът на корелация е положително число, зависимостта е положителна, възходяща – на по-големи стойности на едната променлива, съответстват по-големи стойности на другата променлива. Ако коефициентът на корелация е отрицателно число, зависимостта е отрицателна, низходяща – на по-големи стойности на едната променлива съответстват по-ниски стойности на другата променлива.

Има два начина за тълкуване на коефициента. Единият използва директно получената стойност:

- при стойност 0 – липсва корелация;
- до 0.3 – слаба корелация;
- от 0.3 до 0.5 – умерена корелация;
- от 0.5 до 0.7 – значителна корелация;
- от 0.7 до 0.9 – силна корелация;
- над 0.9 – много силна корелация;

Вторият начин за тълкуване на коефициента е чрез така нареченият коефициент на детерминация, r^2 . Стойността му се получава при повдигане на втора степен на коефициента r и отразява качеството на предиктивния модел, който бихме могли да построим върху данните. При линейната регресия този коефициент (отбелязван с R^2) отразява каква част от варирането в данните се отразява от модела (187).

3.5.4.6.2. Рангова корелация на Спирман (Spearman's rho).

Коефициентът на рангова корелация на Спирман се използва, когато данните не са линейни и има нужда от непараметричен метод за определяне на корелацията. Първоначалните числови или ординарни данни се подреждат по големина и се заменят с рангове и разликите между ранговете при всяко от наблюденията на двете

променливи, се пресмятат. Коефициентът на Спирман се означава с гръцката буква ρ („ро“, от там и наименованието „Spearman's rho“) или r_s .

Както корелационният коефициент на Пирсън r , така и Спирман ρ приема стойности от -1.00 до $+1.00$. Интерпретира по сходен начин и се използва, когато изследваната извадка е по-малка. Коефициентът на Спирман използва ранговете на наблюденията и по този начин минимизира влиянието на екстремалните стойности (аутлаери). Докато корелационният коефициент на Пирсън r е силно чувствителен към аутлаери.

За оценка на статистическата значимост на коефициентите на корелация нулевата хипотеза гласи, че зависимост между изучаваните параметри не съществува и полученият коефициент на корелация, може да се счита за случаен. В случай, че стойността на изчисления по данни от извадката коефициент на корелация е по-малък от критичната стойност ($p < 0.05$), за вярна се приема нулевата хипотеза, която твърди, че няма зависимост между променливите величини.

3.5.4.6.3. Посторояване на корелационно - базирани мрежи.

За да се анализират биологичните взаимоотношения между изследваните микроРНКи, бяха конструирани мрежи, базирани на коефициента на корелация (КБМ). Мрежовият анализ, базиран на корелация, наскоро се превърна в популярен метод за извличане на данни, тъй като позволява намаляване на сложността на многоизмерни данни, същевременно запазвайки по-голямата част от информацията, необходима за интерпретация (188). Разработени са редица свойства на корелационните мрежи, за да се позволи интерпретацията им:

- степен на възел: броят на други възли, към които даден възел е значително корелиран;
- централност: мярката за най-кратките пътища между всеки два възела, които минават през въпросния възел, и
- гъстота на мрежата: съотношението на съществуващите ръбове към общия брой възможни ръбове в мрежата.
- коефициентът на клъстериране: мярка за степента, до която възлите в мрежата са склонни да се групират заедно
- средният коефициент на групиране: използва се за оценка на характеристиката на плътното съседство на мрежата.

Използвайки тези свойства лесно могат да се открият силно свързани възли, известни като хъбове.

За изчисляване на корелационните коефициенти и p -стойностите беше избрана непараметричната рангова корелация на Спирман. Корелационните матрици бяха генерирани чрез SciPy - екосистема, базирана на Python (189). Единствено статистически значимите корелации (p -стойност $\leq 0,05$) се визуализират чрез корелационни топлинни карти с помощта на библиотеките Matplotlib и Seaborn в Python (190,191). За проектирането на КБМ бяха взети коефициент на корелация $\geq 0,5$ и коригирана от Benjamini-Hochberg FDR p -стойност $\leq 0,05$. Мрежите бяха визуализирани с помощта на NetworkX (192).

3.5.5. Прогнозиране на взаимодействията между микроРНК и целеви гени.

МикроРНКите регулират генната експресия, като подпомагат разграждането или потискат транслацията на целевите транскрипти. иРНК-мишени са каталогизирани в бази данни въз основа на експериментално валидиране и компютърно прогнозиране с помощта на различни алгоритми.

Генните-мишени на изследваните 14 микроРНК бяха предсказани с помощта на пакета "multiMiR" в софтуера R (193). Пакетът multiMiR, с уеб сървър на <http://multimir.org>, е изчерпателна колекция от предвидени и валидирани взаимодействия на микроРНКите и техните иРНК-мишени, техните асоциации с различни заболявания и лекарства. multiMiR включва няколко функции, които не са налични в съществуващите пакети на R. Компиляцията от 50 милиона записа от четиринадесет различни бази данни, разделени в три групи дава възможност за извличане на нужната информация, без да се посещават всички 14 бази данни. Първата група съдържа три бази данни за иРНК-мишени (TarBase, miRTarBase и miRecords), базирани на експериментално доказани взаимодействия. Във втората група са осем бази данни, при които взаимодействията микроРНК-мРНК са прогнозирани по различни алгоритми - TargetScan, DIANA-microT, miRanda, MicroCosm, PITA, EIMMO, miRDB и PicTar. Третата група обединява базите данни, свързани с експресията на микроРНКите и асоциацията им с различни заболявания и лекарства (PhenomiR, miR2Disease и Pharmaco-miR) (193). За целите на дисертационния труд бяха използвани данните за иРНК мишените от 14 бази данни.

3.5.6. Анотация на генната онтология (GO) и анализ на сигналните пътища.

Генната онтология (Gene Ontology - GO) цели да уеднакви описанията на генните продукти от различните бази данни и да ги подреди йерархично. GO описва биологичната функция на гените в три аспекта - асоциирани биологични процеси (БП), клетъчни компоненти (КК) и молекулни функции (МФ). Използва се за класификация на обогатен набор от гени-мишени на изследваните микроРНКи (194).

Чрез анализ на сигналните пътища е възможно да се търсят пътищата, в които участват гените, регулирани от изследваните микроРНК-ите при АК и ПК. По този начин може да се оцени, кои биомолекули са от решаващо значение за разбирането на изследваните явления. Чрез този анализ се откриват съответните групи от свързани гени и се идентифицират биологични пътища, които са обогатени в даден списък с гени повече, отколкото би се очаквало по случайност. За тази цел беше използвана WikiPathway базата данни (<https://www.wikipathways.org>) (195).

Онтологичният анализ, анализът на сигналните пътища и визуализацията на резултатите бяха осъществени с помощта на различни модули от пакета GSEArPy Python като коригирана р-стойност от $p < 0,05$ беше използвана, за да покаже значително обогатените GO теримни и засегнатите сигнални пътища (196).

3.5.7. Диаграма на Вен.

Диаграмата на Вен е графична форма, чрез която се представя извадковото пространство, елементарните изходи и събития. Тези диаграми използват кръгове, за да визуализират подобие/припокриването и разликите между данните. Математичното

определение е следното: диаграма, в която математически набор или набор от категорийни твърдения се представят чрез припокриващи се кръгове в рамките на поле, така че всички комбинации на съответните части са представени чрез разнообразни различни области в диаграмата (187). Диаграмата на Вен беше използвана за представяне на общите и специфично експресирани микроРНК при АК и ПК.

3.6. Машинно самообучение.

Машинното самообучение (Machine learning; ML) представляват група алгоритми за изграждането на математически модел въз основа на примери от данни, известни като „данни за обучение“ или параметри, за да се идентифицират модели и да се вземат решения (197). Основава се на законите на вероятността и статистика. Бива три основни вида – контролирано самообучение (Supervised learning), неконтролирано самообучение (Unsupervised learning) и подсилено обучение (Reinforcement learning).

За целите на настоящия дисертационен труд бяха тествани няколко алгоритъма за машинно самообучение с цел оценка на способността на изследваните микроРНК да разграничават двата подтипа НДБК – АК и ПК. Както и охарактеризиране на експресията на микроРНКите при всеки подтип. Беше използван Scikit-learn, Python пакет, който интегрира широк набор от съвременни алгоритми за машинно самообучение (198). Процесът на приложно машинно самообучение включва следните етапите:

- Дефиниране на проблема
- Анализ на данни
- Подготовка на данни - почистване на придобитите данни, трансформация на данните
- Оценка на алгоритъма - избор на алгоритъм за машинно самообучение; Конфигуриране на избрания алгоритъм; Трениране и валидиране на модела
- Подобряване на резултата
- Представяне на резултата

3.6.1. Подготовка на данните и почистване на данните.

Данните, които използвахме в различните модели на машинно самообучение са получените RQ стойности на изследваните 14 микроРНК при АК и ПК. Преди да бъдат обработени и анализирани, получените данните за експресията на микроРНКите, трябва да бъдат представени в подходящ за анализ вид. Събраните RQ стойности (параметри) за АК и ПК се обединяват в една обща таблица като параметрите са като променливи по колони. Всеки ред представлява отделен запис (отделен пациент).

След това е необходимо почистване на данните от липсващи стойности. В случая имаме две места, при които данните са със стойност данни с липсващи стойности (NaN). Когато има липсващи данни проведеният след това анализ може да бъде неправилен, а тестваните модели да бъдат неточни, защото повече алгоритми за прогнозиране не могат да работят с липсващи данни. Често използвани подходи за решаване на този проблем са липсващите данни да бъдат заместени със средни стойности, заменени с „0“, заменени с най-често срещани стойности или тези стойности да бъдат директно премахнати. Често експресията на микроРНКите е много ниска и не може да бъде детектирана от RT-qPCR и поради тази причина липсва запис.

Затова е избран методът, при което липсващите стойности се заменят с „0“, т.е. експресията е много ниска.

Поради хетерогеността при АК и ПК, решихме да запазим „екстремните стойности“ (outliers). В случая те могат да представляват интерес от биологична гледна точка. По-висока или по-ниска експресия може да съответства на по-висок туморен стадий, наличие на нодални или далечни метастази, да зависи от степента на диференциация или морфологичния подтип АК и ПК.

3.6.2. Нормализиране на данни.

Получените RQ стойности при АК и ПК имат различен диапазон и е необходимо те да бъдат нормализирани. Едни от най-широко използваните методи за нормализация са: MinMax, Zscore, StandardScaler, RobustScaler и нелинейна трансформация - квантилни и силови трансформации (199,200).

Избраният в този дисертационен труд тип за нормализиране на данните е силова трансформация. Силите трансформации са семейство параметрични, монотонни трансформации, които се прилагат, за да направят данните с Гаусово разпределение. Това е полезно при моделиране на данни, характеризирани с голяма дисперсия и се изисква нормалност. Понастоящем силите трансформация поддържат трансформацията на Бокс-Кокс и трансформацията на Йео-Джонсън. Оптималният параметър за стабилизиране на дисперсията и минимизиране на изкривяването се оценява чрез максимално правдоподобие. Бокс-Кокс изисква входните данни да са строго положителни, докато Йео-Джонсън поддържа както положителни, така и отрицателни данни. По подразбиране към трансформираните данни се прилага нормализация с нулева средна стойност и единична вариация (198).

3.6.3. Моделиране.

След като събраните данни са подготвени, изчистени и нормализирани те се подават на моделиращата функция и се извършва обучението на алгоритъма. Моделът представлява математически израз на параметрите на моделиращата функция, заедно с входните характеристики за всяка прогноза, клас и действие за регресионни, класификационни и подсилващи категории (201). В машинното самообучение има разделяне на алгоритмите на 3 основни типа категории като изборът на подходящ алгоритъм за решаване на даден проблем зависи основно от вида, размера на данните и целта, която искаме да постигнем (202). Основните категории машинно обучение са:

- Контролирано самообучение (Supervised Learning) – се използва за напредна прогноза на базата на примери. Алгоритъм за контролирано самообучение търси закономерности в етикетите на стойността. Контролирано самообучение бива Класификация (Classification) – за прогнозиране на категориални данни; Регресия (Regression) – за прогнозиране на скаларни данни. Тук спадат линейна регресия (Linear Regression), Naive Bayes класификатор, логистична регресия (Logistic Regression; LR), Метод на опорните вектори (Support Vector Machine; SVM), дърво на решения (Decision Tree; DT), случайна гора (Random Forest; RF), AdaBoost и др. алгоритми.
- Неконтролирано самообучение (Unsupervised Learning) - данните нямат свързани с тях етикети. Целта на този тип алгоритми е да се организират

данните или да се опише тяхната структура. Към обучението без учител спадат: Клъстерен анализ (Clustering) – групиране на набор от обекти; Асоциативен анализ (Association Analysis) – метод базиран на правила за откриване на връзки между променливи и данни; Намаляване на размерността (Dimensionality Reduction). Към обучението без учител се числят алгоритмите – К - най-близките съседи (K-Means Clustering), йерархично клъстериране (Hierarchical Clustering), DBSCAN Clustering и др. алгоритми.

- Подсилено обучение (Reinforcement Learning) – комбинация от предходните две. При него алгоритъмът взема решение и старите решение ефектират бъдещето. Програма/алгоритъм, който променя стратегия си и учени в движение. Тук спадат дълбоките невронни мрежи, Метод на Монте Карло, SARSA и др. алгоритми.

3.6.4. Избрани ML алгоритми за класификационен анализ.

След извършен сравнителен анализ в този дисертационен труд на типовете алгоритми са избрани 4 класификационни алгоритъма за работа - дърво на решения, логистична регресия, случайна гора и метод на опорните вектори. Тези модели са избрани спрямо целите на проучването и данните се използват за прогнозиране на категория – дали туморите се класифицират като АК (клас 1) или ПК (клас 2). Съща така бе проведено йерархично клъстериране и метод на К - най-близките съседи при АК и ПК.

3.6.5. Анализ на методи за валидиране.

При тренирането на даден модел за машинно самообучение и за оценка на неговото представяне при нови за него данни, разделихме всички данни на две части – данни за трениране на дадения алгоритъм и данни за тестване на модела. Функцията Train-Test Split от scikit learn разделя данните на две части в зададено процентно съотношение. В случая данни се разделят 70:30 като по-голямата част се използва за трениране, а по-малката за тестване.

Понякога моделите се справят твърде добре върху данните за обучение, нагаждайки се към тях. Това се получава поради дисперсията в данните за трениране и се отразява негативно на модела. Всеки модел представлява сложен математически алгоритъм, включващ в себе си допълнителни специфични настройки (хиперпараметри). В случая, за да избегнем прекомерното нагаждане (overfitting), се налага тези хиперпараметри да се оптимизират. Най-често прилаганите подходи за справяне с прекомерното нагаждане на модел към данни са:

- Кръстосано валидиране (Cross-validation); данните за трениране се разделят на случаен принцип в k на брой подмножества с еднакъв размер. От тях се запазва една част като данни за валидиране и тестване на хиперпараметрите, а останалите част се използват като данни за обучение. Процесът се повтаря 5 пъти в случая. Предимството на този метод е, че всички наблюдения се използват за обучение, а за валидиране точно веднъж. Целта на кръстосаното валидиране, избягвайки от прекомерно нагаждане да провери способността на

модела да предсказва данни, които не са били използвани при оценката му (203). При задачи за класификация с малък набор от данни за оптимизиране на хиперпараметрите на моделите е най-подходящо използването на кръстосано валидиране (204).

- Обучение с повече данни;
- Регуларизация - добавяне на допълнителни ограничения, с цел да се постигне баланс на това, колко далеч от целта сме (bias) и в същото време колко близо сме до целта (variance). Чрез регуларизация се извършва контрол на сложността на модела. По този начин вкарваме предположения и алгоритъмът ни не се съобразява твърде много с данните. Съществуват 3 начина за минимизиране на функциите на загуба - L1 (Least Absolute Shrinkage and Selection Operator; LASSO) минимизира сумата от абсолютните разлики между прогнозни и действителни стойности, L2 (Ridge Regularization) минимизира сумата от квадратните разлики между прогнозни и действителни стойности и Elastic Net (комбинация от L1 и L2)(205–207).

3.6.6. Градиентно спускане.

Градиентното спускане е алгоритъм за оптимизация, при които се цели минимизиране на някои функции и намиране на локалния минимум чрез движение по градиента. В машинното самообучение се използва градиентно спускане, за да се подобрят хиперпараметрите на използвания модел (208).

3.6.7. Метрики за оценка на класификационни модели.

Съществуват различни метрики за оценка на работата на класификационния модел върху набор от тестови данни (209).

3.6.7.1. Матрица за оценка на класификацията (Confusion Matrix).

За да се оцени резултата на даден модел върху набор от тестови данни, за които са известни истинските стойности, се прилага матрица за оценка на класификацията (Фигура. 3.11). Чрез нея визуално се представя специфичността (точността), чувствителността, прецизността, и цялостното представяне на модела (210). Матрицата представя честотите и видовете на грешките и така позволява да се види, къде приложеният модел прави грешни предположения. Тя да възможност за изчисляване на основните метрики за оценка на модела.

		РЕАЛНОСТ	
		Позитивни	Негативни
ПРЕДСТАВАНИ	Негативни	Истински положителни резултати (TP) Предсказани правилно като клас 1	Фалшиви положителни резултати (FP) Предсказани грешно от клас 2
	Позитивни	Фалшиви негативни резултати (FN) Предсказани грешно от клас 1	Истински отрицателни резултати (TN) Предсказани правилно като клас 2

Фигура 3.11 Матрица за оценка на класификацията

Матрицата за оценка на класификацията съдържа четири квадранта:

- Истински положителни резултати (TP) – обекти, класифицирани като положителни и в действителност са в положителния клас
- Истински отрицателни (TN) – обекти, класифицирани като отрицателни и в действителност са в отрицателния клас
- Фалшиви положителни резултати (FP) - обекти, класифицирани като положителни, но в действителност са в отрицателния клас или Type I Error
- Фалшиви негативни резултати (FN) - обекти, класифицирани като отрицателни, но в действителност са в положителния клас или Type II Error

3.6.7.2. Основни метрики за ефективност.

Точността (accuracy) е главният показател за оценка на различните модели за класификация. Тя представя общия процент на правилните прогнози получени от модела. Изчислява се по формулата:

$$\text{Точност} = \frac{TP + TN}{TP + TN + FP + FN}$$

където TP - истински положителни, TN - истински отрицателни, FP - фалшиви положителни и FN - фалшиви негативни резултати.

Точността дава обща оценка за представянето на конкретния алгоритъм, но не дава информация поотделно за всеки от двата класа. Поради тази причина в процеса по оценяване се взимат предвид и останалите метрики – прецизност, чувствителност и F1-мярката, положителна прогностична стойност (PPV) и отрицателна прогностична стойност (NPV).

Прецизност (precision) е съотношението на правилно прогнозираните положителни случаи към общите прогнозни положителни случаи. Прецизността показва вероятността до колко пациентите, класифицирани като част от даден клас (АК или ПК), наистина принадлежат на него (честотата на истински негативните резултати, представя се в %). Високата прецизност е свързана с ниската степен на фалшиво предсказани положителни резултати. Изчислява се по формулата:

$$\text{Прецизност} = \frac{TP}{TP + FP}$$

Чувствителност (recall) - това е съотношението на правилно прогнозираните положителни случаи към всички наблюдения за даден клас. Чувствителността показва възможността колко от пациентите в даден клас са били класифицирани правилно (честотата на истинските позитивни резултати, представя се в %):

$$\text{Чувствителност} = \frac{TP}{TP + FN}$$

Двете метрики са така дефинирани, че когато стойността на едната е висока, то другата намалява и обратното. В зависимост от класификационната задача е възможно да избегнем допускането на единия тип грешка, докато другият тип грешка не е от особена важност.

Средна хармонична стойност (F1 Score) е средно претеглената стойност между специфичност и чувствителност. Вземат се под внимание фалшивите положителни резултати и фалшиво негативните резултати. Метриката е подходяща, когато искаме да вземем и двете статистики, измерващи грешката на модела.

$$F1\ score = \frac{2 * \text{Прецизност} * \text{Чувствителност}}{\text{Прецизност} + \text{Чувствителност}}$$

Освен гореизброените метриците за оценка на даден класификационен модел за машинно самообучение, съществува и начин получените резултати след трениране да се визуализират.

ROC кривите (Receiver Operating Characteristic Curve) представляват графични диаграми, илюстриращи изпълнението на бинарна класификация система при вариране на дискриминираща гранична стойност. Графиката на ROC кривите показва ефективността на модел за класификация (211). Тя очертава два параметъра: Истински положителен размер (True Positive Rate; TPR) и Фалшив положителен размер (False Positive Rate; FPR).

Ако имаме случайни предположения верните положителни предположения ще имат равен брой с фалшивите положителни предположения ($TPR = FPR$).

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{TN + FP}$$

Площта под кривата - AUC (area under the curve) е ефективна и комбинирана стойност на чувствителността и специфичността, обобщена мярка за сигурност във всички възможни прагове за класификация (212). Варира в стойност от 0 до 1.

3.6.8. Йерархичен клъстерен анализ.

Този метод се използва за визуализация на взаимовръзките между нивата на експресия на подбрани микроРНКи и серия от състояния. Йерархичното клъстериране е метод, често използван при създаването на филогенетични дървета. Принципът на йерархичните алгоритми е изграждането на клъстерите в резултат на многократно присъединяване на двата "най-близки" гени или групи от гени (или други разглеждани обекти). Този подход се използва с цел да се сведе до минимум променливостта в рамките на групата. Връзката между микроРНКите се представя графично като дърво (дендрограма), чиито дължини на клоните отразяват степента на сходство между гените. Тези дендограми са полезни в способността им да представляват различни степени на сходство и по-далечни отношения между групи от тясно свързани микроРНКи, а също така изискват да се направят сравнително малък брой предположения за характера на данните. Изчислените дърветата могат да бъдат използвани за подредба на гените в таблица с данни, така че микроРНКи или групи от микроРНКи с подобни модели на експресия са разположени близко един до друг. С помощта на този метод се дава възможност за един и същи вид интуитивен подход в целия анализ на експресионни данни. Въпреки че този подход е един от основните, който няма присъща специфичност към конкретния метод, използван при усвояването на геномните данни, той се е доказал в практиката като изключително полезен.

Йерархичните алгоритми винаги дават едно конкретно дърво, дори ако то не отразява истинската структура на реалната популация, от която произлизат данните. Макар дендрограмите да се считат за прости инструменти за визуализация, трябва да се подчертае, че това е много специално представяне на данните, което може да бъде напълно неуместно.

При йерархичното клъстериране се използват няколко варианта за изчисляване на разстоянията между обектите или групите:

- Ward's rule - правило на Ward

При него разстоянието между клъстерите представлява сума от квадратите на отклоненията от точките до центроидите. Целта на Ward е да минимизира сумата на квадратните разлики във всички клъстери и да минимизира дисперсията. Чрез този метод се получават клъстери със сходен брой наблюдения, но той е много чувствителен към силно отклоняващи се проби (outliers). Възможно е разстоянието между два клъстера да бъде по-голямо от d_{\max} - максималната стойност в оригиналната матрица. Ако това се наблюдава, то сходството ще бъде негативно.

- Maximum или complete linkage

Представлява клъстериране, при което на всяка стъпка разстоянията между групите се определят от техните най-далечни членове. При този вид йерархично групиране използваните критерии не са локални като се минимизира максималното разстояние между наблюденията на двойките клъстери. Това води до дефиниране на компактни клъстери с малки диаметри (разстояние между членовете) в големи, разнородни групи, като и увеличена чувствителност към големи различия в стойностите.

- Average linkage rule

Разстоянието между клъстерите е средното разстояние между наблюденията в един клъстер (група) и наблюденията в друг клъстер (група), тоест разстоянията между групите се определят от усреднените разстояния между всеки два техни члена. То представлява компромис между чувствителността на complete linkage клъстерирането на показатели с големи различия в стойностите и тенденцията на single linkage клъстериране да образува дълги клонове в дендрограмите, което не съответстват на идеята за целта на групирането (тоест окрупняване и компактно представяне на данните).

- Single linkage rule

Разстоянието между два клъстера е минималното разстояние между наблюденията в единия клъстер и наблюденията в другия клъстер, тоест разстоянията между групите се определят от техните най-близки членове. При това групиране сходството на двете групи е сходството на техните най-подобни членове. Критериите за сливане са локални. Обръща се внимание единствено на зоната, където двата клъстера са най-близки един до друг. Други, по-отдалечени части на клъстера и цялостната структура на клъстера не се взимат под внимание. Този метод е подходящ, когато групите (клъстерите) са очевидно разделени. Когато стойностите (наблюденията) лежат много близко едни до други, то чрез Single метода се идентифицират дълги вериги подобни клъстери, които могат да имат относително големи разстояния, разделящи наблюденията в двата края на веригата и да не дават реална представа за структурата на популацията от обекти.

Lance и Williams разработват формула, която обяснява методите за клъстериране (213). В текущото проучване бе използвано правилото на Ward при йерархичния клъстерен анализ.

Друга важна характеристика при клъстерирането е метриката, която повлиява формата на клъстерите. Някои елементи могат да се разполагат близко един до друг или далечни в зависимост от използваната метрика. Например в двумерното пространство разстоянието между точката (1,0) и началото (0,0) е винаги 1, според usual norms. Разстоянието между точката (1,1) и началото (0,0) може да бъде $\sqrt{2}$ при използване на Manhattan метричната система, ако се използва Евклидовото разстояние или 1 при максималната метрична система. В текущото проучване беше използвано Евклидовото разстояние при клъстерния анализ.

3.6.9. Метод на k-най-близките съседи.

Методът на k-най-близките (KNN) е непараметричен класификатор, при който групирането се осъществява на три стъпки. На първата стъпка се избират началните центроиди. След инициализацията алгоритъмът циклично преминава между другите две стъпки. При първата стъпка всяка извадка се разпределя към най-близкия до нея центроид. Втората стъпка създава нови центроиди, като взема средната стойност на всички проби, разпределени към всеки предишен центроид. Изчислява се разликата между старите и новите центроиди и алгоритъмът повтаря последните две стъпки, докато тази стойност е по-малка от праговата. С други думи, алгоритъмът се повтаря, докато центроидите не се преместят значително.

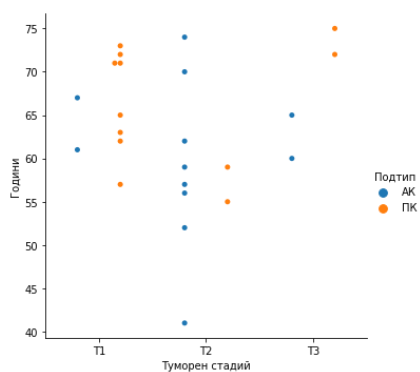
ЧАСТ 4

РЕЗУЛТАТИ

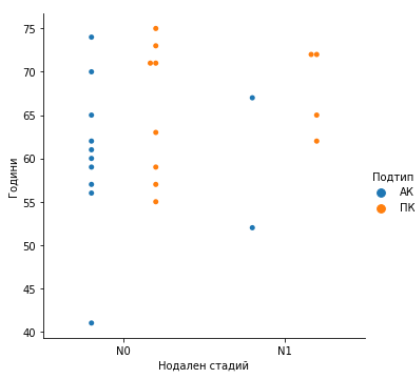
4.1. РЕЗУЛТАТИ ОТ микроРНК МИКРОЧИПОВИЯ АНАЛИЗ.

Средната възраст на пациентите с АК е 60 години като най-младия участник с АК е на 39 години, а най-възрастния участник е на 76 години. При пациентите с ПК средната възраст е 62 години, най-младия пациент е на 40 години, а най-възрастния на 75 години.

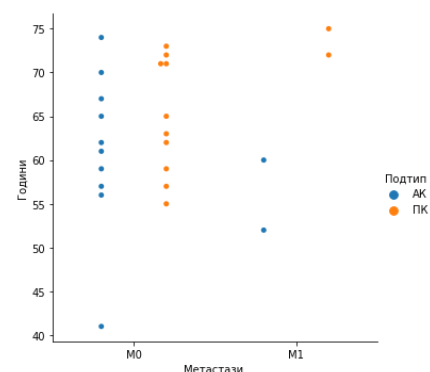
При АК с най-голям относителен дял са пациентите в декадата от 50 и 59 години (39%), следвани от пациентите на възраст между 60 до 69 години (32%). Над 70 години са 19% и между 40-49 години са 10%. При ПК туморите, преобладават пациентите над 70 години - 32%. Между 50 - 59 години и 60 - 69 години са 29% и в декадата от 40-49 години – 10% (Фигура 4.1.- 4.3).



Фигура 4.1. Разпределение на пациентите по възрастови групи и туморен стадий



Фигура 4.2. Разпределение на пациентите по възрастови групи и нодален стадий.



Фигура 4.3. Разпределение на пациентите по възрастови групи и наличие/отсъствие на метастази.

От включените в изследването АК и ПК разпределението по пол и възраст беше както следва – от АК 7 (58%) принадлежаха на мъже на средна възраст 61 г. и 4 (42%) на жени на средна възраст 62 (Таблица 3.1). При ПК 8 от туморите принадлежаха на мъже (67%) на средна възраст 68 г. и съответно на жени (33%) на средна възраст 61 г. (Таблица 3.1). Като цяло 15 (63%) от НДКБК принадлежаха на мъже на средна възраст 64 г. и 9 (38%) на жени на средна възраст 61 г.

При АК преобладаваха туморите в стадий Т2 – 8 тумора, следвани от туморите в стадий Т1 и Т3 с по 2 тумора. При ПК най-много бяха туморите в стадий Т1 – 8 тумора и с 2 тумора в стадий Т2 и Т3.

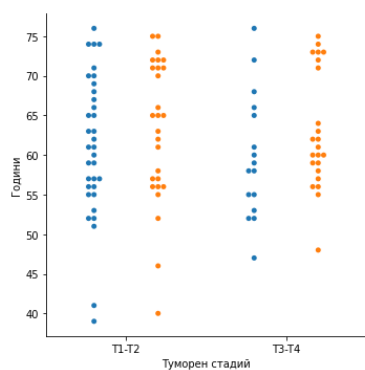
Лимфни метастази бяха наблюдавани при общо 6 (21%) от НДКБК, от които преобладаваха тези при ПК (4 тумора) и АК (2 тумора). Наличие на далечни метастази бяха наблюдавани при двама пациента, диагностицирани с АК и двама с ПК.

По отношение на диференциацията при АК 2 тумора са добре диференцирани (G1), 5 тумора са умерено диференцирани (G2) и 5 тумора са слабо диференцирани (G3). При ПК 3 тумора са G1, 2 тумора са умерено диференцирани (G2) и 6 тумора са G3.

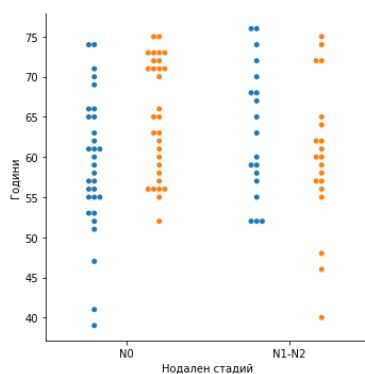
По отношение на морфологията при АК тумора 3 тумора са смесени с преобладаваща солидна компонента и 9 тумора са ацинарно-папиларни (с различно преобладаване на ацинарната и папиларната компонента). Като при 4 тумора от тази

група се наблюдава обилно слузообразуване. При ПК 4 тумора са без кератинизация и 8 тумора с кератинизация.

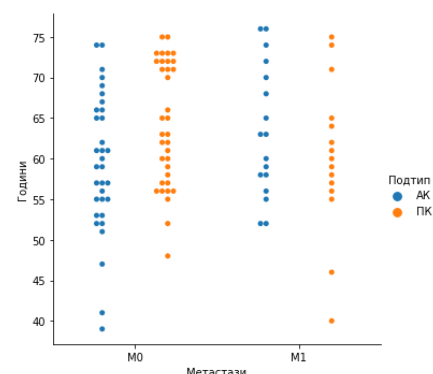
В групата на туморите, при които беше извършена валидацията на експресията на избрани микроРНК, чрез RT-qPCR преобладаваха мъжете – 33 при АК и 42 при ПК. Бяха включени 17 жени с АК и 8 с ПК. Разпределението на пациентите по възрастови групи и туморен стадий, нодален стадий и наличие/отсъствие на метастази е представено на фигури Фигура 4.4 - Фигура 4.6. Преобладаваха пациенти, диагностицирани в ранен туморен стадий (T1 –T2), без лимфни и далечни метастази.



Фигура 4.4. Разпределение на пациентите по възрастови групи и туморен стадий



Фигура 4.5. Разпределение на пациентите по възрастови групи и нодален стадий.

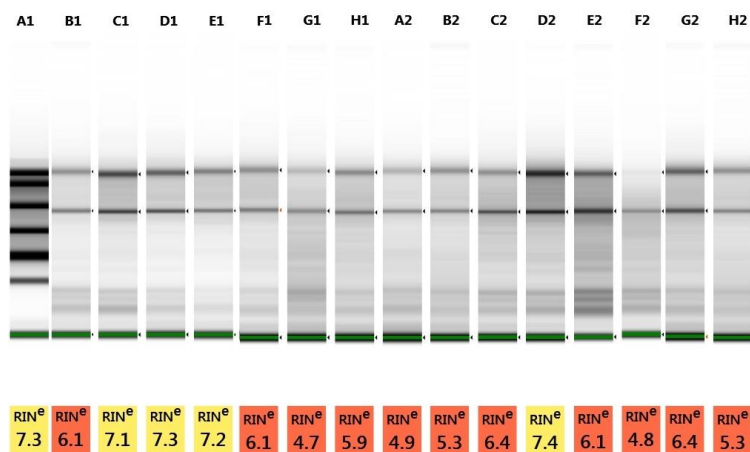


Фигура 4.6. Разпределение на пациентите по възрастови групи и наличие/отсъствие на метастази.

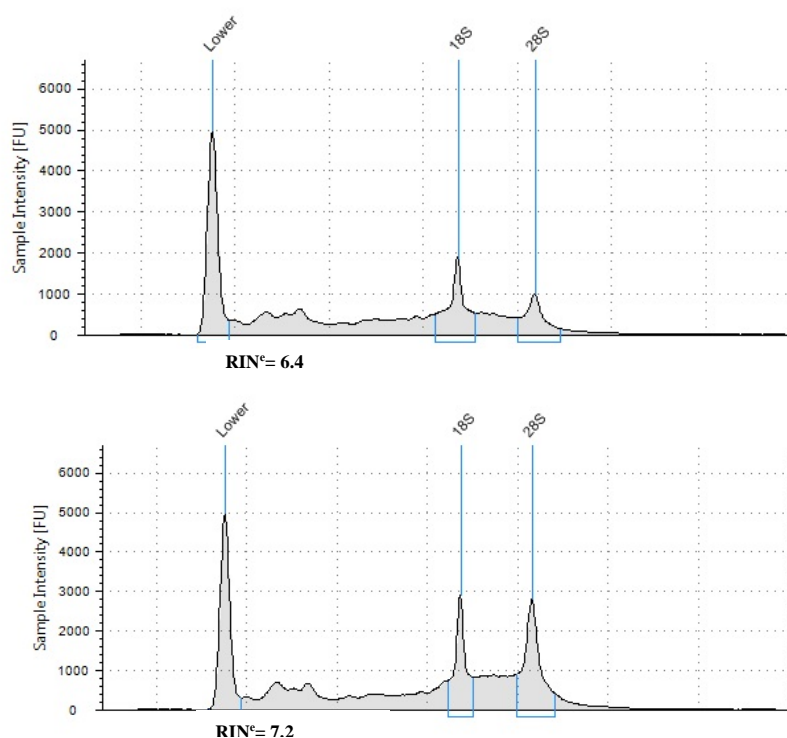
4.1.1. Измерване на концентрацията и определянето на качеството на изолираните РНК с помощта на Nano Drop и Agilent 2200 TapeStation System. Изготвяне на 50 ng/μl разредки за miRNA експресионното профилиране чрез микрочипов анализ.

Концентрациите на изолираните РНК-и от туморните и нормалните белодробни тъкани бяха първоначално определени на NanoDrop, за да се провери тяхното качество, след това бяха измерени на TapeStation, според протокола на производителя. Концентрациите на РНК, изолирани от туморните белодробни тъкани и измерени на NanoDrop, варират от 300 до 3700 ng/μl, а на нормалните тъкани от 70 до 500 ng/μl. Спрямо определените концентрации на NanoDrop бяха приготвени 50 ng/μl разредки от пробите и тези разредки бяха измерени на TapeStation. Стойностите за RINe (RNA integrity number), показващ РНК интегритета, качеството на изолираната РНК, варираше от 3 до 8. За провеждането на микроРНК експресионното профилиране подбрани бяха проби с RINe над 6.

Фигури 4.7, 4.8 и представят резултатите получени след анализ със системата TapeStation.



Фигура 4.7. TapeStation снимка на гел с профилите на разделяне на всяка проба и установените стойности на RIN^e. A1 – молекулен маркер; B1-H1 – АК тумори; A2-H2 – ПК тумори



Фигура 4.8. Електрофореграми на тотална РНК от проби с различни RIN^e.

Бяха пуснати шест чипа – три с тумори от плоскоклетъчен белодробен карцином и прилежащите нормални тъкани и три с тумори от аденокарцином на белия дроб и прилежащите нормални тъкани. След 20 часа хибризация на 55 °C микрочиповете бяха промити и сканирани, съгласно инструкциите на фирмата производител. Данните от сканирането бяха обработени чрез Agilent Feature Extraction. Бяха отчетени стойностите на параметрите, които показват качеството на анализиранияте проби. Данните са представени на Таблица 4.1А, Фигура 4.9А. Поради наличието на силен фонен шум осъществихме оптимизирахме – промиване на чипа с ацетонитрил. След сканиране наблюдавахме подобрене на резултатите Таблица 4.1Б, Фигура 4.9Б. По този начин оптимизирахме протокола и добавихме допълнителна стъпка – промиване с ацетонитрил след промиването с работните буфери.

Таблица 4.1 А) Стойности на параметрите, които показват качеството на анализираниите проби след промиване с работните буфери. Б) Стойности на параметрите, които показват качеството на анализираниите проби след промиване ацетонитрил.

А

Evaluation Metrics for miRNA_QCMT_Jun14
Excellent (3) ; Good (2) ; Evaluate (1)

Metric Name	Value	Excellent	Good	Evaluate
IsGoodGrid	1.00	>1	NA	<1
AddErrorEstimateGreen	20.78	<5	5 to 12	>12
AnyColorPrctFeatPopn...	2.31	<8	8 to 15	>15
gNonCtrlMedPrctCVBGS..	5.78	0 to 10	10 to 15	<0 or >15
gTotalSignal75pctile	0.10			
LabelingSpike-InSignal	4.39		>2.50	<2.50
HybSpike-InSignal	4.22		>2.50	<2.50
StringencySpike-InRat...	0.95			

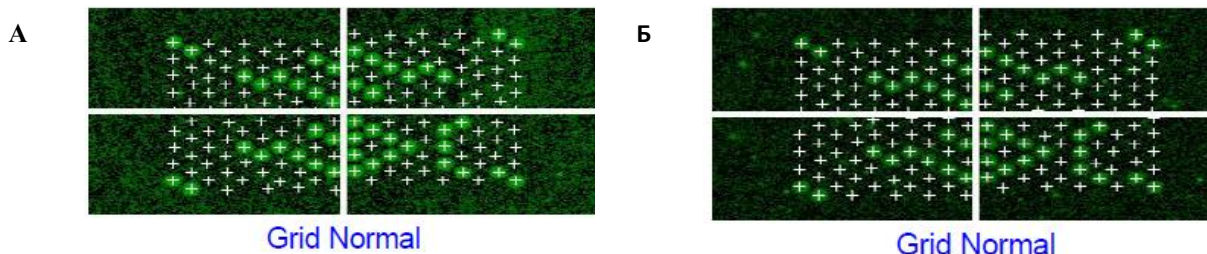
• Excellent • Good • Evaluate

Б

Evaluation Metrics for miRNA_QCMT_Jun14
Excellent (4) ; Good (2)

Metric Name	Value	Excellent	Good	Evaluate
IsGoodGrid	1.00	>1	NA	<1
AddErrorEstimateGreen	4.33	<5	5 to 12	>12
AnyColorPrctFeatPopn...	3.26	<8	8 to 15	>15
gNonCtrlMedPrctCVBGS..	6.96	0 to 10	10 to 15	<0 or >15
gTotalSignal75pctile	0.10			
LabelingSpike-InSignal	4.36		>2.50	<2.50
HybSpike-InSignal	4.23		>2.50	<2.50
StringencySpike-InRat...	0.99			

• Excellent • Good • Evaluate



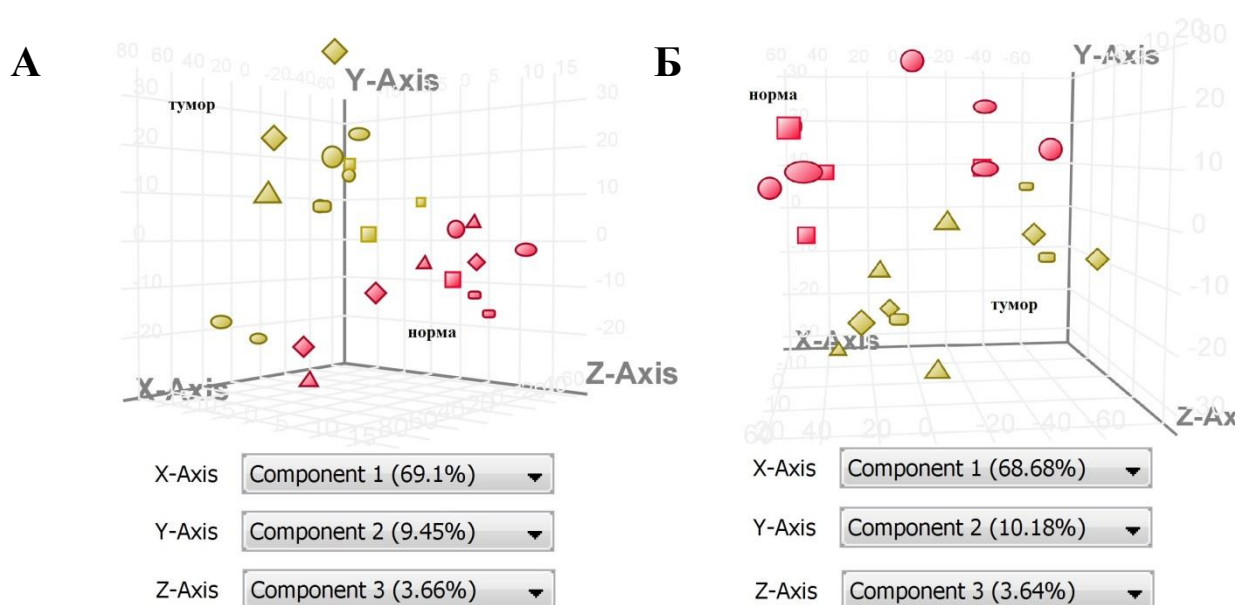
Фигура 4.9. А) Откриване на ъгълите на поле от чипа и разпознаване на грида. Наличие на фонен шум. Б) Откриване на ъгълите на поле от чипа и разпознаване на грида. Промиване с ацетонитрил и нормален фонен шум.

4.1.2. Подбор на метод за нормализация на резултатите от микрочиповия анализ.

Следващата стъпка, след сканиране на чиповете, е нормализация на суровите данни. Gene spring софтуерът позволява избор между следните методи за нормализация: Quantile, Percentile shift, Scale, Normalize to control genes, Normalize to External value. Изборът на подходящ метод за нормализиране на суровите данните от микрочиповете е важна стъпка за отстраняване на варирането и за създаване на възможност за тяхното сравняване (177,179,214). От фирмата производител на микрочиповете препоръчват нормализация на данните от miRNA (микроРНК) микрочиповете да е 90th percentile shift. Този метод за нормализация е използван в редица проучвания (215,216) и е описан в раздел 3. Материали и Методи.

4.1.3. Контрол на качеството на резултатите и PCA (principle component analysis - анализ на основните компоненти).

Преди анализирането на резултатите бе извършена проверка на качеството на отделните полета с проби от микрочиповете. Нормализираните данни на всички микроРНК се използват за генерирането на PCA 3D плот, чрез които се визуализират глобалните микроРНК профили на туморните и нормалните тъкани. В резултат на този анализ от групата на ПК карциноми беше изключен един тумор и една нормална тъкан, а от групата на АК беше изключена една нормална тъкан, тъй като не бяха с добри параметри и това не позволяваше добро разделяне. На Фигура 4.10 са показани резултатите от PCA анализът.



Фигура 4.10. PCA, показващ разграничаването на туморите и нормалните тъкани на две групи въз основа на miRNA експресионните профили. Легенда: в червено са отбелязани нормите, в жълто туморите. А) аденокарцином Б) плоскоклетъчен белодробен карцином

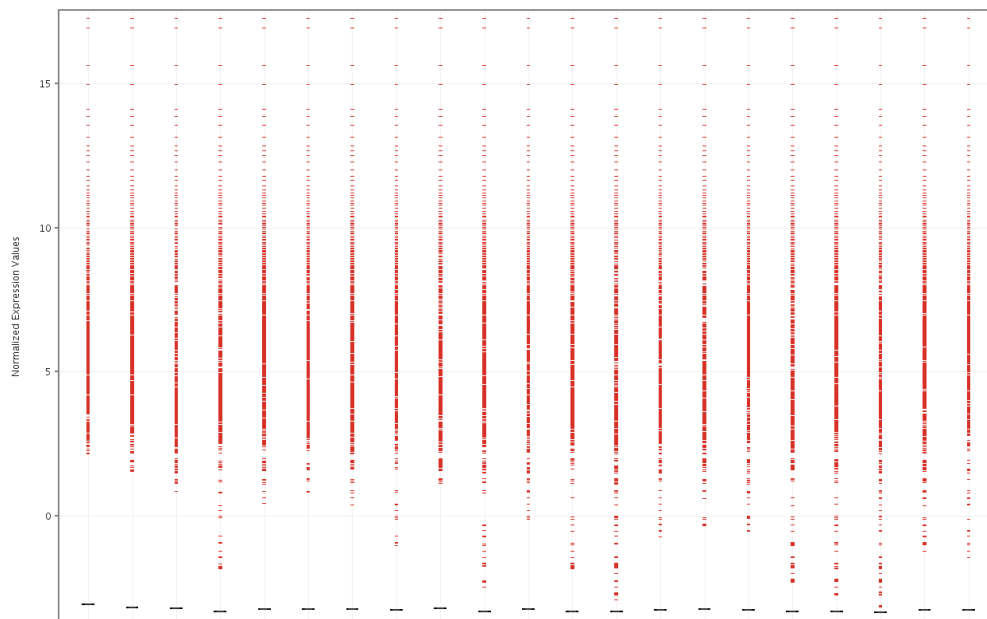
4.1.4. Филтриране на получените резултати.

Първоначално филтрирахме микроРНК-ите според това дали са детектирани или не - Filter by Flags. В платформата на Agilent микрочиповете всяка една микроРНК е представена с няколко копия различни олигонуклеотидни хибридационни проби. Ако всички повторения на дадена микроРНК са без отчетен сигнал, то приеманата стойност за тази микроРНК е 0, а ако дадена микроРНК е с отчетен сигнал във всички свои повторения се приема 1. Тези стойности се отбелязват в колонката gIsGeneDetected. Когато има сигнал само от едно от повторенията на дадена микроРНК, то тази тя се изключва от анализа тъй като резултата не е надежден. След това филтриране микроРНК-и не отпадат.

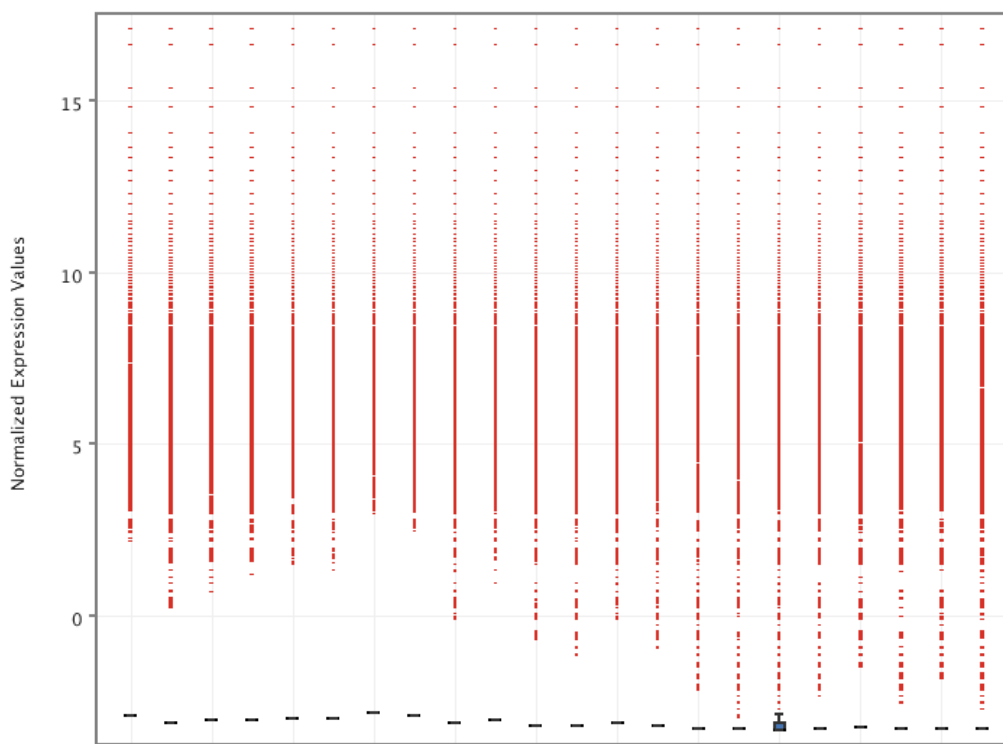
Осъществихме и филтриране според експресията (Filter Probesets by Expression) върху суровите данни, което ни позволи да се изключат микроРНК-и с много нисък

интензитети или тези, които са достигнали максимална сатурация. Предпоследното филтриране, което приложихме е филтриране за грешки (Filter Probesets by Error). Така се изключват тези микроРНК-и, при които коефициента на вариране (CV) е над 50%. Така от останаха при аденокарциномите – 2196 микроРНКи, а при плоскоклетъчните - 2346 микроРНКи.

Всички проби в Box Whiskers plot (Фигура 4.11 и Фигура 4.12) изглеждат сходно и това показва, че няма проби, които да бъдат изключени от анализа поради големите си различия в нормализираните стойности на интензитетите. Последно приложихме филтриране според промяната в експресията ≥ 2.0 на туморите спрямо нормите. Така при групата на АК останаха 107 микроРНКи, а при ПК 240.



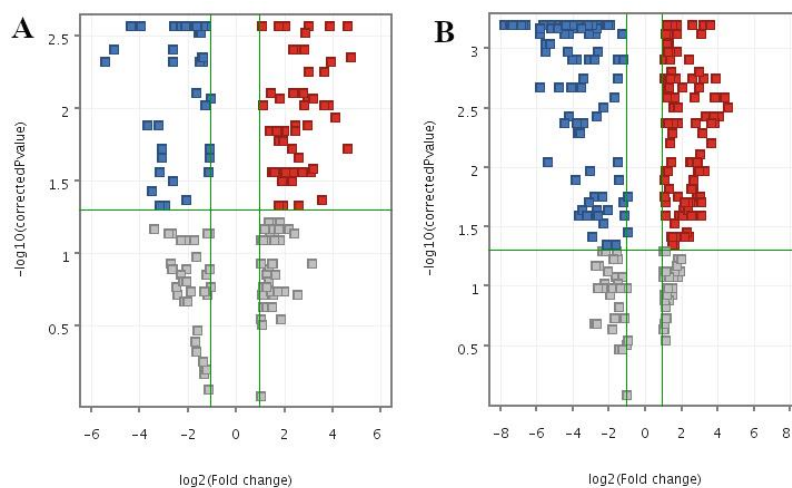
Фигура 4.11. Box Whiskers plot, показващ максималните и минималните стойности на всички проби; след прилагане на филтрите при анализа на данни при аденокарциноми и прилежащите им нормални тъкани



Фигура 4.12 BoxWhiskersplot, показващ максималните и минималните стойности на всички проби; след прилагане на филтрите при анализа на данни при плоскоклетъчните белодробни карциноми и прилежащите им нормални тъкани

4.1.5. Установяване на микроРНКи с променена експресия при АК и ПК в сравнение с прилежащите им нормални тъкани.

При сравнението между микроРНК експресионните профили на туморните тъкани на пациентите, при двете групи – АК и прилежащите им нормални тъкани и ПК и прилежащите им нормални тъкани, бе използван Mann–Whitney unpaired тест и Benjamini–Hochberg FDR за множествени сравнения. За значими разлики приехме промяна в експресията $FC \geq 2$ и p стойност ≤ 0.05 . На Фигура 4.13 е представен Volcano plot, визуализиращ по x-оста логаритъм от нивата на експресия на микроРНК-ите и по y-оста статистическата значимост на получените резултати. МикроРНК-ите, показали статистически значима разлика в експресията са представени в Таблицы 4.2- 4.5. При АК 107 микроРНКи са с променена експресия спрямо нормалните тъкани. От тях 68 са с увеличена и 39 с намалена експресия. При ПК 240 микроРНКи показаха статистически значима разлика в експресията. От тях 144 са с увеличена и 96 с намалена експресия. МикроРНКите, промяната в експресията и коригираната статистическа значимост са представени на Таблица 4.2 – 4.5.



Фигура 4.13. Volcano plot, визуализиращ нивата на експресия на микроРНК-ите и статистическата значимост на получените резултати. А) аденокарцином В) плоскоклетъчен белодробен карцином

Таблица 4.2. микроРНК-и показали статистически значима повишена експресия в туморна тъкан от АК спрямо нормална тъкан

микроРНК наименование	Регулация	р коригирана стойност	Промяна в експресията (FC)	logFC
hsa-let-7i-3p	повишена	0,008880214	5,036897	2,332535
hsa-miR-10a-3p	повишена	0,018734166	3,3082647	1,726075
hsa-miR-141-3p	повишена	0,011220638	2,173805	1,120223
hsa-miR-141-5p	повишена	0,005520372	14,835485	3,89098
hsa-miR-148a-5p	повишена	0,009959764	3,5163665	1,814085
hsa-miR-15b-3p	повишена	0,008880214	2,6883876	1,426741
hsa-miR-182-5p	повишена	0,003423761	24,329082	4,60461
hsa-miR-183-5p	повишена	0,005136078	26,979858	4,753811
hsa-miR-200a-5p	повишена	0,011220638	14,271401	3,835055
hsa-miR-200b-5p	повишена	0,011220638	12,479959	3,641541
hsa-miR-200c-5p	повишена	0,01608505	3,9692483	1,988866
hsa-miR-20a-3p	повишена	0,035487946	4,9870715	2,318193
hsa-miR-21-3p	повишена	0,003423761	4,0238457	2,008575
hsa-miR-21-5p	повишена	0,006423011	7,830559	2,969115
hsa-miR-210-3p	повишена	0,003423761	3,304915	1,724613
hsa-miR-214-5p	повишена	0,031034533	6,621358	2,727127
hsa-miR-215-5p	повишена	0,047834966	11,711017	3,549794
hsa-miR-26b-3p	повишена	0,01608505	2,629708	1,394903
hsa-miR-299-5p	повишена	0,004458654	6,958345	2,798744
hsa-miR-301b-3p	повишена	0,003423761	7,8708377	2,976517
hsa-miR-3065-3p	повишена	0,031034533	2,798849	1,484834
hsa-miR-31-3p	повишена	0,031034533	8,414936	3,072952
hsa-miR-31-5p	повишена	0,021514926	24,083584	4,589978
hsa-miR-3132	повишена	0,014781993	7,7578917	2,955665
hsa-miR-335-3p	повишена	0,01608505	3,3980398	1,764703
hsa-miR-340-3p	повишена	0,004458654	5,4209967	2,438558
hsa-miR-3591-3p	повишена	0,003423761	14,701666	3,877908

hsa-miR-3620-5p	повишена	0,028895691	8,950988	3,162047
hsa-miR-375	повишена	0,01608505	3,2141502	1,684437
hsa-miR-376b-3p	повишена	0,009959764	8,274848	3,048733
hsa-miR-382-5p	повишена	0,031034533	5,130276	2,359036
hsa-miR-409-3p	повишена	0,006423011	12,291252	3,61956
hsa-miR-421	повишена	0,01608505	3,1695185	1,664264
hsa-miR-4254	повишена	0,035487946	3,7360494	1,901514
hsa-miR-432-5p	повишена	0,031034533	3,999462	1,999806
hsa-miR-4449	повишена	0,009959764	7,807456	2,964853
hsa-miR-4451	повишена	0,003670778	7,146252	2,837187
hsa-miR-4483	повишена	0,031034533	4,1003013	2,03573
hsa-miR-450a-5p	повишена	0,011220638	6,8976216	2,786099
hsa-miR-4633-5p	повишена	0,021514926	4,931059	2,301898
hsa-miR-4750-5p	повишена	0,009959764	9,030754	3,174846
hsa-miR-493-5p	повишена	0,004458654	5,0781455	2,344302
hsa-miR-495-3p	повишена	0,014781993	5,3696446	2,424827
hsa-miR-503-5p	повишена	0,01608505	5,4881806	2,456328
hsa-miR-505-5p	повишена	0,018734166	3,8377619	1,940265
hsa-miR-514a-3p	повишена	0,031034533	3,0009358	1,585412
hsa-miR-542-3p	повишена	0,024962183	5,876831	2,555039
hsa-miR-550a-3p	повишена	0,01608505	3,9064121	1,965844
hsa-miR-605-5p	повишена	0,008880214	6,814134	2,76853
hsa-miR-6073	повишена	0,014781993	5,0289073	2,330245
hsa-miR-625-5p	повишена	0,005520372	19,524307	4,2872
hsa-miR-628-5p	повишена	0,01608505	3,2572668	1,703662
hsa-miR-629-3p	повишена	0,018734166	4,2672963	2,093322
hsa-miR-629-5p	повишена	0,003670778	7,2937202	2,866655
hsa-miR-650	повишена	0,013285798	14,309939	3,838946
hsa-miR-6516-3p	повишена	0,041264985	4,85152	2,278437
hsa-miR-6516-5p	повишена	0,003423761	7,3933396	2,886226
hsa-miR-654-3p	повишена	0,035487946	4,3100133	2,107692
hsa-miR-6741-3p	повишена	0,004155135	8,831817	3,14271
hsa-miR-6743-3p	повишена	0,007796806	14,396902	3,847687
hsa-miR-6804-3p	повишена	0,004155135	9,692731	3,276903
hsa-miR-6826-5p	повишена	0,003423761	2,067314	1,047758
hsa-miR-6834-3p	повишена	0,003423761	23,665821	4,564733
hsa-miR-6876-5p	повишена	0,028895691	9,941143	3,313412
hsa-miR-7-1-3p	повишена	0,006423011	10,354289	3,372157
hsa-miR-7-5p	повишена	0,014781993	14,566601	3,864592
hsa-miR-744-3p	повишена	0,004155135	11,383588	3,508884
hsa-miR-95-3p	повишена	0,028895691	9,92171	3,310589

Таблица 4.3. микроРНК-и показали статистически значима намалена експресия в туморна тъкан от АК спрямо нормална тъкан

микроРНК наименование	Регулация	p коригирана стойност	Промяна в експресията (FC)	logFC
hsa-miR-101-5p	намалена	0,014781993	-9,344661	-3,22414

hsa-miR-1207-5p	намелена	0,003423761	-3,1422064	-1,65178
hsa-miR-1225-5p	намелена	0,003423761	-2,8946111	-1,53337
hsa-miR-125b-2-3p	намелена	0,031034533	-8,884574	-3,1513
hsa-miR-133a-3p	намелена	0,035487946	-6,1939783	-2,63087
hsa-miR-1343-5p	намелена	0,004458654	-33,57634	-5,06937
hsa-miR-139-5p	намелена	0,047834966	-4,2133293	-2,07496
hsa-miR-140-3p	намелена	0,031034533	-2,247318	-1,1682
hsa-miR-144-3p	намелена	0,005520372	-6,214397	-2,63561
hsa-miR-1915-3p	намелена	0,005520372	-2,5604696	-1,35641
hsa-miR-195-3p	намелена	0,003423761	-15,5512085	-3,95895
hsa-miR-23a-5p	намелена	0,024962183	-8,375734	-3,06622
hsa-miR-2861	намелена	0,003670778	-2,9280784	-1,54995
hsa-miR-30a-5p	намелена	0,003423761	-5,9120617	-2,56366
hsa-miR-30c-2-3p	намелена	0,021514926	-8,467571	-3,08195
hsa-miR-30d-5p	намелена	0,005520372	-2,8768535	-1,52449
hsa-miR-3188	намелена	0,005520372	-43,267292	-5,43521
hsa-miR-3610	намелена	0,041264985	-11,191286	-3,4843
hsa-miR-3656	намелена	0,005136078	-2,6879265	-1,42649
hsa-miR-3665	намелена	0,009959764	-2,0713503	-1,05057
hsa-miR-378a-3p	намелена	0,021514926	-2,1282322	-1,08966
hsa-miR-378i	намелена	0,024962183	-2,1689517	-1,117
hsa-miR-4270	намелена	0,003423761	-4,867893	-2,2833
hsa-miR-4281	намелена	0,003423761	-2,8054564	-1,48824
hsa-miR-4433b-3p	намелена	0,014781993	-12,770636	-3,67476
hsa-miR-4466	намелена	0,003670778	-2,7933626	-1,482
hsa-miR-4516	намелена	0,003423761	-4,338719	-2,11727
hsa-miR-4687-3p	намелена	0,003423761	-2,8992429	-1,53568
hsa-miR-4689	намелена	0,003423761	-20,426617	-4,35238
hsa-miR-4739	намелена	0,003423761	-2,3932948	-1,259
hsa-miR-4763-3p	намелена	0,003423761	-2,5890677	-1,37243
hsa-miR-6068	намелена	0,008880214	-3,114682	-1,63908
hsa-miR-6088	намелена	0,003423761	-3,0602877	-1,61367
hsa-miR-6090	намелена	0,003423761	-3,2764704	-1,71214
hsa-miR-638	намелена	0,003423761	-3,1360958	-1,64897
hsa-miR-6722-3p	намелена	0,004458654	-6,2456465	-2,64285
hsa-miR-6821-5p	намелена	0,011220638	-2,4149294	-1,27198
hsa-miR-6848-3p	намелена	0,031034533	-10,132953	-3,34098
hsa-miR-762	намелена	0,005136078	-2,594818	-1,37563

Таблица 4.4. микроРНК-и показали статистически значима повишена експресия в туморна тъкан от ПК спрямо нормална тъкан

микроРНК наименование	Регулация	p коригирана стойност	Промяна в експресията (FC)	logFC
hsa-let-7b-3p	повишена	0,009450995	6,9629602	2,799701
hsa-miR-106b-3p	повишена	0,034239274	3,2127438	1,683806
hsa-miR-1180-3p	повишена	0,016110394	5,6314855	2,493516
hsa-miR-1182	повишена	0,030307354	4,0752764	2,026898

hsa-miR-1227-3p	повишена	0,016110394	3,5722535	1,836834
hsa-miR-1228-5p	повишена	0,030307354	4,1426644	2,050559
hsa-miR-1238-3p	повишена	0,014931935	2,0225632	1,016185
hsa-miR-1273c	повишена	0,021086326	5,723962	2,517014
hsa-miR-1281	повишена	0,00290732	2,571711	1,362729
hsa-miR-1306-3p	повишена	0,016110394	4,944136	2,305718
hsa-miR-1825	повишена	0,002144072	6,194722	2,63104
hsa-miR-193b-5p	повишена	0,034239274	4,9103827	2,295836
hsa-miR-196a-5p	повишена	0,016110394	8,129771	3,023215
hsa-miR-196b-5p	повишена	0,01828969	6,5592093	2,713522
hsa-miR-197-5p	повишена	0,014931935	2,528761	1,338431
hsa-miR-21-3p	повишена	0,004523036	2,3429286	1,228313
hsa-miR-210-3p	повишена	0,005004228	3,5730357	1,83715
hsa-miR-2276-3p	повишена	0,028413987	5,5440593	2,470943
hsa-miR-2278	повишена	0,045117393	4,3773937	2,130072
hsa-miR-3125	повишена	0,01828969	4,853183	2,278931
hsa-miR-3132	повишена	0,030307354	3,9277897	1,973718
hsa-miR-3147	повишена	0,030307354	4,778745	2,256632
hsa-miR-3148	повишена	0,009450995	10,903545	3,446725
hsa-miR-3154	повишена	0,030307354	4,3434176	2,118831
hsa-miR-3158-5p	повишена	0,045117393	3,4378452	1,781505
hsa-miR-3162-3p	повишена	0,003532344	2,4011977	1,263754
hsa-miR-32-3p	повишена	0,030307354	5,78749	2,532938
hsa-miR-345-3p	повишена	0,016110394	5,848755	2,54813
hsa-miR-3614-5p	повишена	0,008346809	11,603369	3,536472
hsa-miR-3620-5p	повишена	0,028413987	6,029007	2,59192
hsa-miR-3646	повишена	0,045117393	4,4279895	2,146652
hsa-miR-3654	повишена	0,007799984	15,24002	3,929793
hsa-miR-370-3p	повишена	0,008346809	12,831146	3,681578
hsa-miR-371a-5p	повишена	0,008346809	10,264218	3,359552
hsa-miR-3917	повишена	0,034239274	3,7711031	1,914987
hsa-miR-424-3p	повишена	0,030307354	6,3564763	2,668227
hsa-miR-4290	повишена	0,016110394	6,2810354	2,651002
hsa-miR-431-3p	повишена	0,016110394	6,2038565	2,633165
hsa-miR-4312	повишена	0,030307354	3,4529123	1,787814
hsa-miR-4323	повишена	0,030307354	3,0808277	1,623318
hsa-miR-4419b	повишена	0,005004228	11,538603	3,528397
hsa-miR-4428	повишена	0,008346809	4,54941	2,185679
hsa-miR-4436b-5p	повишена	0,013202623	6,280769	2,650941
hsa-miR-4449	повишена	0,01828969	5,7501974	2,523612
hsa-miR-4462	повишена	0,004523036	6,201169	2,63254
hsa-miR-4481	повишена	0,004523036	8,36529	3,064416
hsa-miR-4484	повишена	0,003532344	2,7726414	1,471261
hsa-miR-4485-5p	повишена	0,014931935	2,258655	1,175464
hsa-miR-4486	повишена	0,030307354	4,1276426	2,045318
hsa-miR-4496	повишена	0,014931935	8,162092	3,028939
hsa-miR-449b-3p	повишена	0,034239274	3,6325662	1,860989
hsa-miR-4505	повишена	0,045117393	2,1794972	1,123995
hsa-miR-4507	повишена	0,028413987	2,3322527	1,221724

hsa-miR-4513	повишена	0,008346809	8,085316	3,015304
hsa-miR-4522	повишена	0,045117393	3,8487191	1,944378
hsa-miR-4535	повишена	0,024385218	5,994929	2,583743
hsa-miR-4538	повишена	0,030307354	4,0777826	2,027785
hsa-miR-4539	повишена	0,009450995	10,860715	3,441047
hsa-miR-4636	повишена	0,011185372	9,948228	3,31444
hsa-miR-4644	повишена	0,021086326	4,6032567	2,202655
hsa-miR-4646-3p	повишена	0,008346809	6,866486	2,779572
hsa-miR-4647	повишена	0,039612453	6,446025	2,68841
hsa-miR-4649-3p	повишена	0,003232973	2,4495964	1,292544
hsa-miR-4653-3p	повишена	0,003232973	2,4719908	1,305673
hsa-miR-4688	повишена	0,024385218	5,177824	2,372346
hsa-miR-4716-5p	повишена	0,007799984	9,587915	3,261217
hsa-miR-4728-3p	повишена	0,008346809	6,8586516	2,777925
hsa-miR-4732-5p	повишена	0,030307354	4,4222474	2,14478
hsa-miR-4738-3p	повишена	0,011185372	9,500814	3,248051
hsa-miR-4746-3p	повишена	0,007799984	2,7513008	1,460114
hsa-miR-4749-3p	повишена	0,003021522	5,251813	2,392816
hsa-miR-4758-5p	повишена	0,011185372	16,172337	4,015456
hsa-miR-4793-3p	повишена	0,016110394	5,912246	2,563706
hsa-miR-483-3p	повишена	0,045117393	4,9339185	2,302734
hsa-miR-494-3p	повишена	0,021086326	2,328804	1,219589
hsa-miR-5190	повишена	0,030307354	3,9238179	1,972258
hsa-miR-550a-5p	повишена	0,01828969	6,4257903	2,683874
hsa-miR-5699-5p	повишена	0,014931935	13,172954	3,719507
hsa-miR-574-5p	повишена	0,007799984	2,330212	1,220461
hsa-miR-595	повишена	0,030307354	4,7910547	2,260343
hsa-miR-601	повишена	0,021086326	7,1884212	2,845675
hsa-miR-619-5p	повишена	0,004523036	2,6292298	1,39464
hsa-miR-622	повишена	0,024385218	8,427953	3,075182
hsa-miR-629-3p	повишена	0,034239274	3,9716263	1,98973
hsa-miR-650	повишена	0,028413987	6,740216	2,752795
hsa-miR-6512-5p	повишена	0,006768253	8,497485	3,087036
hsa-miR-6516-5p	повишена	0,016110394	4,5416155	2,183206
hsa-miR-663b	повишена	0,039612453	5,0675855	2,341299
hsa-miR-665	повишена	0,014931935	6,714665	2,747315
hsa-miR-6730-3p	повишена	0,034239274	3,3200555	1,731207
hsa-miR-6731-3p	повишена	0,003532344	8,672306	3,116416
hsa-miR-6732-3p	повишена	0,009450995	7,085099	2,824788
hsa-miR-6734-5p	повишена	0,003232973	2,4980998	1,320831
hsa-miR-6740-5p	повишена	0,004523036	2,3542647	1,235277
hsa-miR-6741-3p	повишена	0,008346809	7,359548	2,879617
hsa-miR-6743-3p	повишена	0,021086326	6,4703684	2,693848
hsa-miR-6743-5p	повишена	0,030307354	3,4168253	1,772656
hsa-miR-6745	повишена	0,004523036	10,093825	3,335401
hsa-miR-6746-5p	повишена	0,034239274	3,5467567	1,8265
hsa-miR-6751-3p	повишена	0,016110394	5,083593	2,345849
hsa-miR-6752-3p	повишена	0,006768253	2,5761395	1,365211
hsa-miR-6753-3p	повишена	0,013202623	5,9539332	2,573843

hsa-miR-6756-3p	повишена	0,008346809	6,4799275	2,695978
hsa-miR-6757-3p	повишена	0,034239274	3,2170844	1,685754
hsa-miR-6759-3p	повишена	0,008346809	6,4672995	2,693163
hsa-miR-6760-3p	повишена	0,014931935	3,732345	1,900082
hsa-miR-6763-3p	повишена	0,007799984	7,8476605	2,972263
hsa-miR-6765-5p	повишена	0,005004228	9,2943125	3,216348
hsa-miR-6767-5p	повишена	0,005879957	7,842166	2,971252
hsa-miR-6769a-5p	повишена	0,030307354	4,160573	2,056782
hsa-miR-6769b-5p	повишена	0,016110394	2,6072454	1,382526
hsa-miR-6772-5p	повишена	0,005879957	16,887386	4,077874
hsa-miR-6775-3p	повишена	0,030307354	3,6838362	1,881209
hsa-miR-6779-3p	повишена	0,008346809	6,0516405	2,597326
hsa-miR-6780b-5p	повишена	0,005879957	2,8273003	1,499425
hsa-miR-6785-5p	повишена	0,009450995	5,3530664	2,420366
hsa-miR-6788-5p	повишена	0,006768253	17,62149	4,139264
hsa-miR-6796-5p	повишена	0,005004228	9,571801	3,258791
hsa-miR-6797-5p	повишена	0,016110394	6,92942	2,792735
hsa-miR-6799-5p	повишена	0,004523036	13,520305	3,757056
hsa-miR-6804-3p	повишена	0,021086326	6,0390024	2,59431
hsa-miR-6812-3p	повишена	0,001980057	12,048807	3,590818
hsa-miR-6813-3p	повишена	0,013202623	2,01824	1,013098
hsa-miR-6826-5p	повишена	0,003021522	2,0723226	1,051249
hsa-miR-6856-5p	повишена	0,004523036	8,432757	3,076005
hsa-miR-6857-5p	повишена	0,003667966	16,763884	4,067285
hsa-miR-6858-3p	повишена	0,003021522	2,4181032	1,273876
hsa-miR-6858-5p	повишена	0,011185372	2,1978493	1,136092
hsa-miR-6861-3p	повишена	0,030307354	3,9687223	1,988675
hsa-miR-6872-5p	повишена	0,016110394	5,834934	2,544716
hsa-miR-6877-5p	повишена	0,004523036	25,806473	4,689661
hsa-miR-6879-5p	повишена	0,007799984	2,0425103	1,030343
hsa-miR-6880-3p	повишена	0,003232973	11,812307	3,562219
hsa-miR-6886-3p	повишена	0,016110394	6,2342434	2,640214
hsa-miR-6886-5p	повишена	0,030307354	4,0847874	2,030261
hsa-miR-6891-3p	повишена	0,016110394	5,2326274	2,387536
hsa-miR-6893-5p	повишена	0,009450995	3,4856844	1,801442
hsa-miR-7114-3p	повишена	0,016110394	4,6424594	2,214889
hsa-miR-744-3p	повишена	0,024385218	5,0652633	2,340637
hsa-miR-7515	повишена	0,003667966	22,115452	4,466983
hsa-miR-7844-5p	повишена	0,030307354	4,0869904	2,031039
hsa-miR-7851-3p	повишена	0,030307354	3,9097667	1,967083
hsa-miR-9-3p	повишена	0,030307354	4,83164	2,272513
hsa-miR-9500	повишена	0,014931935	8,632543	3,109786

Таблица 4.5. микроРНК-и показали статистически значима намалена експресия в туморна тъкан от ПК спрямо нормална тъкан

микроРНК наименование	Регулация	р коригирана стойност	Промяна в експресията (FC)	logFC
hsa-let-7a-5p	намелена	0,003667966	-8,714135	-3,12336
hsa-let-7b-5p	намелена	0,002527865	-4,2492824	-2,08722
hsa-let-7c-5p	намелена	0,003021522	-4,388827	-2,13384
hsa-let-7d-5p	намелена	0,00425998	-17,694942	-4,14527
hsa-let-7e-5p	намелена	0,003667966	-25,459307	-4,67012
hsa-let-7f-5p	намелена	0,00425998	-9,990323	-3,32053
hsa-let-7g-5p	намелена	0,004523036	-12,984713	-3,69874
hsa-let-7i-5p	намелена	0,013202623	-4,0499325	-2,0179
hsa-miR-1-3p	намелена	0,008346809	-17,436295	-4,12402
hsa-miR-100-5p	намелена	0,003021522	-13,123912	-3,71413
hsa-miR-100-3p	намелена	0,003021522	-17,862677	-4,15888
hsa-miR-101-5p	намелена	0,005004228	-8,341182	-3,06025
hsa-miR-101-3p	намелена	0,0030215217	-17,862677	-4,1588764
hsa-miR-103a-3p	намелена	0,003667966	-3,3932369	-1,76266
hsa-miR-107	намелена	0,005004228	-5,643177	-2,49651
hsa-miR-125b-5p	намелена	0,003021522	-7,489845	-2,90494
hsa-miR-126-3p	намелена	0,003021522	-65,08808	-6,02432
hsa-miR-126-5p	намелена	0,009450995	-26,434418	-4,72435
hsa-miR-130a-3p	намелена	0,003021522	-9,701295	-3,27818
hsa-miR-133a-3p	намелена	0,01828969	-7,0426726	-2,81612
hsa-miR-133b	намелена	0,001980057	-79,923485	-6,32055
hsa-miR-138-5p	намелена	0,009450995	-9,00127	-3,17013
hsa-miR-139-5p	намелена	0,006768253	-14,650818	-3,87291
hsa-miR-140-3p	намелена	0,001980057	-5,704569	-2,51212
hsa-miR-142-3p	намелена	0,014931935	-5,2000155	-2,37852
hsa-miR-143-5p	намелена	0,001980057	-39,71562	-5,31163
hsa-miR-144-3p	намелена	0,003021522	-34,770836	-5,11981
hsa-miR-145-3p	намелена	0,005004228	-30,2562	-4,91916
hsa-miR-145-5p	намелена	0,001980057	-9,7083	-3,27922
hsa-miR-150-5p	намелена	0,003667966	-8,272661	-3,04835
hsa-miR-15a-5p	намелена	0,003667966	-9,250584	-3,20954
hsa-miR-15b-5p	намелена	0,008346809	-6,2632985	-2,64692
hsa-miR-16-5p	намелена	0,007799984	-9,966979	-3,31716
hsa-miR-181a-3p	намелена	0,003532344	-33,040115	-5,04615
hsa-miR-181a-5p	намелена	0,002144072	-3,799132	-1,92567
hsa-miR-181b-5p	намелена	0,003667966	-3,132015	-1,64709
hsa-miR-185-5p	намелена	0,005004228	-6,5515018	-2,71183
hsa-miR-193a-3p	намелена	0,00425998	-6,2202854	-2,63698
hsa-miR-195-3p	намелена	0,01828969	-4,922339	-2,29934
hsa-miR-195-5p	намелена	0,003021522	-74,25374	-6,21439
hsa-miR-199a-3p	намелена	0,006768253	-11,681119	-3,54611
hsa-miR-199a-5p	намелена	0,005879957	-5,3011365	-2,4063

hsa-miR-218-5p	намелена	0,003232973	-132,12706	-7,04578
hsa-miR-22-5p	намелена	0,01828969	-8,568639	-3,09907
hsa-miR-222-3p	намелена	0,021086326	-2,2087471	-1,14323
hsa-miR-23a-3p	намелена	0,004523036	-5,365813	-2,4238
hsa-miR-23a-5p	намелена	0,003232973	-11,734796	-3,55272
hsa-miR-23b-3p	намелена	0,003667966	-6,822757	-2,77035
hsa-miR-24-3p	намелена	0,008346809	-2,6289	-1,39446
hsa-miR-25-3p	намелена	0,034239274	-2,4081726	-1,26794
hsa-miR-26a-5p	намелена	0,002220771	-8,034105	-3,00614
hsa-miR-27a-3p	намелена	0,006768253	-5,6259255	-2,49209
hsa-miR-27b-3p	намелена	0,004523036	-5,1309695	-2,35923
hsa-miR-29a-3p	намелена	0,003021522	-6,4013658	-2,67838
hsa-miR-29b-3p	намелена	0,003667966	-6,9939594	-2,80611
hsa-miR-29c-3p	намелена	0,003021522	-7,811022	-2,96551
hsa-miR-3065-5p	намелена	0,005004228	-13,602806	-3,76583
hsa-miR-30a-3p	намелена	0,001980057	-126,85476	-6,98703
hsa-miR-30a-5p	намелена	0,001980057	-8,594262	-3,10337
hsa-miR-30b-3p	намелена	0,003532344	-14,957311	-3,90278
hsa-miR-30b-5p	намелена	0,003021522	-35,44236	-5,1474
hsa-miR-30c-2-3p	намелена	0,00290732	-18,071358	-4,17563
hsa-miR-30c-5p	намелена	0,003021522	-20,806171	-4,37894
hsa-miR-30d-3p	намелена	0,003532344	-12,7699795	-3,67468
hsa-miR-30d-5p	намелена	0,001980057	-4,244131	-2,08547
hsa-miR-30e-3p	намелена	0,003667966	-38,121853	-5,25255
hsa-miR-30e-5p	намелена	0,003021522	-2,276849	-1,18704
hsa-miR-32-5p	намелена	0,01828969	-7,89558	-2,98105
hsa-miR-326	намелена	0,002527865	-17,316021	-4,11404
hsa-miR-331-3p	намелена	0,006768253	-2,568184	-1,36075
hsa-miR-335-5p	намелена	0,005004228	-34,269066	-5,09884
hsa-miR-338-3p	намелена	0,003021522	-79,36461	-6,31042
hsa-miR-342-3p	намелена	0,002220771	-4,213197	-2,07492
hsa-miR-342-5p	намелена	0,00425998	-27,308702	-4,77129
hsa-miR-34a-5p	намелена	0,00425998	-9,197053	-3,20117
hsa-miR-34c-5p	намелена	0,01828969	-29,863344	-4,9003
hsa-miR-361-5p	намелена	0,003532344	-3,9141662	-1,9687
hsa-miR-3610	намелена	0,039612453	-4,5718303	-2,19277
hsa-miR-362-3p	намелена	0,009450995	-17,283339	-4,11131
hsa-miR-374c-5p	намелена	0,016110394	-8,011072	-3,002
hsa-miR-375	намелена	0,005879957	-11,739739	-3,55333
hsa-miR-4284	намелена	0,045117393	-2,199938	-1,13746
hsa-miR-4291	намелена	0,024385218	-2,23119	-1,15781
hsa-miR-4306	намелена	0,008346809	-2,0431116	-1,03077
hsa-miR-4324	намелена	0,003232973	-21,45688	-4,42337
hsa-miR-4516	намелена	0,014931935	-2,1456094	-1,10139
hsa-miR-451a	намелена	0,003021522	-98,50006	-6,62205
hsa-miR-455-5p	намелена	0,016110394	-9,880085	-3,30452
hsa-miR-486-5p	намелена	0,001980057	-14,332465	-3,84122
hsa-miR-497-5p	намелена	0,002220771	-17,123777	-4,09793
hsa-miR-500a-5p	намелена	0,034239274	-3,6371903	-1,86282

hsa-miR-501-3p	намелена	0,024385218	-5,821932	-2,5415
hsa-miR-511-3p	намелена	0,01828969	-14,248615	-3,83275
hsa-miR-551b-3p	намелена	0,008346809	-37,568012	-5,23143
hsa-miR-98-5p	намелена	0,009450995	-21,11362	-4,4001
hsa-miR-99a-5p	намелена	0,003021522	-16,09832	-4,00884
hsa-miR-99b-5p	намелена	0,001980057	-4,0189724	-2,00683

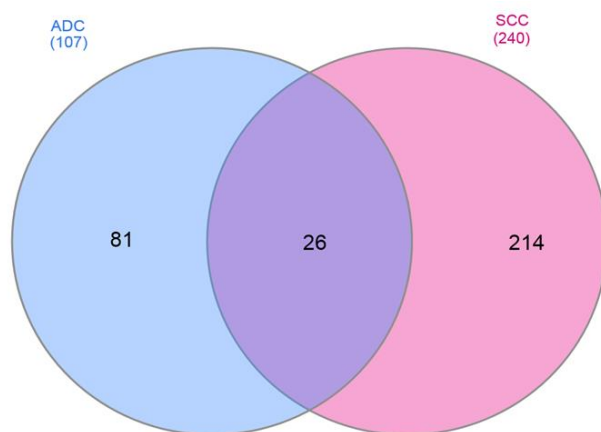
4.1.6. Резултати от йерархичен клъстерен анализ.

За да определим дали експресионните профили на туморните и нормалните тъкани в двете изследвани групи се разграничават, проведохме йерархичен клъстерен анализ, като анализът спомогна при подбора на микроРНК-и за последваща валидация в по-голяма извадка. При АК клъстерния анализ установи, че има добро разделяне на туморните и нормалните тъкани като се обособяват два отделни клъстера (групи) (Фигура 4.16). Това се дължи основно на микроРНК-ите с повишена експресия. При АК тумори се разграничават 2 клъстера – по-малък, състоящ се от 5 пациента и по-голям със 7 пациента. В първия клъстер се намират основно пациенти в напреднал стадии, нискодиференцирани диференцирани тумори, с наличие на нодални метастази и смесена морфология – ацинарно-папилярни и солидни структури, както и наличие на муцинозна компонента. Във втория клъстер са пациентите с ранен туморен стадии, добре диференцирани или умерено диференцирани тумори, без нодални метастази и ацинарна, папиларна или смесена ацинарно-папиларна морфология.

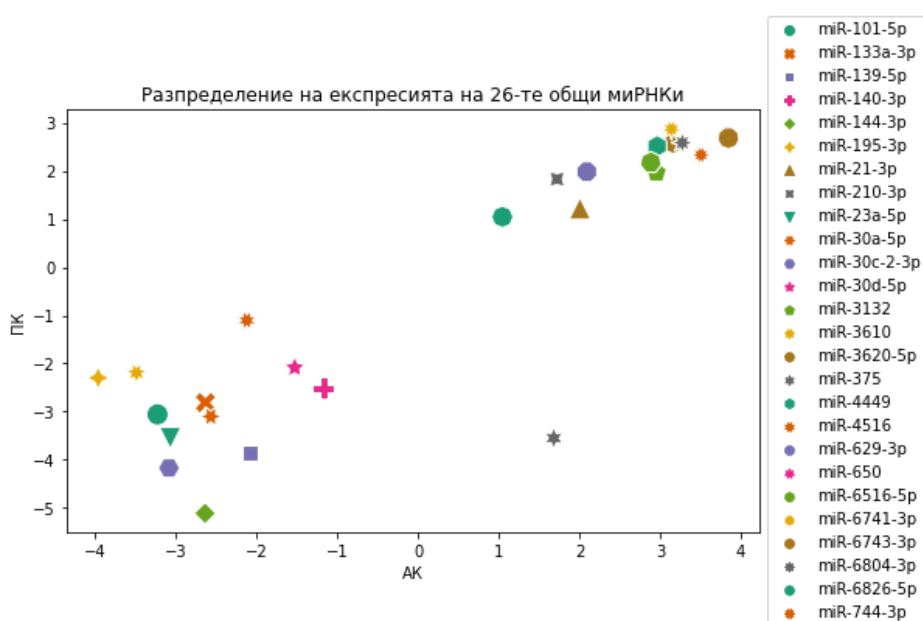
При ПК също наблюдавахме добро разделяне на туморните и нормалните тъкани на две групи (Фигура 4.17). При ПК туморите се разграничават 2 клъстера – единият с 5, а другият с 6 пациента. В клъстера с 6 пациента са пациентите в напреднал стадии, нискодиференцирани диференцирани тумори, с наличие на нодални и далечни метастази. Според морфологията те се разделят на 2 подкълъстера – при трима пациента се наблюдава кератинизация, а останалите трима са без кератинизация. Във втория клъстер са пациентите с ранен туморен стадии, добре диференцирани или умерено диференцирани тумори, без нодални метастази, като при един от тях не се наблюдава кератинизация.

4.1.7. Избор на микроРНКи за валидация.

На фигура 4.14 е представена Вен диаграма, показваща броя на общите и различните микроРНКи, които са с променена експресия при АК и ПК. За АК са характерни 81 микроРНКи, а за ПК специфични са 214 микроРНКи. Като 26 микроРНКи са дерегулирани общо и при двата подтипа. За да представим визуално разпределението на общите 26 микроРНКи използвахме скатър графиката, представена на Фигура 4.15.

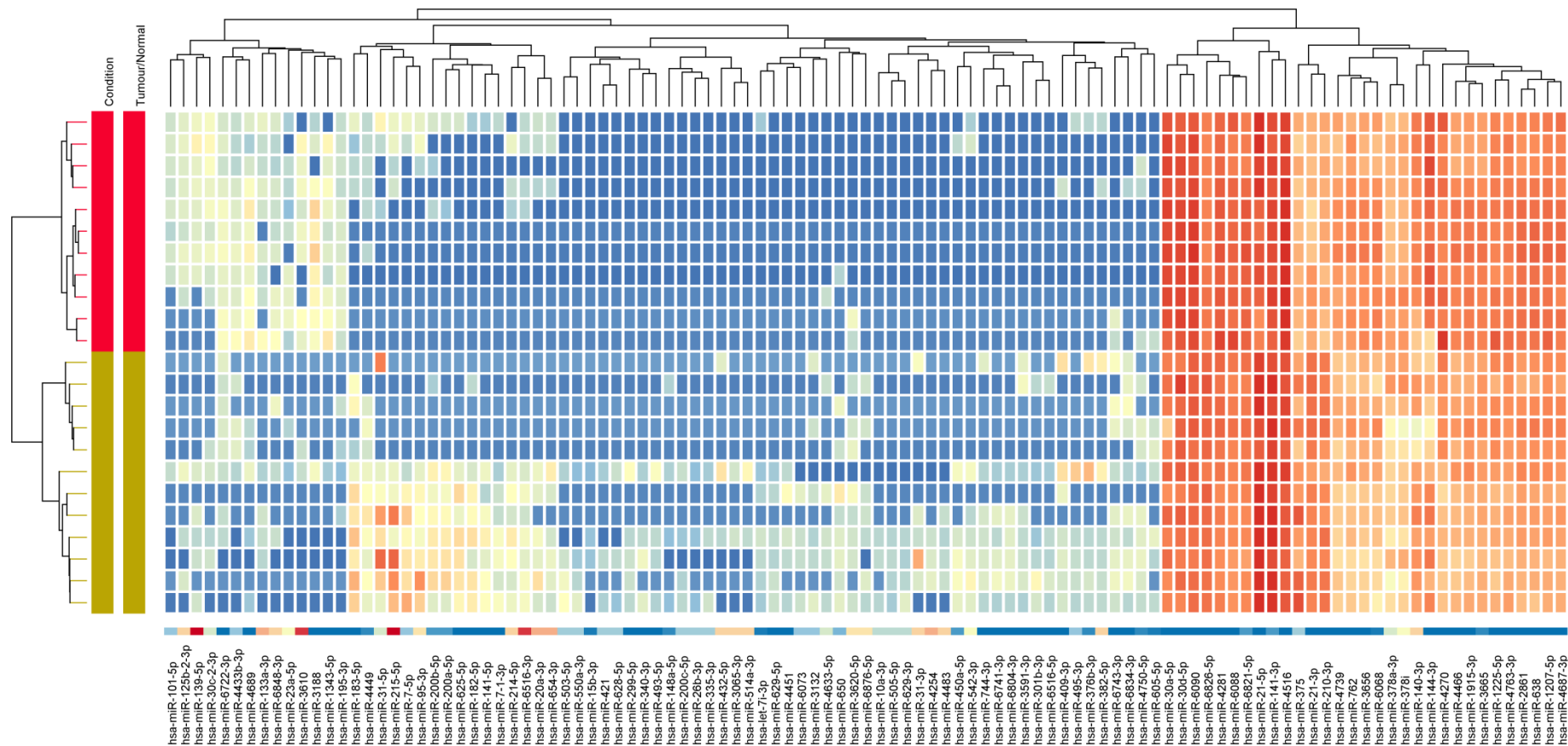


Фигура 4.14. Вен диаграми с броя на общите и различните микроРНК при АК и ПК (217)

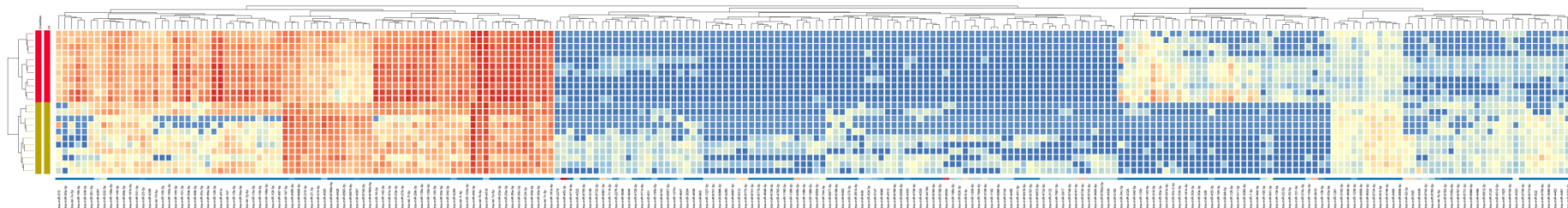


Фигура 4.15. Скатър графика, показваща разпределението на експресията на общите микроРНК при двата подтипа белодробен карцином.

От 26-те общи микроРНК за двата подтипа белодробен карцином бяха избрани 8 микроРНК - miR-744-3p, miR-650, miR-375, miR-21-3p, miR-140-3p, miR-144-3p, miR-195-3p и miR-139-5p. При внимателно разглеждане на специфичните за АК и ПК бяха подбрани още 6 микроРНК - 3 характерни за АК - miR-4689, miR-7-5p и miR-95-3p и 3 характерни за ПК - miR-101-3p, miR-195-5p, и miR-6785-5p. Всички 14 избрани микроРНК показват статистически значима промяна в експресията в сравнение с прилежащите нормални тъкани.



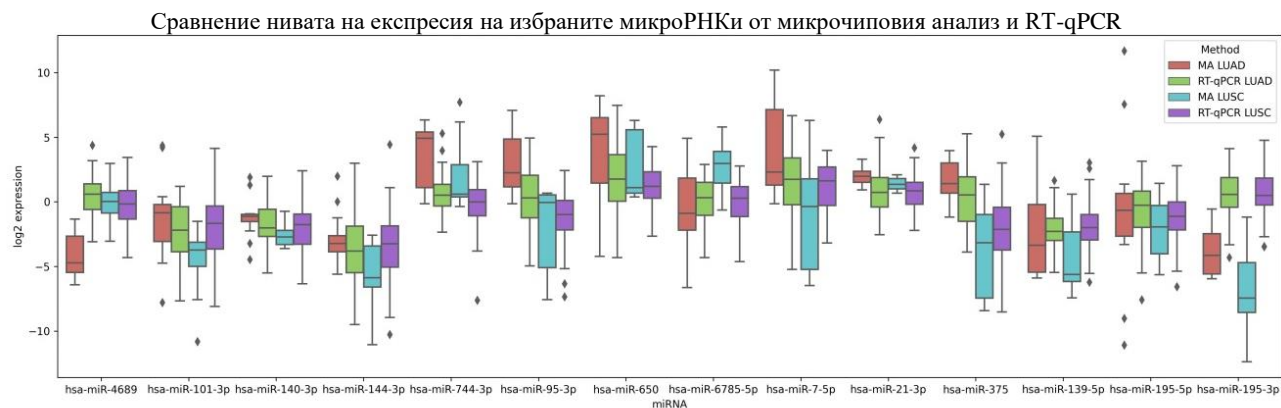
Фигура 4.17. Йерархичен клъстерен анализ въз основа на диференциално експресираниите микроРНК-и при туморите от пациенти, диагностирани с аденокарцином на белия дроб и прилежащите нормални тъкани след прилагане на филтриране на микроРНК-ите ($p \leq 0,05$ и 2.0 пъти промяна в експресията). Използвано е Евклидовото разстояние при клъстерирането, правило на Ward, Mann–Whitney unpaired тест и Benjamini–Hochberg FDR корекция на p стойността. Обозначения: T- тумор; N- норма; червено – повишена експресия, синьо – намалена експресия.



Фигура 4.16. Йерархичен клъстерен анализ въз основа на диференциално експресираниите микроРНК-и при туморите от пациенти, диагностицирани с плоскоклетъчен карцином на белия дроб и прилежащите нормални тъкани след прилагане на филтриране на микроРНК-ите ($p \leq 0,05$ и 2.0 пъти промяна в експресията). Използвано е Евклидовото разстояние при клъстерирането, правило на Ward, Mann–Whitney unpaired тест и Benjamini–Hochberg FDR корекция на p стойността. Обозначения: T- тумор; N- норма; червено – повишена експресия, синьо – намалена експресия.

4.2. Валидация на избраните микроРНКи чрез RT-qPCR.

Валидацията на 14-те избрани от микрочиповия анализ микроРНКи беше извършена чрез RT-qPCR при 50 АК и 50 ПК и прилежащите им нормални тъкани. Сравнение на нивата на експресия, определена чрез двата метода, е представено на Фигура 4.18.



Фигура 4.18. Сравнение на нивата на експресия на избраните за валидация 14 микроРНКи, установена чрез два метода – микрочипов анализ (МА) и RT-qPCR. С червени бокс плотове е експресията, установена чрез МА при АК; зелени бокс плотове – експресията, установена чрез RT-qPCR при АК; сини бокс плотове – експресията, установена чрез МА при ПК; лилави бокс плотове – експресията, установена чрез RT-qPCR при ПК. По х-оста са 14 микроРНКи, а по у-оста е представен \log_2 от експресията.

За да проверим дали има разлика в разпределението на резултатите между двата метода на изследване – микрочипов анализ и RT-qPCR, проведехме Kolmogorov-Smirnov тест. Тестът показва разлики в разпределението на експресията на изследваните микроРНК между двата набора от данни както в групите АК, така и при ПК. Така за групата на АК се валидираха 8 микроРНКи - 3 с повишена експресия: miR-7-5p ($p = 0.144$), miR-375 ($p = 0.117$), miR-6785-3p ($p = 0.060$) и 5 с намалена експресия: miR-101-3p ($p = 0.353$), miR-139-5p ($p = 0.051$), miR-140-3p ($p = 0.179$), miR-144-3p ($p = 0.179$) и miR-195-5p ($p = 0.427$). Шест микроРНКи не се валидираха с RT-qPCR при АК - miR-4689 ($p = 0.0001$), miR-744-3p ($p = 0.001$), miR-95-3p ($p = 0.003$), miR-650 ($p = 0.004$), miR-21-3p ($p = 0.001$), miR-195-3p ($p = 0.0001$). При ПК се валидираха общо 10 микроРНКи - 3 с повишена експресия: miR-7-5p ($p = 0.051$), miR-21-3p ($p = 0.057$), miR-650 ($p = 0.051$), и 7 с намалена експресия: miR-4689 ($p = 0.980$), miR-140-3p ($p = 0.053$), miR-744-3p ($p = 0.115$), miR-95-3p ($p = 0.455$), miR-375 ($p = 0.301$), miR-195-5p ($p = 0.265$) и miR-144-3p ($p = 0.051$). Експресията на четири микроРНКи не се валидираха с RT-qPCR при ПК - miR-101-3p ($p = 0.038$), miR-6785-3p ($p = 0.002$), miR-139-5p ($p = 0.001$) и miR-195-3p ($p = 0.0001$).

4.3. Нива на експресия на 14-те изследвани микроРНКи в тъканни проби от рак на белия дроб.

Нивата на експресия на подбраните за валидация 14 микроРНКи бяха изследвани и анализирани при 50 пациента с АК и 50 пациента с ПК. МикроРНКте, които бяха изследвани са както следва: miR-744-3p, miR-650, miR-375, miR-21-3p, miR-140-3p, miR-144-3p, miR-195-3p, miR-139-5p, miR-101-3p, miR-7-5p, miR-95-3p, miR-195-5p, miR-4689 и miR-6785-5p.

На Таблица 4.6 и Таблица 4.7 са представени брой и процент пациенти, средна стойност на относителната експресия съответно в групите показващи намалена, без промяна и повишени нива на експресия при АК и ПК.

Таблица 4.6. Относителна експресия на изследваните пациенти с АК. Представени са средните RQ стойности при пациенти, които показват понижени, без промяна и повишени в експресионни нива.

микроРНК	Пациенти с понижени експресионни нива n (%) Средно RQ±SD	Пациенти с нормални нива на експресия n (%) Средно RQ±SD	Пациенти с повишени нива на експресия n (%) Средно RQ±SD	Значимост p, χ^2
miR-4689	9 (18) 0.326 ± 0.045	23 (46) 1.363 ± 0.108	18 (36) 5.462 ± 1.301	0.048
miR-101-3p	34 (68) 0.145 ± 0.025	14 (28) 1.109 ± 0.126	2 (4) 4.368 ± 3.580	>0.0001
miR-140-3p	34 (68) 0.213 ± 0.022	15 (30) 0.974 ± 0.110	1 (2) 3,945	>0.0001
miR-144-3p	44 (88) 0.114 ± 0.020	4 (8) 0.980 ± 0.228	2 (4) 0.297±0.119	>0.0001
miR-744-3p	10 (20) 0.371 ± 0.023	21 (42) 1.136 ± 0.077	19 (38) 5.800 ± 1.981	0.127
miR-95-3p	17 (34) 0.273 ± 0.038	13 (26) 1.083 ± 0.124	20 (40) 8.981 ± 1.944	0.477
miR-650	5 (10) 0.210 ± 0.076	12 (24) 1.099 ± 0.143	33 (66) 14.763 ± 3.541	>0.001
miR-6785-5p	15 (30) 0.346 ± 0.038	15 (30) 1.127 ± 0.108	20 (40) 3.789 ± 0.326	0.645
miR-7-5p	9 (18) 0.254 ± 0.047	12 (24) 1.112 ± 0.095	29 (58) 19.083 ± 4.665	0.001
miR-21-3p	6 (12) 0.317 ± 0.047	23 (46) 1.121 ± 0.092	21 (42) 10.105 ± 3.873	0.005
miR-375	16 (32) 0.228 ± 0.032	16 (32) 1.277 ± 0.115	18 (36) 12.049 ± 2.616	0.923
miR-139-5p	38 (76) 0.185 ± 0.018	9 (18) 0.911 ± 0.145	3 (6) 6.103 ± 3.480	>0.0001
miR-195-5p	20 (40) 0.180 ± 0.033	17 (34) 1.036 ± 0.087	13 (26) 4.104 ± 0.622	0.375
miR-195-3p	8 (16) 0.287 ± 0.058	20 (40) 1.110 ± 0.088	22 (44) 5.693 ± 0.842	0.032

* Брой на пациенти с АК= 50

Таблица 4.7. Относителна експресия на изследваните пациенти с ПК. Представени са средните RQ стойности при пациенти, които показват понижени, без промяна и повишени в експресионни нива.

микроРНК	Пациенти с понижени експресионни нива n (%) Средно RQ±SD	Пациенти с нормални нива на експресия n (%) Средно RQ±SD	Пациенти с повишени нива на експресия n (%) Средно RQ±SD	Значимост p, χ^2
miR-4689	15 (30) 0.217 ± 0.036	23 (46) 1.072 ± 0.081	12 (24) 3.786 ± 0.697	0.144
miR-101-3p	29 (58) 0.140 ± 0.025	14 (28) 0.840 ± 0.091	7 (14) 6.664 ± 2.298	0.001
miR-140-3p	33 (66) 0.196 ± 0.025	14 (28) 0.823 ± 0.098	3 (6) 3.336 ± 1.000	>0.0001
miR-144-3p	40 (80) 0.124 ± 0.018	8 (16) 1.195 ± 0.130	2 (4) 11.806 ± 9.659	>0.0001
miR-744-3p	15 (20) 0.293 ± 0.043	23 (42) 1.185 ± 0.092	12 (38) 3.568 ± 0.528	0.144
miR-95-3p	24 (48) 0.229 ± 0.030	19 (36) 0.945 ± 0.079	7 (16) 3.198 ± 0.422	0.010
miR-650	8 (16) 0.186 ± 0.047	15 (30) 1.241 ± 0.118	27 (54) 5.811 ± 0.877	0.004
miR-6785-5p	13 (26) 0.281 ± 0.042	20 (40) 1.101 ± 0.071	17 (34) 3.701 ± 0.414	0.470
miR-7-5p	9 (18) 0.282 ± 0.041	8 (16) 1.059 ± 0.127	33 (66) 9.668 ± 1.671	>0.0001
miR-21-3p	14 (26) 0.309 ± 0.037	16 (36) 1.398 ± 0.100	20 (40) 5.307 ± 0.914	0.571
miR-375	32 (64) 0.151 ± 0.025	10 (20) 0.870 ± 0.103	8 (16) 8.762 4.173	>0.0001
miR-139-5p	37 (74) 0.215 ± 0.022	10 (20) 0.714 ± 0.092	3 (6) 5.801 ± 1.408	>0.0001
miR-195-5p	27 (54) 0.252 ± 0.028	19 (38) 1.050 ± 0.094	4 (8) 3.881 ± 1.017	>0.0001
miR-195-3p	7 (14) 0.227 ± 0.051	22 (44) 1.123 ± 0.076	21 (42) 6.115 ± 1.286	0.015

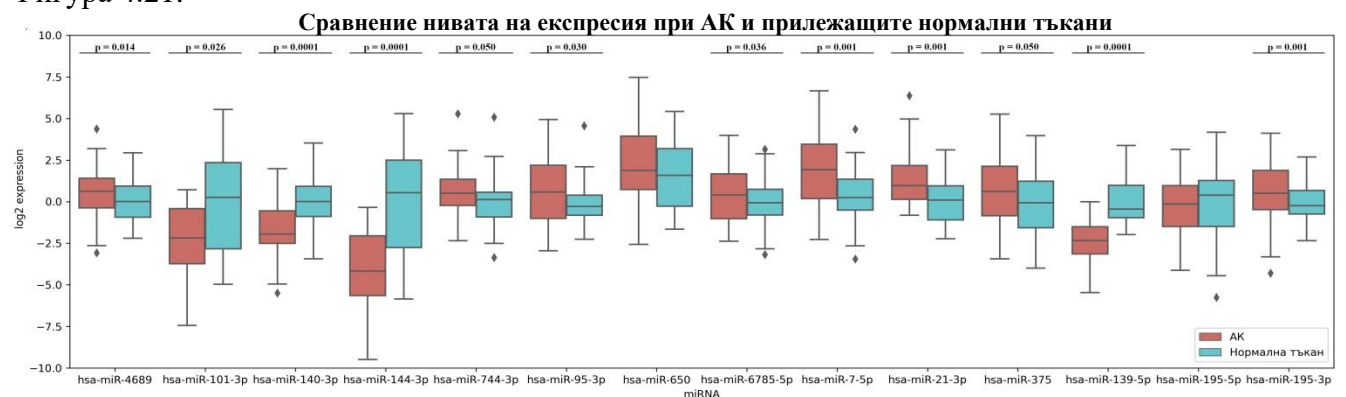
* Брой на пациенти с ПК= 50

За да се определи, кои от изследваните микроРНКи са статистически значимо променени в туморна белодробна тъкан в сравнение с нормална белодробна тъкан, в зависимост от разпределението на данните, се проведе съответния параметричен или непараметричен статистически тест (Independent T-тест или Mann–Whitney unpaired тест). Бяха използвани RQ стойностите за всяка от изследваните микроРНКи.

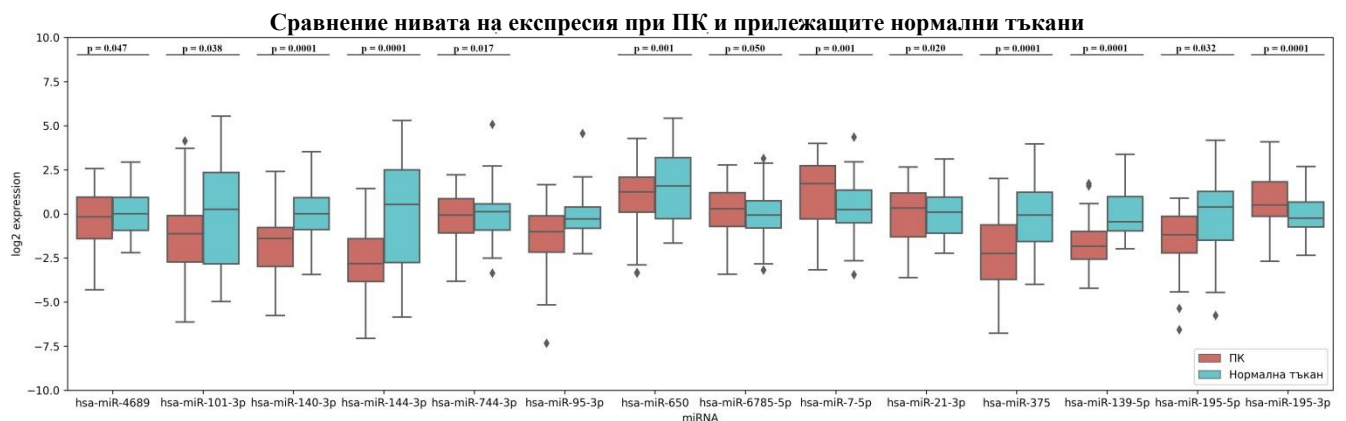
При АК само за miR-650 ($p = 0.784$) и miR-195-5p ($p = 0.231$) не се наблюдава статистически значима разлика в експресионните нива в туморите спрямо нормалните тъкани. Всички останали изследвани микроРНКи при АК бяха със значимо променена експресия в сравнение със съответстващата нормална тъкан. Експресионните нива на miR-4689 ($p = 0.014$), miR-744-3p ($p = 0.050$), miR-95-3p ($p = 0.030$), miR-6785-5p ($p = 0.036$), miR-7-5p ($p = 0.001$), miR-21-3p ($p = 0.001$), miR-375 ($p = 0.050$), miR-195-3p ($p = 0.001$) бяха значително повишени, докато miR-101-3p ($p = 0.026$), miR-140-3p ($p \leq 0.0001$), miR-144-3p ($p \leq 0.0001$), miR-139-5p ($p \leq 0.0001$) бяха значително понижени при аденокарциномите в сравнение с нормалната тъкан. Преобразуваните експресионни нива са представени чрез бокс плот на Фигура 4.19.

При ПК само за miR-95-3p ($p = 0.382$) не се наблюдава статистически значима разлика в експресионните нива в туморите спрямо нормалните тъкани. Всички останали изследвани микроРНКи при ПК бяха със значимо променена експресия в сравнение с нормална тъкан. Експресионните нива на miR-6785-5p ($p = 0.021$), miR-7-5p ($p \leq 0.0001$), miR-21-3p ($p = 0.002$), miR-650 ($p = 0.001$), miR-195-3p ($p \leq 0.0001$) бяха значително повишени, докато miR-4689 ($p = 0.047$), miR-101-3p ($p = 0.038$), miR-140-3p ($p \leq 0.0001$), miR-144-3p ($p \leq 0.0001$), miR-744-3p ($p = 0.017$), miR-375 ($p = 0.0001$), miR-195-5p ($p = 0.032$) бяха значително понижени при плоскоклетъчните карциноми в сравнение с нормалната тъкан. Преобразуваните експресионни нива са представени чрез бокс плот на Фигура 4.20.

При сравняване на експресионните нива на изследваните микроРНКи при АК и ПК 3 микроРНКи не показаха статистически значима разлика в експресията между двата подтипа белодробен карцином - miR-140-3p ($p = 0.777$), miR-6785-5p ($p = 0.950$), miR-195-3p ($p = 0.844$). Повишени при АК сравнение с ПК бяха miR-4689 ($p = 0.015$), miR-744-3p ($p = 0.031$), miR-95-3p ($p = 0.0001$), miR-650 ($p = 0.013$), miR-7-5p ($p = 0.031$), miR-21-3p ($p = 0.007$), miR-375 ($p \leq 0.0001$), miR-195-5p ($p = 0.004$), докато miR-101-3p ($p = 0.037$), miR-144-3p ($p = 0.003$), miR-139-5p ($p = 0.048$) бяха със значително намалена експресия при аденокарциномите в сравнение с плоскоклетъчните карциноми. Преобразуваните експресионни нива са представени чрез бокс плот на Фигура 4.21.

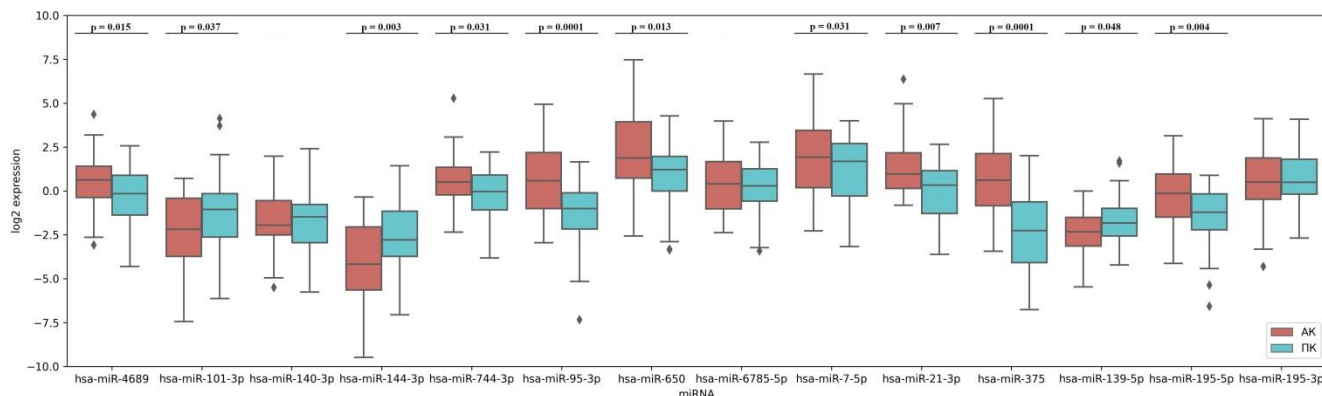


Фигура 4.19. Сравнение нивата на експресия при АК и прилежащите нормални тъкани. По x-оста са микроРНКите, по y-оста \log_2 от експресионните нива. Беше използван Independent T-тест или Mann-Whitney unpaired тест, според разпределението на данните; $p \geq 0.05$ е незначимо



Фигура 4. 20. Сравнение нивата на експресия при ПК и прилежащите нормални тъкани. По x-оста са микроРНКите, по y-оста \log_2 от експресионните нива на микроРНКите. Беше използван Independent T-тест или Mann-Whitney unpaired тест, според разпределението на данните; $p \geq 0.05$ е незначимо;

Сравнение нивата на експресия при АК и ПК



Фигура 4.21. Сравнение нивата на експресия при АК и ПК. По x-оста са микроРНКите, по y-оста \log_2 от експресионните нива на микроРНКите. Беше използван Independent T-тест или Mann-Whitney unpaired тест, според разпределението на данните; $p \geq 0.05$ е незначимо;

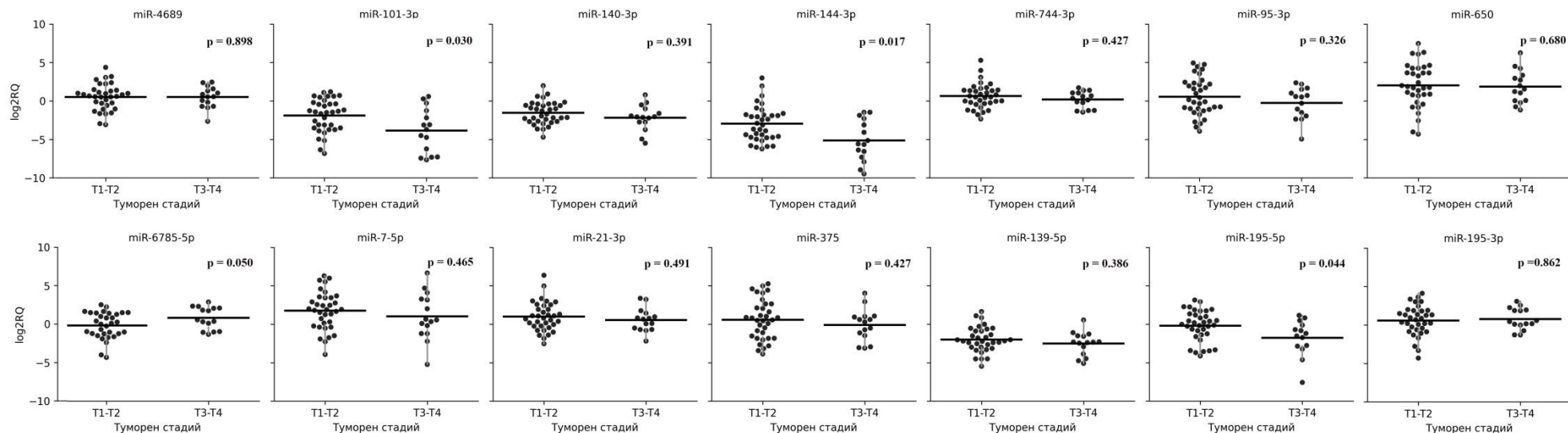
4.4. Асоциация между експресионните нива на микроРНКите и клиникопатологичните характеристики на пациентите.

Асоциацията между получените RQ данни на изследваните микроРНКи и групирането, базирано на клиникопатологичните характеристики на аденокарциномите е представено на Фигури 4.22 - 4.29.

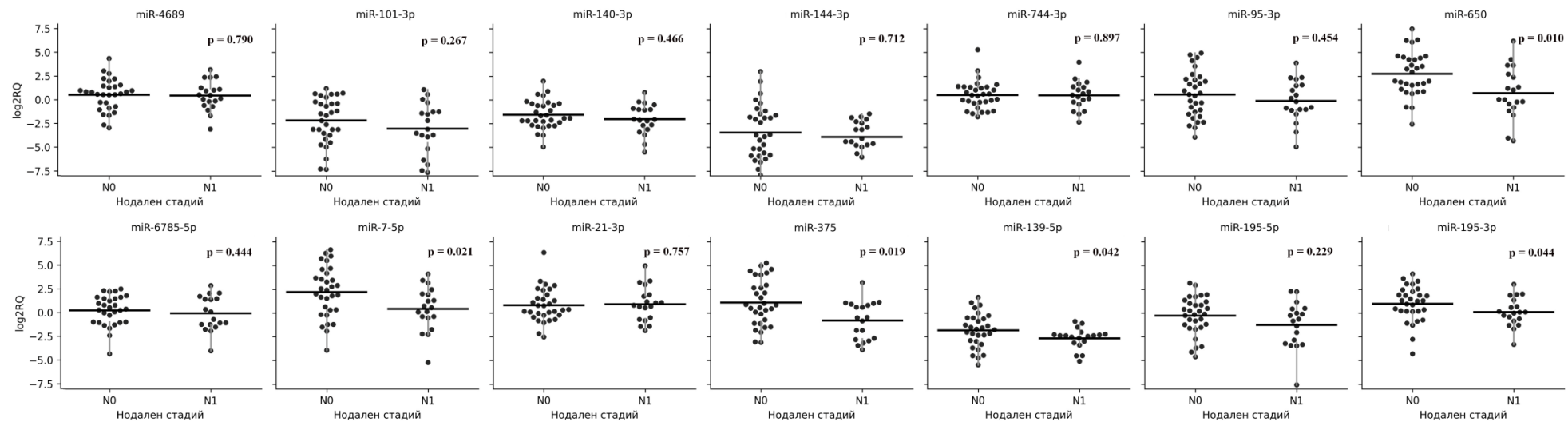
Получените данни показаха статистически значима асоциация между експресионните нива на изследваните микроРНКи с различни клиникопатологичните характеристики. MiR-101-3p се асоциира статистически значимо с туморния статус ($p = 0.030$) и намалена експресия при наличие на метастази ($p = 0.009$). Експресионните нива на miR-144-3p ($p = 0.017$) бяха повишени при пациенти, диагностицирани с T1-T2 спрямо T3-T4. MiR-744-3p ($p = 0.014$), miR-95-3p ($p = 0.048$) се характеризират с повишена експресия при пациенти, негативни за далечни метастази. Експресията на miR-650 е намалена при пациенти, положителни за нодални метастази ($p = 0.010$) и далечни метастази ($p = 0.038$). Повишена експресия на miR-6785-5p ($p = 0.050$) се асоциира с прогресията на T-стадия. Нивата на miR-7-5p намаляват с наличието на лимфни ($p = 0.021$) и далечни метастази ($p = 0.003$). Понижената експресия на miR-375 се асоциира с лимфните ($p=0.019$) и далечните метастази ($p=0.031$). MiR-139-5p се асоциира с нодалния статус ($p=0.042$) и метастазите ($p=0.018$). Понижената експресия на miR-195-5p се асоциира с туморния статус ($p=0.044$) и метастазите ($p=0.035$), докато експресията на другото рамо miR-195-3p се асоциира с нодалния статус ($p=0.044$) и наличието на метастази ($p=0.036$). Спрямо разделянето по морфология на АК експресията на miR-144-3p ($p=0.004$) показва по-висока експресия при смесените АК с преобладаваща солидна компонента. Експресията на miR-140-3p ($p = 0.015$) е най-ниска при ацинарните АК, средна при смесените смесените с преобладаваща микропапиларна компонента карциноми и най-висока при смесените солидни АК. За последните е характерна статистически значимо по-високо ниво на експресия за miR-95-3p ($p = 0.002$), miR-7-5p ($p = 0.012$), miR-375 ($p = 0.007$), miR-139-5p ($p = 0.050$), miR-195-5p ($p = 0.001$) спрямо другите три морфологични варианта АК. За АК с

преобладаваща ацинарна компонента е характерна повишена експресия на miR-650 ($p = 0.001$), miR-6785-5p ($p = 0.049$), miR-4689 ($p = 0.001$) и miR-195-3p ($p = 0.008$). При асоциациите по пол miR-6785-5p ($p = 0.014$) показва значително по-висока експресия при жените сравнение с мъжете. Нивата на miR-4689 ($p = 0.010$) и miR-195-5p ($p = 0.047$) за значително по-високи при пациентите под 60 години. Нито един от изследваните биомаркери не показва статистически значима асоциация с туморната диференциация и локацията на тумора (ляв или десен лоб на белия дроб).

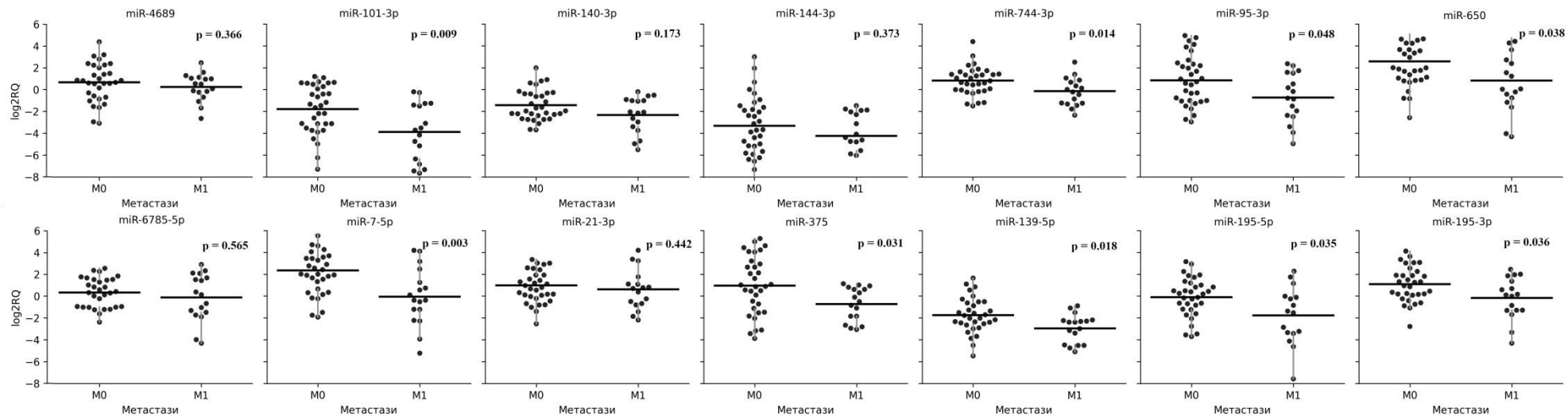
На Фигури 4.30 – 4.37 са представени асоциациите между RQ данните на изследваните микроРНК и групирането, базирано на клиникопатологичните характеристики при ПК. Всички микроРНК показваха статистически значима асоциация между експресионните нива с различни клиникопатологичните характеристики. Не се наблюдава асоциация на нивата на експресия на изследваните биомаркери с туморната диференциация и годините на изследваните пациентите. Значимо повишение нивата на експресия при прогресия на туморния стадий се наблюдава за: miR-4689 ($p = 0.050$), miR-101-3p ($p = 0.015$), miR-140-3p ($p = 0.024$), miR-144-3p ($p = 0.032$), miR-6785-5p ($p = 0.023$), miR-21-3p ($p = 0.049$), miR-375 ($p = 0.040$) и miR-139-5p ($p = 0.046$). Нивата на miR-95-3p са значително намалени при пациенти, негативни за лимфни ($p = 0.022$) и далечни метастази ($p = 0.013$), докато при пациентите с нодални и далечни метастази, експресията на miR-95-3p е повишена. Нивата на miR-744-3p намаляват с наличието на метастази ($p = 0.048$). Двете рамена на miR-195 се асоциират с различни клиничко-патологични характеристики. Експресията на miR-195-3p е повишена при прогресия на туморния статус ($p = 0.022$), а водещото рамо miR-195-5p е с повишена експресия при пациентите, положителни за метастази ($p = 0.018$). Според морфологията на ПК експресията на miR-195-5p е по-висока при туморите, при които не се наблюдава кератинизация, докато нивото на miR-195-3p е по-висока при туморите, характеризирани се с кератинизация. За ПК с кератинизация са характерни и по-високи нива на експресия на miR-21-3p ($p = 0.015$), miR-7-5p ($p = 0.034$). Въпреки че експресията на miR-375 ($p = 0.144$) не достига до значимост, при нея също така се наблюдава повишена експресия при туморите с кератинизация. При ПК без кератинизация се наблюдават по-високи нива на експресия при miR-144-3p ($p = 0.004$), miR-744-3p ($p = 0.046$), miR-101-3p ($p = 0.001$) и miR-95-3p ($p = 0.005$). Туморите локализиращи в десния белодробен лоб показват по-висока експресия на miR-95-3p ($p = 0.003$) и miR-195-5p ($p = 0.035$). При разпределението по пол се вижда по-висока експресия на miR-650 ($p = 0.019$) и miR-195-3p ($p = 0.050$) при жените. Въпреки че експресията на miR-744-3p ($p = 0.152$) не достига до статистическа значимост по-високи нива се са характерни при пациентите от мъжки пол.



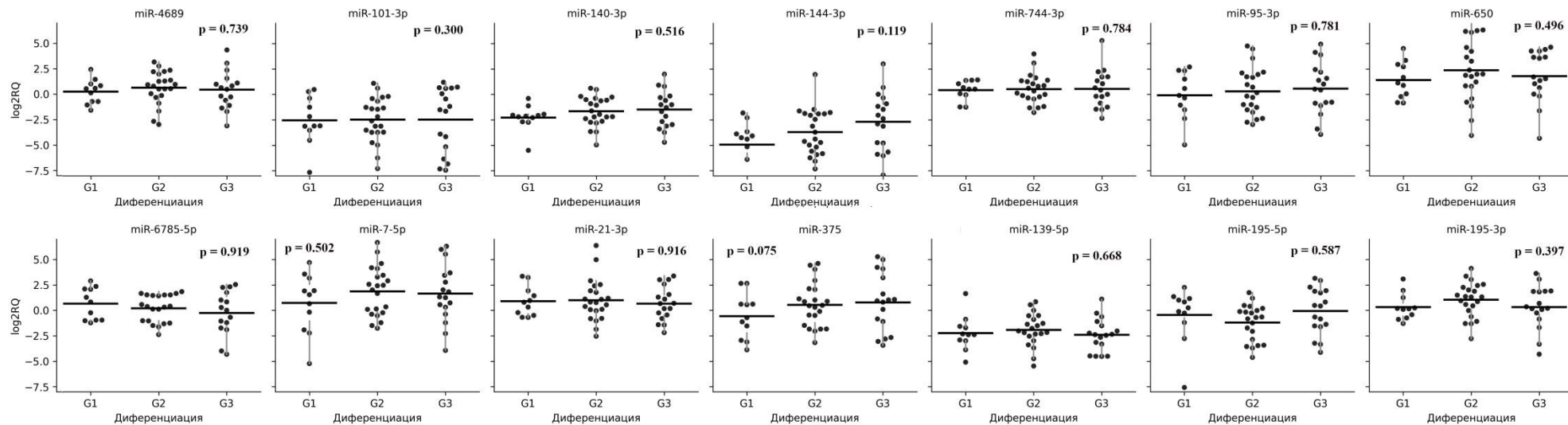
Фигура 4. 22. Разпределение на броя пациенти спрямо туморния стадий, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



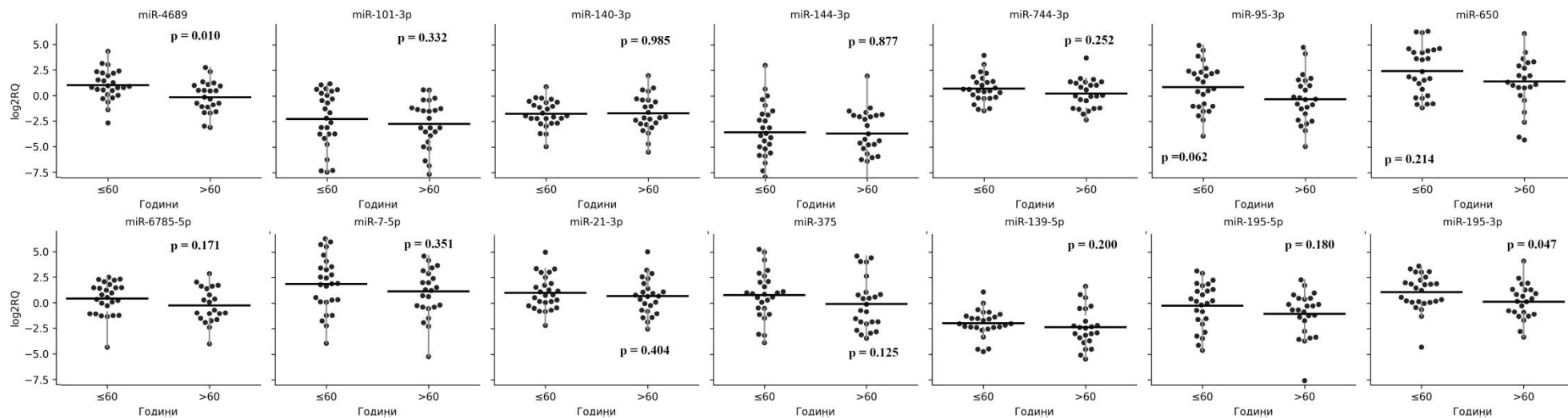
Фигура 4. 23. Разпределение на броя пациенти спрямо нодалния стадий, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



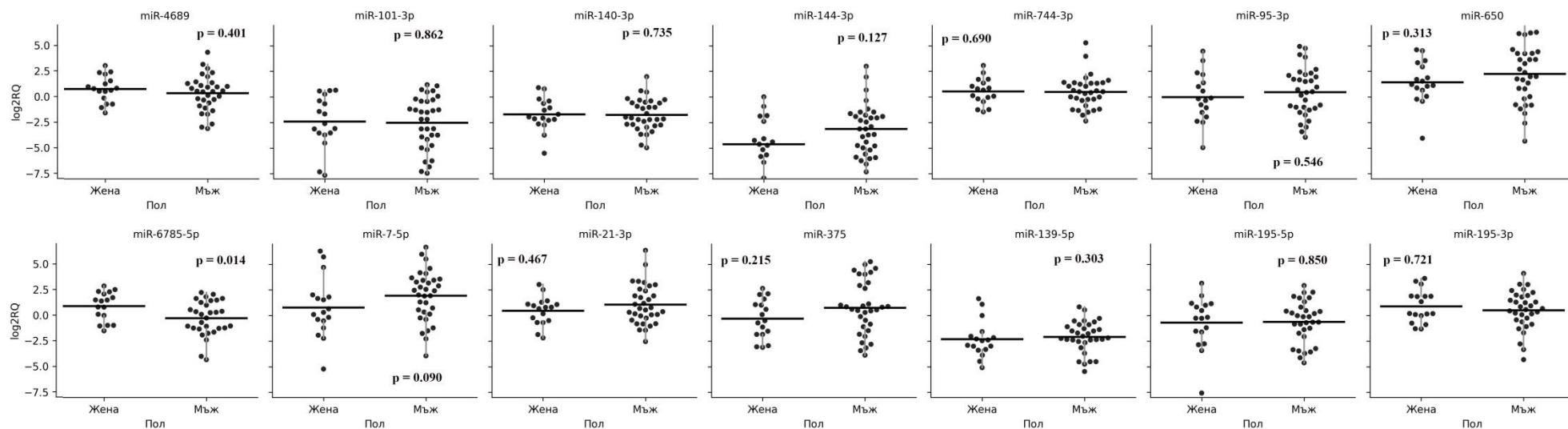
Фигура 4. 24. Разпределение на броя пациенти спрямо наличието/отсъствието на далечни метастази, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



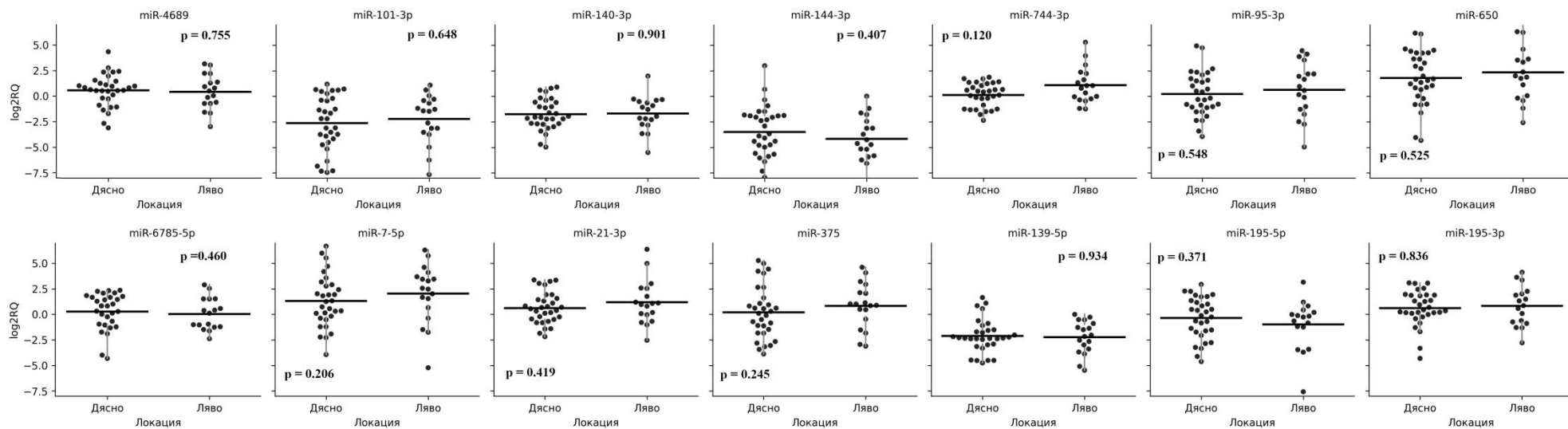
Фигура 4. 25. Разпределение на броя пациенти спрямо диференциацията на туморите, експресията на микроРНКите и резултатите от проведения статистически тест (Kruskal Wallis; $p \geq 0.05$ е незначимо) при АК.



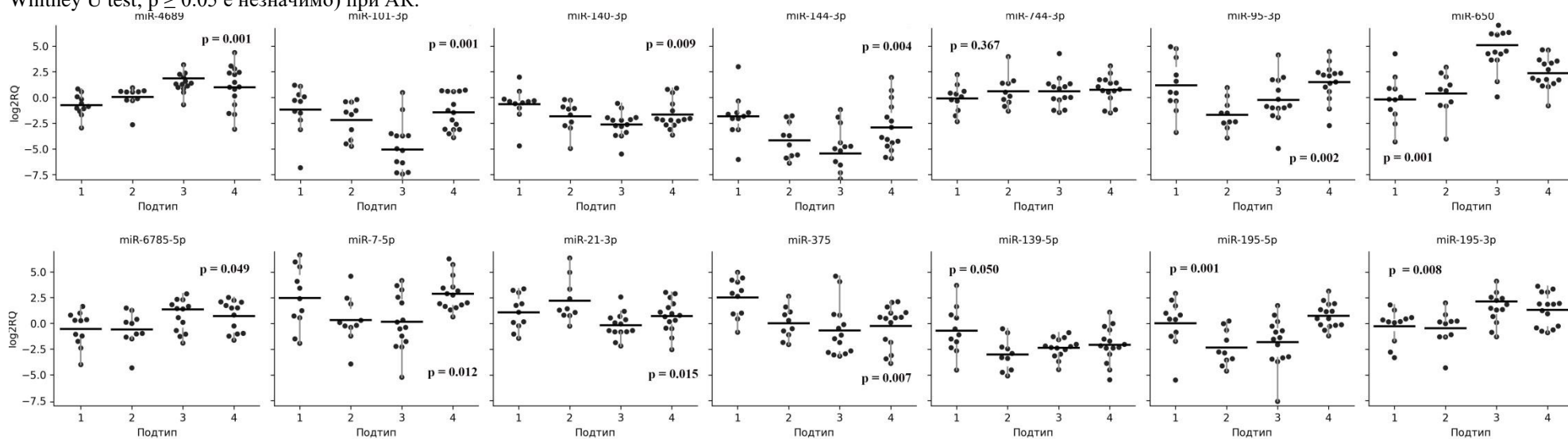
Фигура 4. 26. Разпределение на броя пациенти спрямо години, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



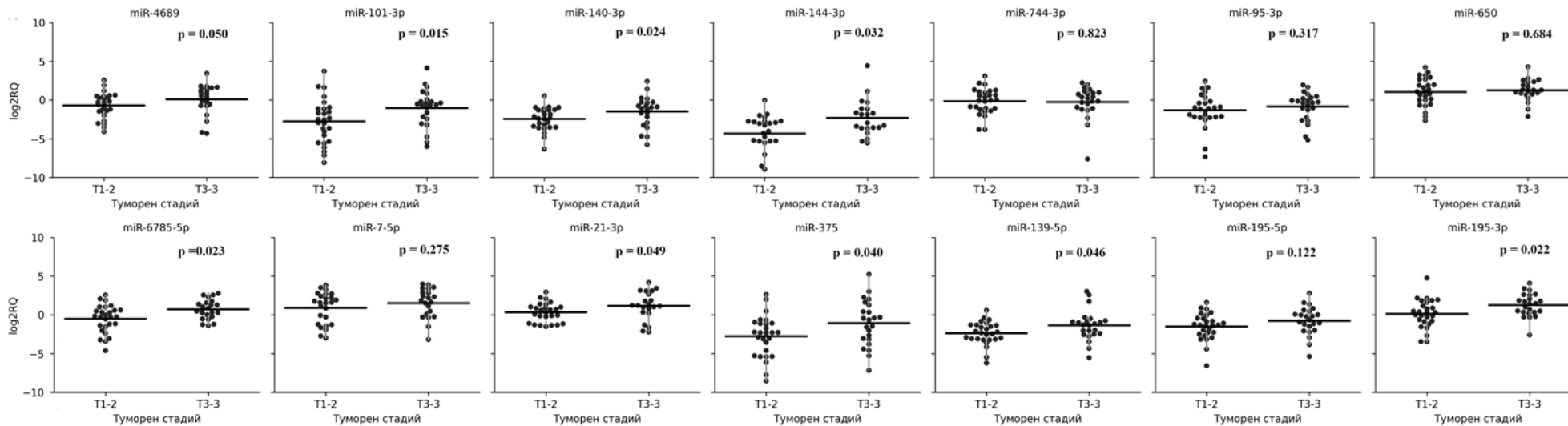
Фигура 4. 27. Разпределение на броя пациенти спрямо пол, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



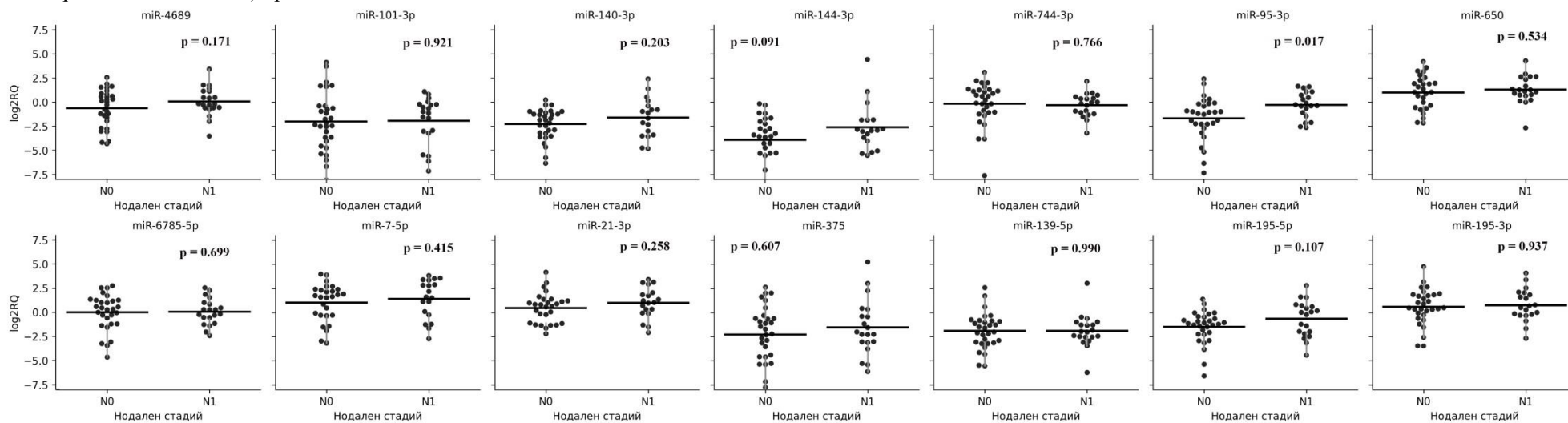
Фигура 4. 28. Разпределение на броя пациенти спрямо локализацията на тумора, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



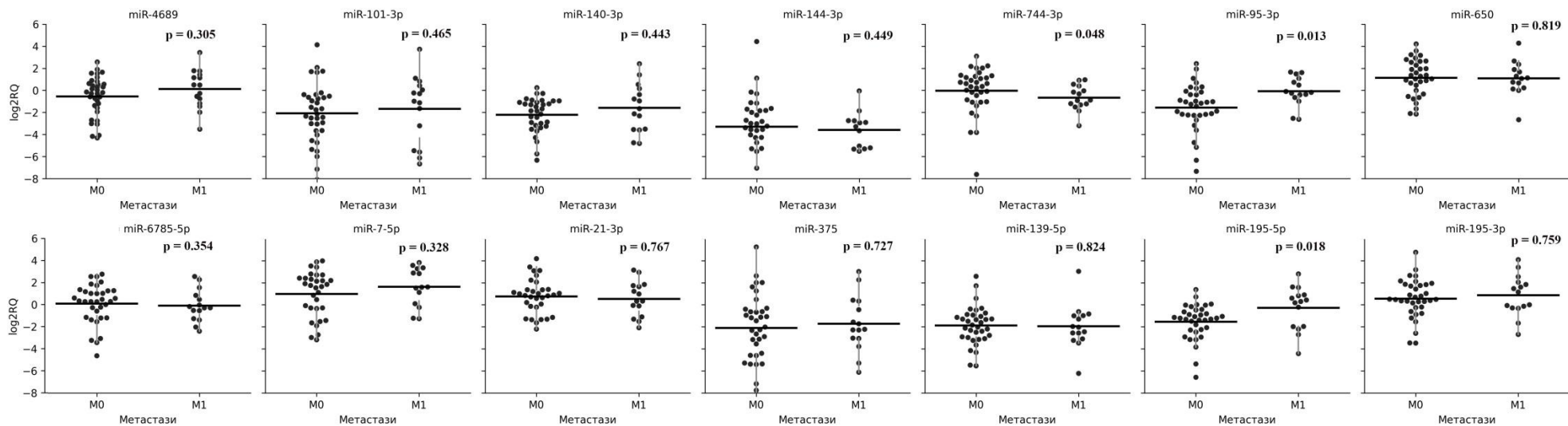
Фигура 4. 29. Разпределение на броя пациенти спрямо морфолгията на АК, експресията на микроРНКите и резултати от проведения статистически тест (Kruskal Wallis; $p \geq 0.05$ е незначимо). Всички карциноми са смесен тип, като при 1 – преобладава солидна компонента; 2 – преобладава микропапиларна компонента; 3 – преобладава ацинарна компонента и 4 – преобладаваща папиларна компонента.



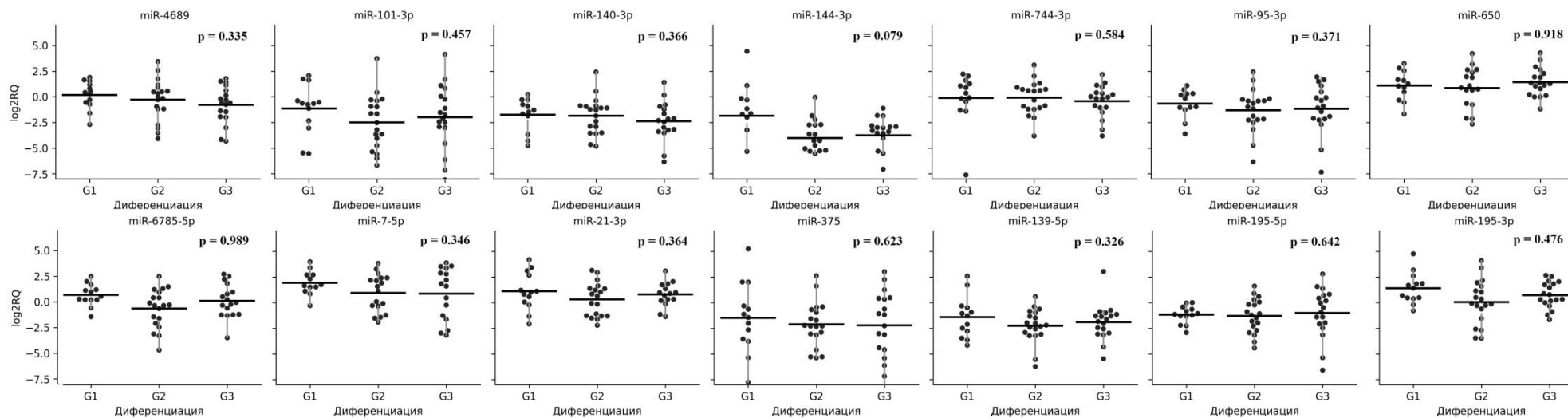
Фигура 4.30. Разпределение на броя пациенти спрямо туморния стадий, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при ПК.



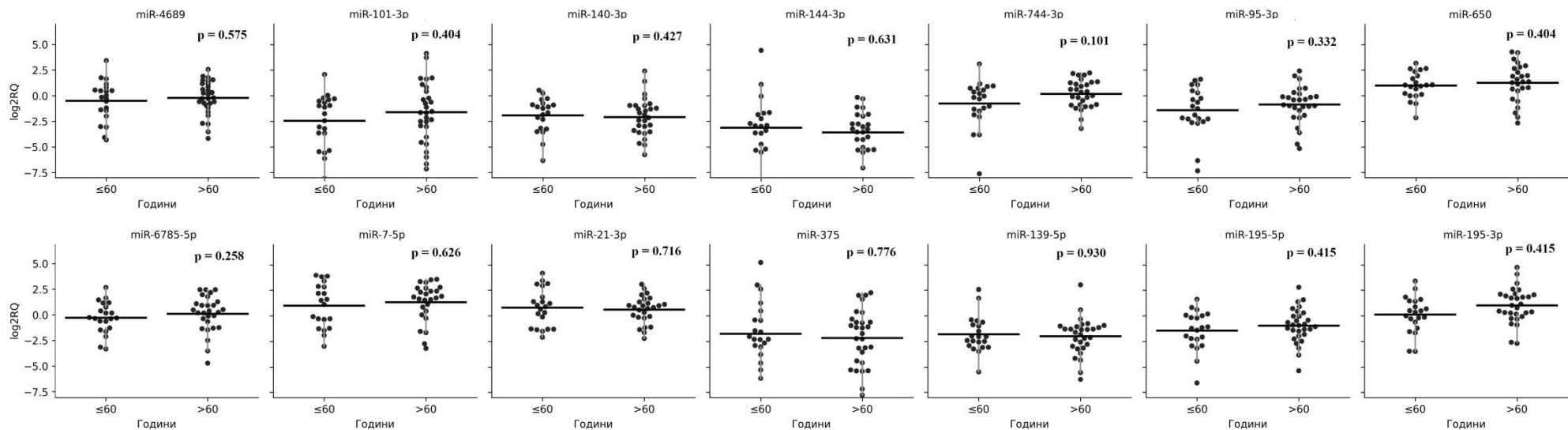
Фигура 4.31. Разпределение на броя пациенти спрямо нодалния стадий, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при ПК.



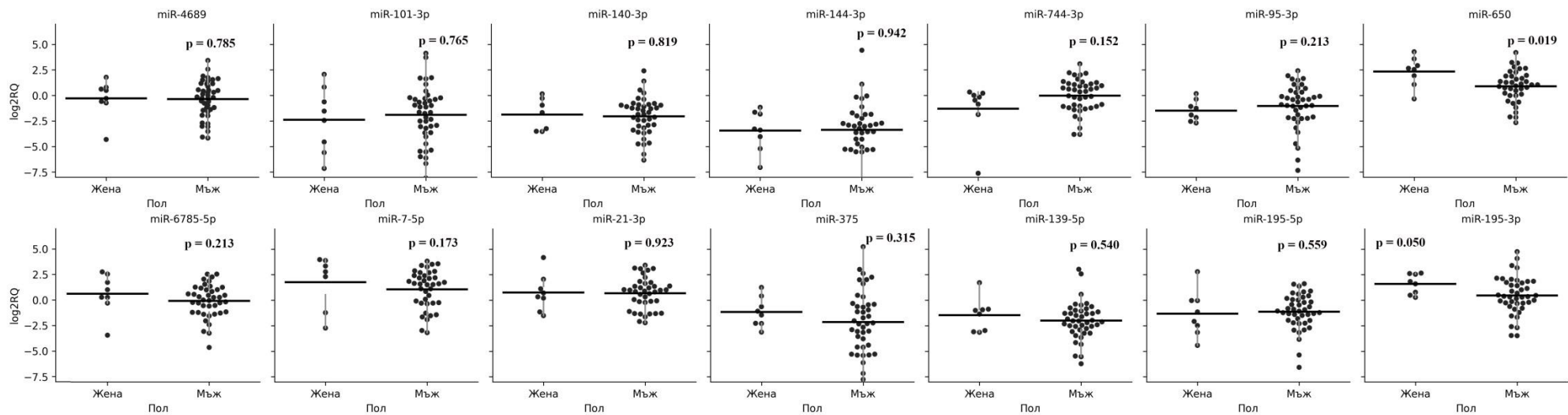
Фигура 4. 32. Разпределение на броя пациенти спрямо наличието/отсъствието на далечни метастази, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при ПК.



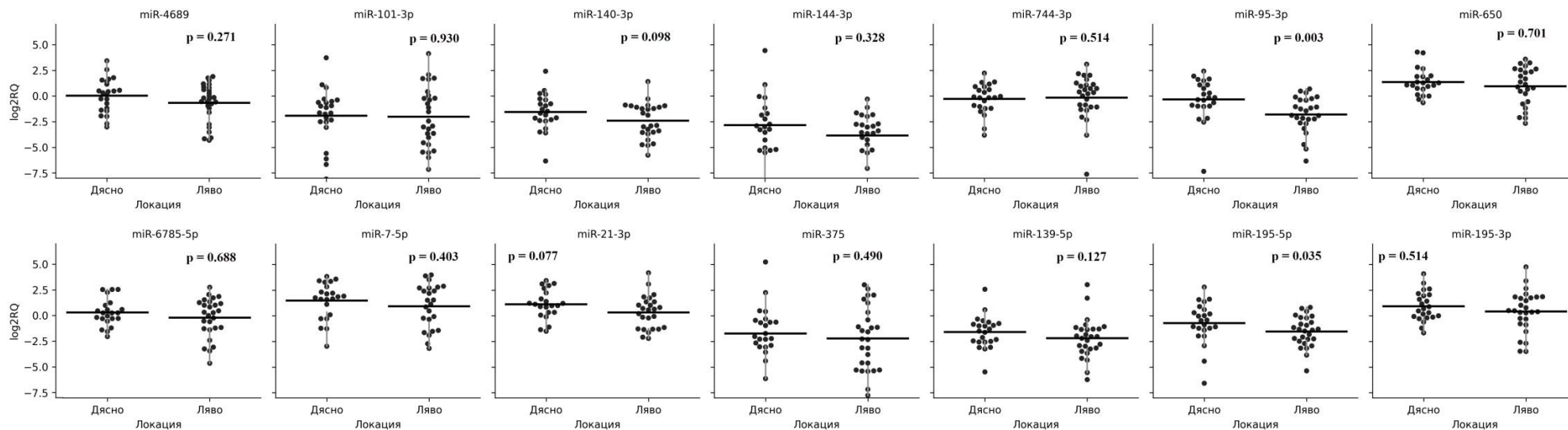
Фигура 4. 33. Разпределение на броя пациенти спрямо диференциацията на туморите, експресията на микроРНКите и резултатите от проведения статистически тест (Kruskal Wallis; $p \geq 0.05$ е незначимо) при ПК.



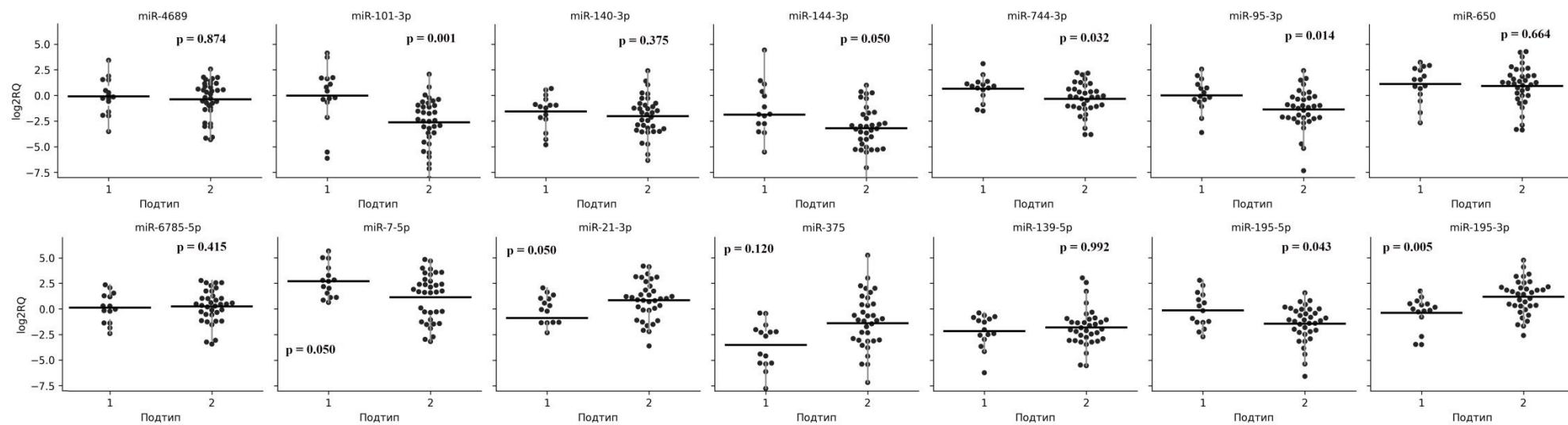
Фигура 4. 34. Разпределение на броя пациенти спрямо години, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при ПК.



Фигура 4. 35. Разпределение на броя пациенти спрямо пола, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при ПК.



Фигура 4.36. Разпределение на броя пациенти спрямо локацията на тумора, експресията на микроРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при ПК.



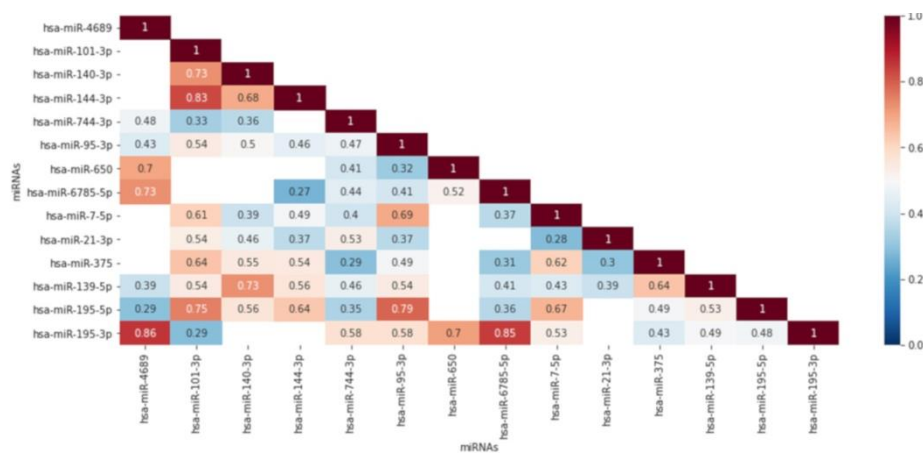
Фигура 4.37. Разпределение на броя пациенти спрямо морфологията на ПК, експресията на микроРНКите и резултати от проведения статистически тест (Kruskal Wallis; $p \geq 0.05$ е незначимо). Всички карциноми са смесен тип, като при 1 – без кератинизация; 2 – с кератинизация.

4.5. Корелационен анализ между изследваните микроРНКи.

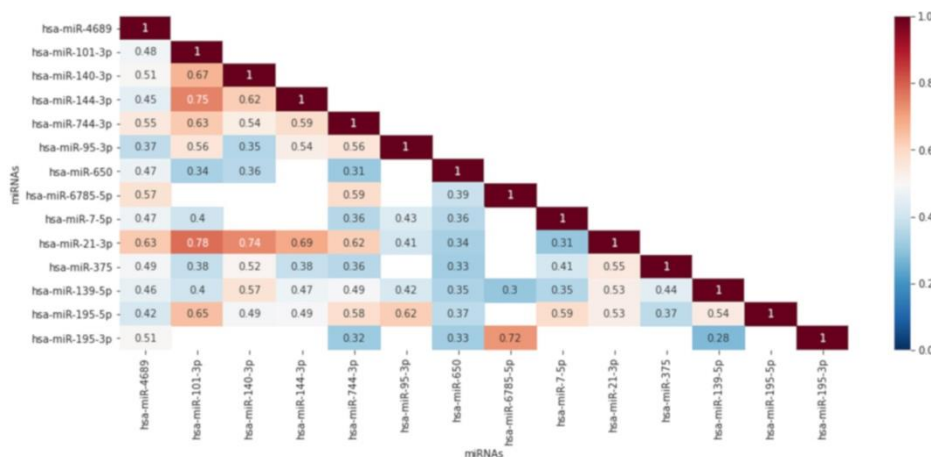
С цел да се анализира дали са свързани и каква е силата на зависимост между нивата на относителна експресия на микроРНКи, в изследваната група пациенти при АК и ПК, се използва корелационен анализ на Spearman. Анализът също ни позволява да направим предположение дали една микроРНКа зависи от една или повече други микроРНКи.

На следващата Фигура 4.38 и Фигура 4.39 са представени съответно матрица с корелационните коефициенти на всички възможни комбинации между микроРНКите. Наблюдаваха се само положителни корелации. В червената скала са отбелязани силните положителните корелации, докато в синята скала са отбелязани по-слабите корелации. Визуализирани са само статистически значимите корелации с $p \leq 0.05$.

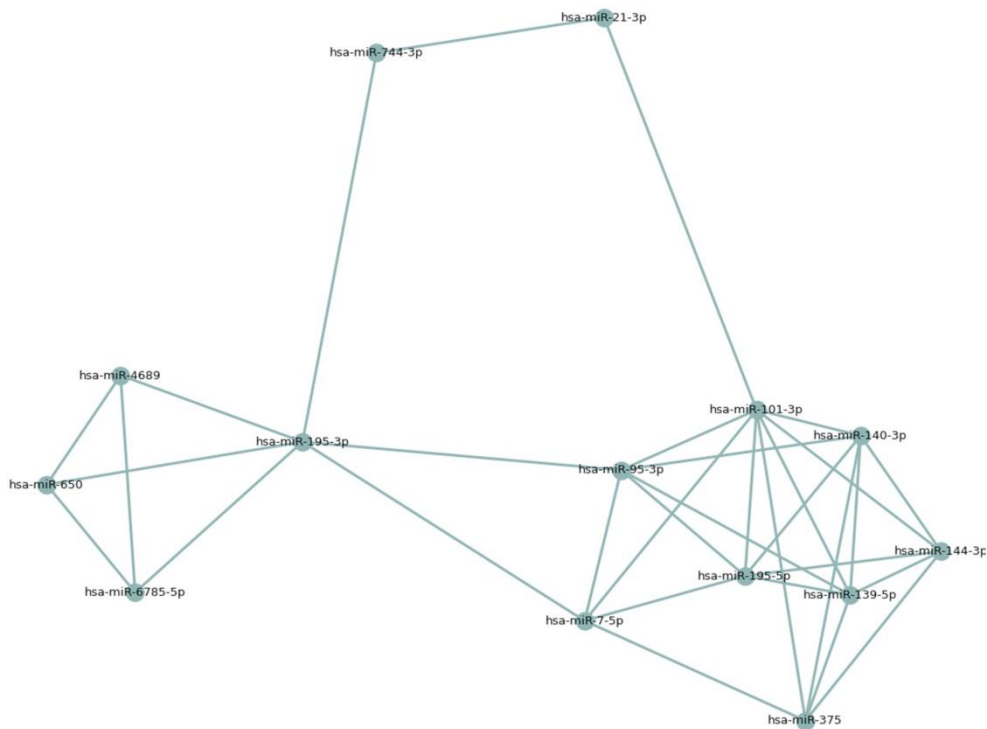
За да се стесни кръгът до най-истинските и най-значими взаимовръзки между изследваните микроРНКи и да се избегнат фалшиво положително корелационни взаимовръзки се проведе допълнителна корекция по Benjamini - Hochberg FDR (BH-FDR). След корекцията, връзките с коефициент на корелация по-голям от 0.500 послужиха за послужиха за построяването на корелационно – базирани мрежи. Мрежите са представени на Фигура 4.40. за АК и Фигура 4.41. за ПК.



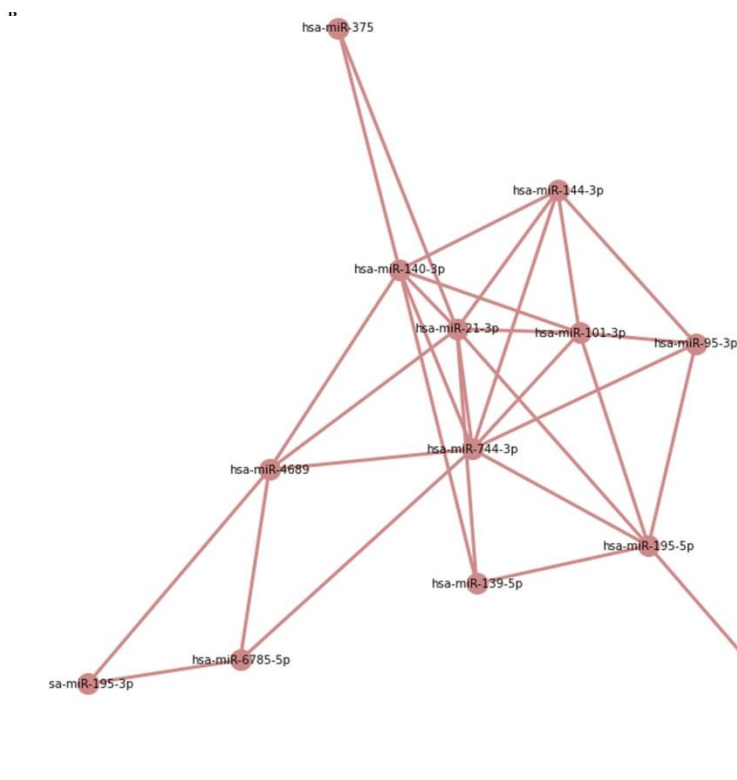
Фигура 4.38. Матрица с корелационните коефициенти на възможните комбинации между микроРНКи при АК. Визуализирани са само статистически значимите корелации с $p \leq 0.05$, без приложена корекция.



Фигура 4.39. Матрица с корелационните коефициенти на възможните комбинации между микроРНКи при ПК. Визуализирани са само статистически значимите корелации с $p \leq 0.05$, без приложена корекция.



Фигура 4.40. Корелационно – базирана мрежа със статистически значимите корелационни връзки между маркерите при АК. С добавена корекция (BH-FDR) и коефициент на корелация ≥ 0.500 .



Фигура 4.41. Корелационно – базирана мрежа със статистически значимите корелационните връзки между маркерите при ПК. С добавена корекция (BH-FDR) и коефициент на корелация ≥ 0.500 .

Ако две мироРНКите взаимодействат помежду си в мрежа, е по-вероятно те да регулират пътищата и целеви гени с подобни функции (218–220). Следователно,

изследването на взаимодействията между микроРНКите в мрежовите модели може да предостави значителни насоки и да даде обяснение на сложните регулаторни механизми при лечението на рака на белия дроб (221). Степента на централност спомага да идентифицираме най-важните възли в мрежата по отношение на броя на взаимодействията. Тенденцията на мрежата да образува кълъстери може да се измери чрез средния коефициент на кълъстериране. Коефициентът на кълъстериране оценява свързаността на всеки възел: възелът има висок коефициент на кълъстериране, ако неговите съседи са склонни да бъдат директно свързани помежду си.

При образуването на графа при АК микроРНКите са представени като възли (върхове), а като дъги (ребра) са представени коралационно-базираните връзки. Броят на възлите е 14, а на дъгите е 33. Средната степен на неориентирания граф е сумата от степените на всички негови възли, разделена на броя на възлите в графа и е мярка за свързаността на графика. При АК стойността на средната степен е 4.7143. Ясно се формираха два кълъстера, свързани чрез miR-21-3p, miR-744-3p и miR-195-3p. В образуването на по-малкия Кълъстер 1 участват четири микроРНКи. Това са miR-195-3p, miR-4689, miR-650 и miR-6785-5p. От тях miR-195-3p има шест коралационно-базираните връзки. Свързва се с трите микроРНКи от Кълъстер 1, с miR-744-3p и с микроРНКите miR-7-5p и miR-95-3p, които са част от Кълъстер 2. Останалите микроРНКи от Кълъстер 1 имат по три коралационно-базираните връзки и се свързват помежду си. miR-101-3p има 8 корелационно-базирани връзки, следвана от miR-139-5p, miR-140-3p, miR-195-5p, miR-195-3p и miR-95-3p с 6 корелационно-базирани връзки. На Таблица 4.8 са представени характеристиките на мрежата при АК.

Таблица 4. 8. Характеристики на мрежата от микроРНКи при АК.

микроРНК	Брой връзки	Степен на централност	Коефициент на кълъстериране
miR-4689	3	0.0	1.0
miR-650	3	0.0	1.0
miR-6785-5p	3	0.0	1.0
miR-195-3p	6	0.438	0.267
miR-21-3p	2	0.05	0
miR-744-3p	2	0.051	0
miR-95-3p	6	0.206	0.6
miR-140-3p	6	0.016	0.8
miR-101-3p	8	0.160	0.536
miR-7-5p	5	0.141	0.5
miR-144-3p	5	0.002	0.9
miR-375	5	0.018	0.7
miR-139-5p	6	0.016	0.8
miR-195-5p	6	0.029	0.733

При образуването на графа при ПК възлите (микроРНКите) са 13, а броят на дъгите е 30. Средната степен на графа е 4.6154. При ПК нямаме ясно се кълъстериране както при АК, но за разлика от АК miR-21-3p и miR-744-3p имат най-много корелационно-базирани връзки – осем. Следвани miR-140-3p с 7 корелационно-базирани връзки. miR-650 не образува силни връзки с нито една микроРНКа. На Таблица 4.9 са представени характеристиките на мрежата при ПК.

Таблица 4. 9. Характеристики на мрежата от микроРНКи при ПК.

микроРНК	Брой връзки	Степен на централност	Коефициент на кълъстиране
miR-4689	5	0.147	0.5
miR-650	-	-	-
miR-6785-5p	5	0.033	0.666
miR-195-3p	2	0.0	1.0
miR-21-3p	8	0.189	0.464
miR-744-3p	8	0.227	0.5
miR-95-3p	4	0.008	0.833
miR-140-3p	7	0.116	0.476
miR-101-3p	6	0.028	0.733
miR-7-5p	1	0.0	0
miR-144-3p	5	0.013	0.8
miR-375	2	0.0	1.0
miR-139-5p	3	0.007	0.666
miR-195-5p	6	0.200	0.498

4.6. Анализ на гените – мишени на микроРНКите.

Установяването на таргетните гени регулирани от диференциално експресирани микроРНК-и, както и клетъчните пътища, в които участват бе осъществено с помощта на multiMir и Enrichr. На Таблица 4.10 са представен броя на неповтарящите се в 11 бази данни гени-мишени на всяка от изследваните микроРНКи.

Таблица 4.10. Броят на гените-мишени на всяка от изследваните микроРНКи

miRNA	Брой валидирани генни-мишени*	Брой предсказани генни-мишени*	Брой валидирани и предсказани генни-мишени*
miR-4689	127	1884	1827
miR-101-3p	2984	2547	3449
miR-140-3p	768	2022	2380
miR-144-3p	422	3183	3071
miR-744-3p	546	1529	1843
miR-95-3p	182	823	915
miR-650	509	7890	7729
miR-6785-5p	825	3000	2692
miR-7-5p	3403	4242	5121
miR-21-3p	1935	1628	2999
miR-375	1410	2200	2818
miR-139-5p	786	2418	2542
miR-195-5p	3158	4144	5925
miR-195-3p	117	2277	2268

*представен е броят на не повтарящите се гени от 11 бази данни

4.7. Онтологичен анализ.

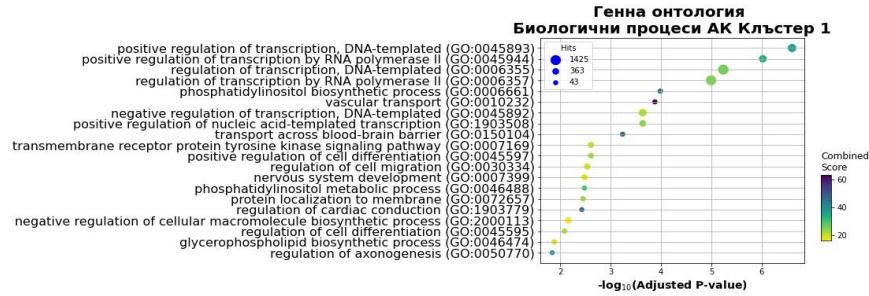
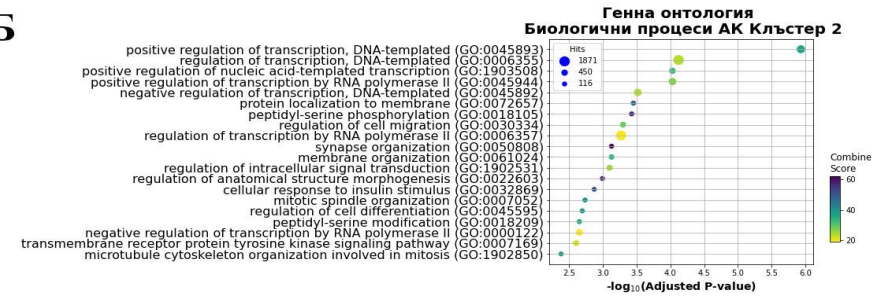
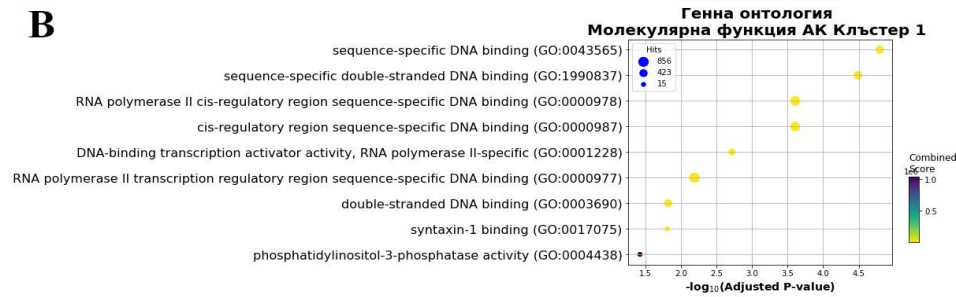
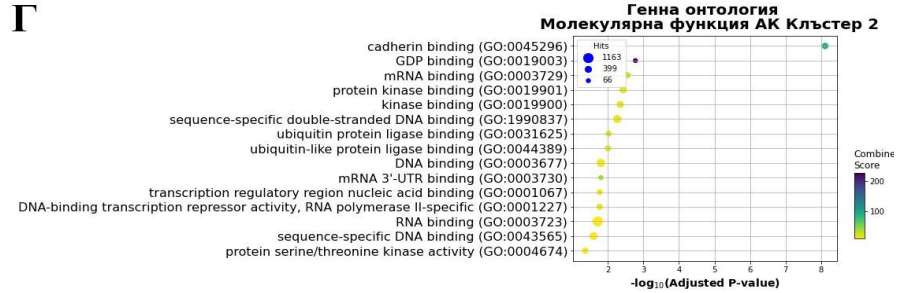
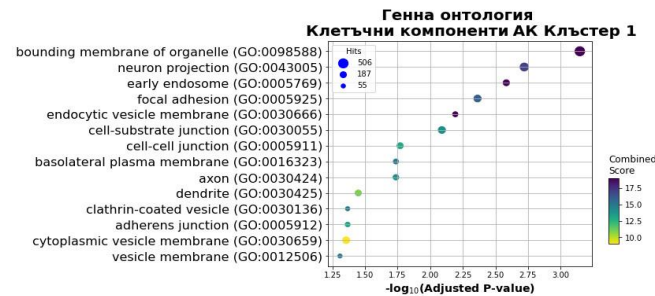
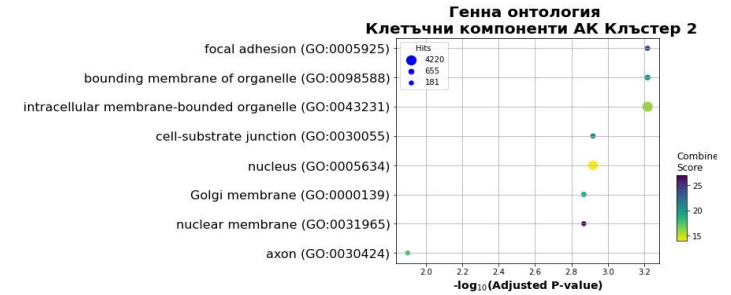
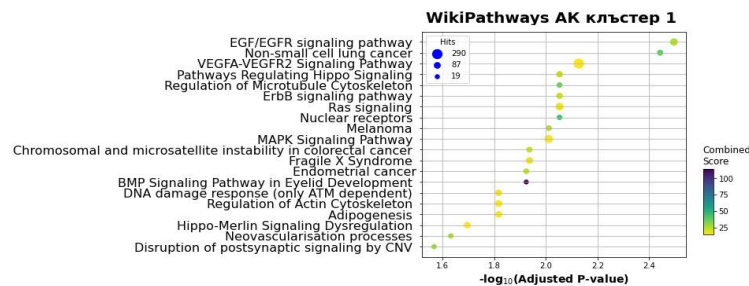
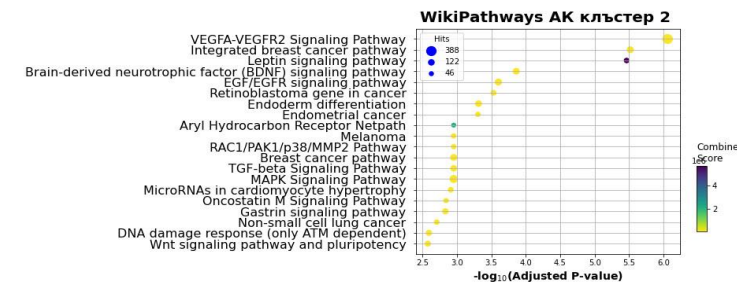
За да определим разпределението на генните-мишени проведохме онтологичен анализ и анализ на най-засегнатите сигнални пътища. При АК извършихме двата

анализа за Клъстер 1 и Клъстер 2 като използвахме валидираните и предсказаните генни-мишени. За ПК в анализите участваха валидираните и предсказаните генни-мишени на микроРНКите, характеризиращи се с най-много корелационно-базирани връзки и заемащи най-централно значение при образуването на мрежата - miR-21-3p, miR-744-3p, miR-140-3p и miR-195-5p.

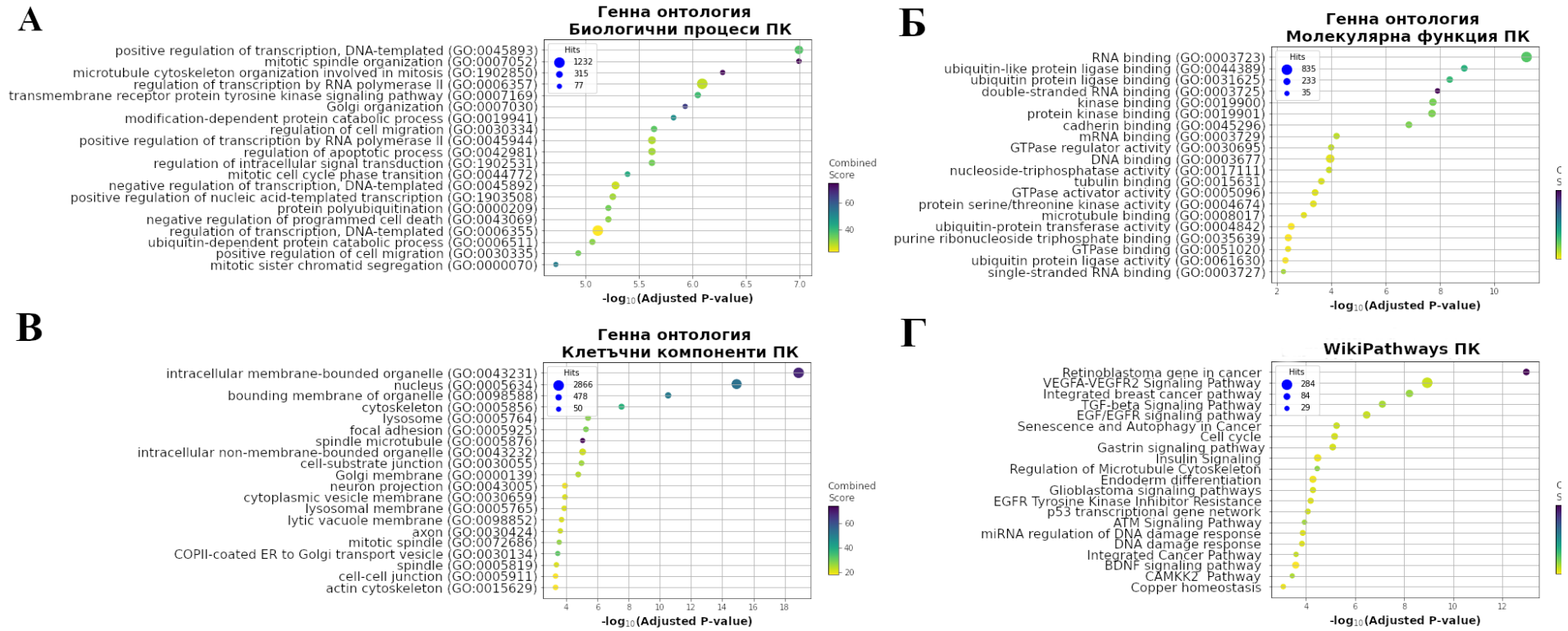
При АК Клъстер 1 валидирани, чрез различни функционални изследвания и предсказани, чрез различни алгоритми, генни-мишени на четирите микроРНКи са 11560 гена. Според онгологичния анализ генните-мишени спадат към регулация на транскрипцията и транслацията, фосфоинозитол и глицерофосфолипидните биосинтетичните процеси, везикуларен транспорт, транспорт през кръвно-мозъчната бариера, регулация на клетъчната миграция и диференциацията. Повечето белтъците са транскрипционни фактори или цис-регулатори на транскрипцията (Фигура 4.42А,В,Д). Според WikiPathway 2021 тези гени, след корекция на р-стойността, обогатяват значимо 30 пътя. На Фигура 4.42Ж са представени двайсетте най-засегнати сигнални пътя. В тях обогатени са 910 генни-мишени.

При АК Клъстер 2 валидирани и предсказани генни-мишени на осемте микроРНКи са 15808 гена. Според WikiPathway 2021, след корекция на р-стойността, тези гени обогатяват значимо 93 пътя. На Фигура 4.42З са представени двайсетте най-засегнати сигнални пътя. В тях обогатени са 1231 генни-мишени. Според онгологичния анализ мишените на изследваните микроРНКи се отнасят към регулация на транскрипцията и транслацията, пептидил-серин модификациите, организация на делителното вретено, белтъците по повърхността на клетъчната мембрана, регулация на клетъчната диференциация и морфогенеза. Повечето белтъци са кадхерин свързващи, G-свързващи регулаторни белтъци, протеин кинази, убиквитин лигази, транскрипционни фактори и белтъци, участващи в биогенезата на различни мРНКи(Фигура 4.42Б,Г,Е).

При ПК валидирани и предсказани генни-мишени на четирите микроРНКи са 9889 гена. Според онгологичния анализ генните-мишени на четирите микроРНКи с най-много корелационно базирани връзки са белтъци, участващи в деленето на клетката, регулация на клетъчната миграция, регулатори на генната експресия, апоптозата и различни модификатори. Повечето белтъци се свързват с РНК, кадхерините, участват в биогенезата на мРНКи, клетъчния транспорт, биват убиквитин лигази или се свързват с пуринов нуклеотид (Фигура 4.43А,Б,В). Според WikiPathway 2021, след корекция на р-стойността, тези гени обогатяват значително 57 пътя. На Фигура 4.43Г са представени двайсетте най-засегнати сигнални пътя. В тях обогатени са 915 генни-мишени.

А**Б****В****Г****Д****Е****Ж****З**

Фигура 4.42. Онтологичен анализ и анализ на сигналните пътища при АК. А, В, Д и Ж резултати за Клъстер 1; Б, Г, Е, З резултати за Клъстер 2



Фигура 4.43. Онтологичен анализ и анализ на сигналните пътища при ПК.

4.8. Най-силно засегнати и важни за раковите клетки пътища повлияни от показалите диференциална експресия микроРНК-и, регулиращи участници в тях.

Нанахан и съавтори характеризират основните свойства на раковите клетки. Това са придобити способности за поддържане на пролиферацията, потискане на супресори на растежа, потискане на клетъчната смърт, репликативно безсмъртие, индуциране/достъп до кръвоносни съдове, инвазия и метастазирание, препрограмиране на клетъчния метаболизъм и избягване на имунния отговор (222,223). Базирайки се на тези характеристики, сред 20 най-засегнати пътя, избрахме основните клетъчни пътища и процеси, променени при АК и ПК.

4.8.1. EGF/EGFR сигнален път.

При Клъстер 1 на АК генните-мишени, участващи в EGF/EGFR сигналния път и регулиращи се от четирите микроРНКи са 121 от 162 гена. Съотношението на шансовете е 2.167, а комбинираният резултат е 26.398. Гените *CREB1* и *RICTOR* се таргетират от три от четирите микроРНКи, които са в Клъстер 1.

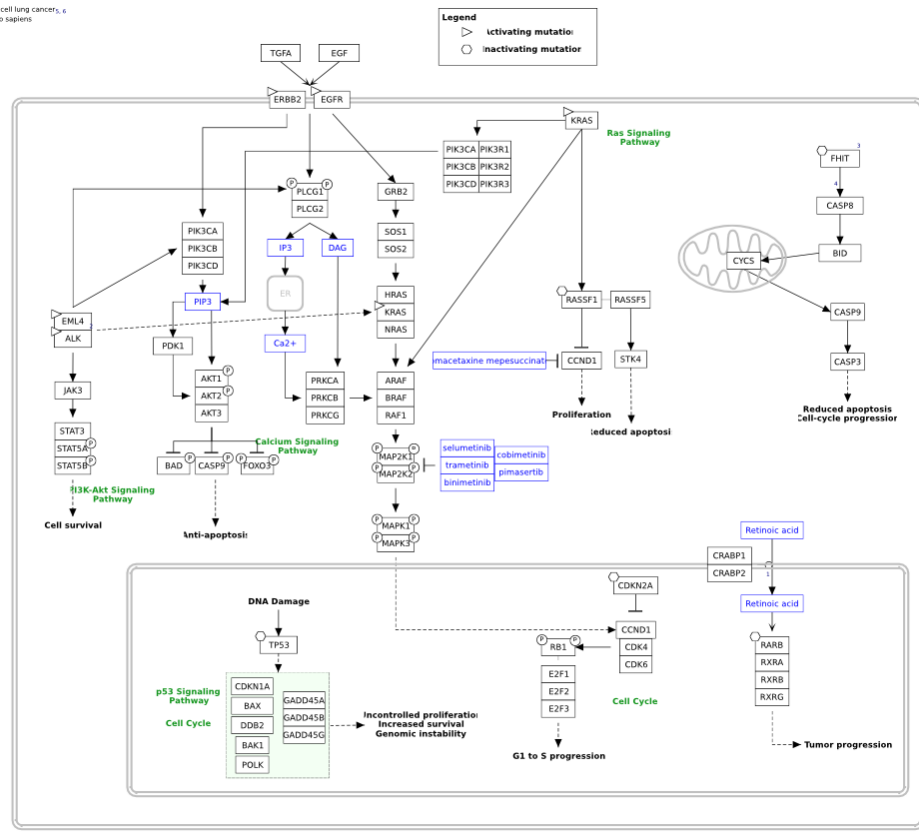
При Клъстер 2 на АК мишени на микроРНК-ите са 150 гена от 162 участника в EGF/EGFR сигналния път. Съотношението на шансовете е 3.336, а комбинираният резултат е 43.777. Гените *CREB1*, *JUN* и *RICTOR* се таргетират от шест от осемте микроРНКи в Клъстер 2.

При ПК мишени на четирите микроРНК-ите са 117 гена от 162 гена, участващи в EGF/EGFR сигналния път. Съотношението на шансовете е 2.678, а комбинираният резултат е 52.828. Гените *PTEN*, *RPS6KA5* и *RICTOR* се регулират от четирите микроРНКи.

4.8.2. Сигнален път при НДРБД.

При АК мишени, участващи в НДРБД сигналния път и регулиращи се от четирите микроРНКи са 59 от 72 гена. Съотношението на шансовете е 3.326, а комбинираният резултат е 37.795. Гените *CDKN1A*, *E2F2*, *KRAS*, *MAPK1*, *PDK1*, *PIK3R1*, *RASSF1* и *STK4* са регулирани от микроРНК-ите, участващи във образуването на Клъстер 1.

При Клъстер 2 на АК гени мишени на микроРНК-ите са 69 гена от 72 участника в този сигнален път. Съотношението на шансовете е 6.121, а комбинираният резултат е 43.777. Гените *CDK6*, *CDKN1A*, *AKT3*, *CASP3*, *CCND1*, *E2F3*, *MAPK1*, *RARB* и *TGFA* се таргетират от най-много микроРНКи в този клъстер.



Фигура 4.44. Сигнален път при НДРБД (195).

4.8.3. VEGFA-VEGFR2 сигнален път.

Най - много микроРНК-и от Клъстер 1 на АК регулират гените *CREB1*, *MAPKAPK5*, *QKI*, *RICTOR*, *ACKR3*, *ALDOA*, *BCL2L1*, *CALR*, *FMNL3*, *GAB1*, *HDAC9*, *MAPK1*, *MDM2* и *CDH5*. Общо се таргетираат 290 от 432 гена като съотношението на шансовете е 1.504, а комбинираният резултат е 15.389.

Най - много микроРНК-и от Клъстер 2 при АК регулират гените *QKI*, *ADAM10*, *BCL2*, *CREB1*, *FMNL3*, *JUN*, *PRRC2C*, *RICTOR*, *TXNIP*, *ARF6*, *CCND1*, *CDC42*, *EGR3*, *FOXO1*, *MAP3K5* и *VEGFA*. Общо се таргетираат 388 от 432 гена като съотношението на шансовете е 2.372, а комбинираният резултат е 48.352.

При ПК се таргетираат 284 от 432 гена като съотношението на шансовете е 1.990, а комбинираният резултат е 52.412. Най-много се търгетираат гените - *RPS6KA*, *ADAM10*, *PTPN14*, *RICTOR*, *MAP2K3*, *RAB11A*, *QKI*, *JUN*, *BIRC5*, *KDR*, *VEGFA*, *FOXO3*.

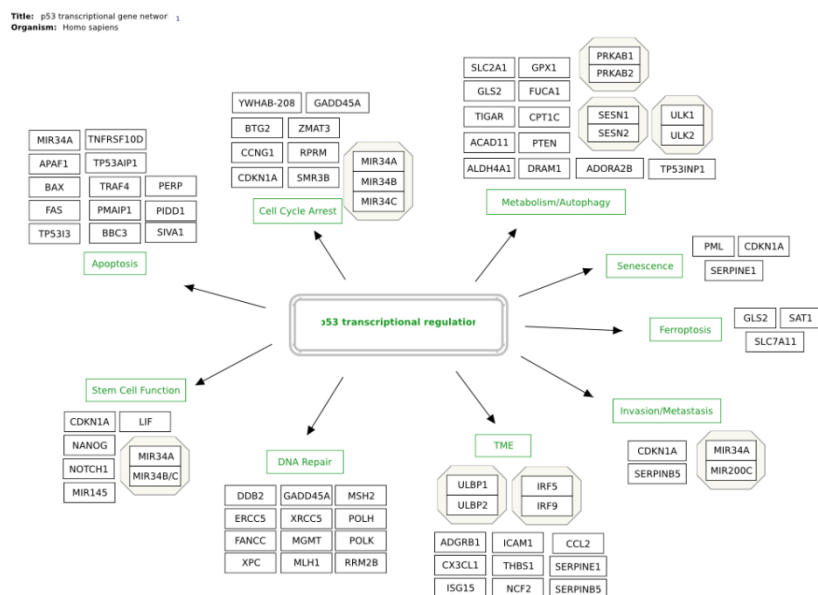
4.8.4. Сигнални пътища свързани с ДНК поправката.

Най - много микроРНК-и от Клъстер 1 на АК регулират гените *CDKN1A*, *SCP2*, *BAX*, *KRAS*, *PDK1*, *PIK3C3*, *PTEN*, *WNT1* и *WNT5A*. Общо се таргетираат 81 от 110 гена като съотношението на шансовете е 2.047, а комбинираният резултат е 16.032.

Най - много микроРНК-и от Клъстер 2 при АК регулират гените *BCL2*, *CREB1*, *HIPK2*, *HMGCR*, *IRS1*, *JUN*, *PKIA*, *VEGFA*, *BMPR2*, *CASP3*, *CCND1*, *CDC42*, *FOXO1*, *MAPK1*, *NF1*, *SMARCA4*, *TGFBR1*, *WEE1*, *CTNBN1*, *MTOR*, *NAB1*, *SP1*, *BRCA1*, *CHEK1*,

MAX, MDM2, MSH6, MYC, PML, PTEN, SMAD1, SMAD2, SMAD7, EGFR, KRAS, MSH2, RB11, STK11, BCL2, CDKN1A, IRS1, JUN, AKT3, MAP3K4, MAPK1, MAPK8, PIK3CA, RHOA, APC, BBC3, GSK3B, PIK3C2A, SOD2, ABL1, MYC, PTEN, KRAS, TP53 и *WNT1*.

При ПК се засягат различни пътища, свързани с поправката на ДНК молекулата. Най-много се таргетират гените *PTEN, WEE1, PKIA, JUN, BCL2, CHEK1, SMAD7, VEGFA, BRCA1, AURKA, MDM2, TP53, EGFR, MTOR, APOBEC3G, KRAS, BTG2, CDKN1A, ZMAT3, SERPINE1, SAT1, PMAIP1, PML, XPC, MGMT, POLH, MLH1, MSH2, CCND1, CDK6, CCND2, CDC25A, DDB2, MCM7, MRE11, GADD45G, CDK1, CDK4, CDKN1A, ATF2*



Фигура 4.45. Механизмите, лежащи в основата на туморно-супресорната функция на p53, неговите регулатори и цели (195).

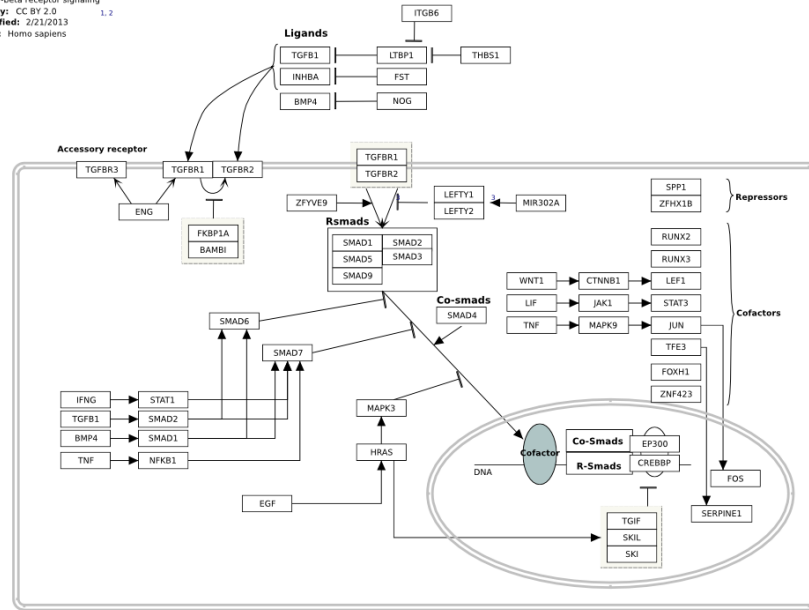
4.8.5. TGF-бета сигнален път.

В този път са засегнати общо 122 гена (Клъстер 2 АК) от общо 132 като съотношението на шансовете е 3.252, а комбинираният резултат е 34.633. Важни таргети на микроРНК-ите са: *CDKN1A, JUN, SKIL, THBS1, SKI, FOS, ITCH, KLF6, MET, RAF1, SP1, SMURF1, SMURF2* и *RUNX2*.

При ПК се засягат 76 гена в този сигнален път като съотношението на шансовете е 3.217, а комбинираният резултат е 68.930. Важни таргети на микроРНК-ите са:

PRKAR2A, CCND1, APP, CDKN1A, PAK2, MAP2K3, YAP1, SMURF2, NEDD9, ATF2, NUP153, SKIL, JUN, MAPK14, SMURF1, SIK1, JUNB и *RUNX2*.

Title: TGF-beta receptor signaling
 Availability: CC BY 2.0
 Last modified: 2/21/2013
 Organism: Homo sapiens

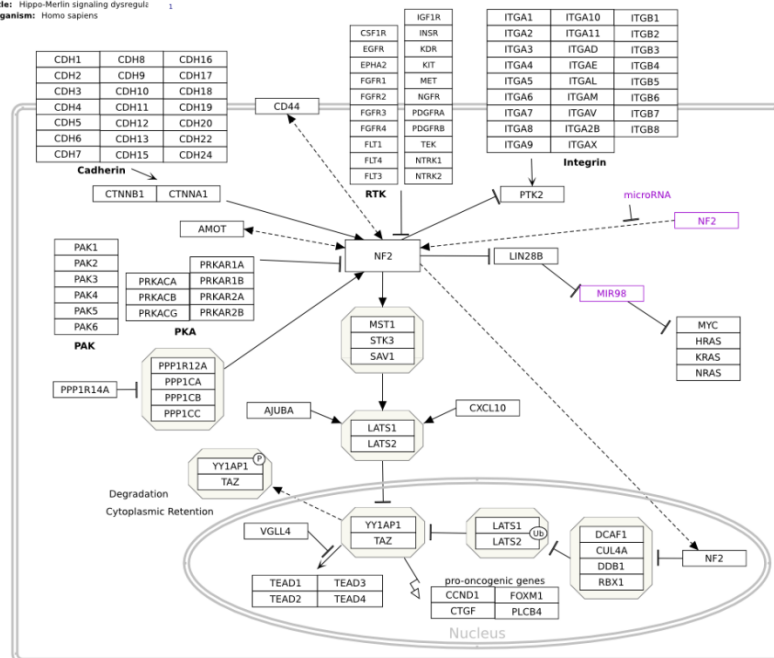


Фигура 4.46. TGF-бета сигнален път (195).

4.8.6. Hippo-Merlin сигнален път.

В този път са засегнати общо 87 гена (Клъстер 1 АК) от общо 120 като съотношението на шансовете е 1.932, а комбинираният резултат е 14.387. Важни таргети на микроРНК-ите са: *GNAI3*, *ITGA11*, *ITGA3*, *NGFR*, *PDGFRA*, *AMOT*, *CD44*, *CDH4*, *GNAS*, *KRAS*, *NF2*, *NTRK2*, *RAC1*, *ITGB3*, *PRKAA2*, *PRKAB2* и *PB4*.

Title: Hippo-Merlin signaling dysregul
 Organism: Homo sapiens

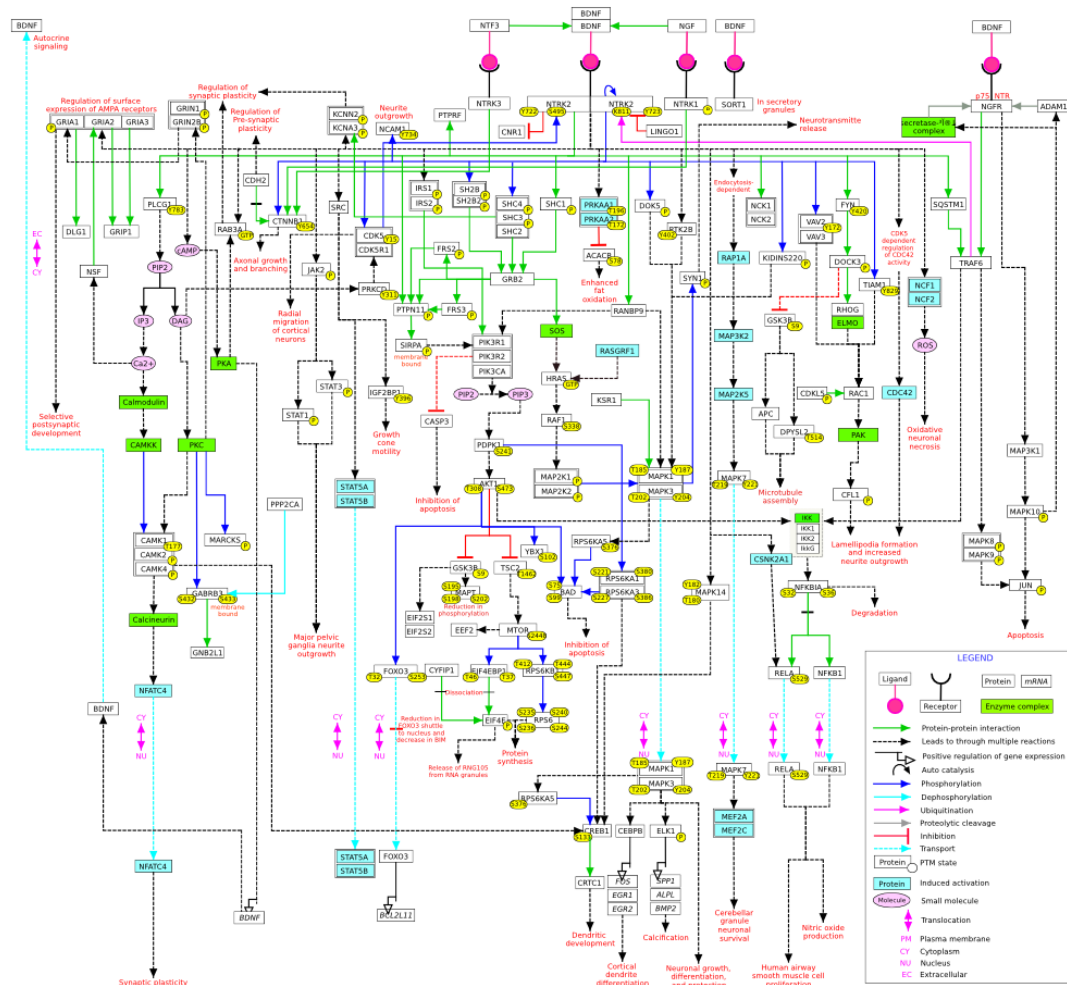


Фигура 4.47. Hippo-Merlin сигнален път (195).

4.8.7. BDNF сигнален път.

В този път са засегнати общо 135 гена от общо 152 като съотношението на шансовете е 4.003, а комбинираният резултат е 55.802. Важни засегнати гени от микроРНКите в Клъстер 2 на АК_{ca}: *CREB1*, *IRS1*, *JUN*, *RPS6KA3*, *BDNF*, *CASP3*, *CDC42*, *EIF2S1*, *MAP3K2*, *MAPK1*, *MAPK8*, *PRKAA1*, *RPS6KA5*, *APC*, *BCL2L11*, *CDK5R1*, *EIF4E*, *FOS*, *FRS2*, *GSK3B*, *IRS2*, *JAK2*, *MAP2K1*, *MEF2A*, *MTOR*, *NTRK2*, *RAF1*, *CDH2*, *FOXO3*, *KSR1*, *STAT3*, *IGF2BP1* и *PTPN11*.

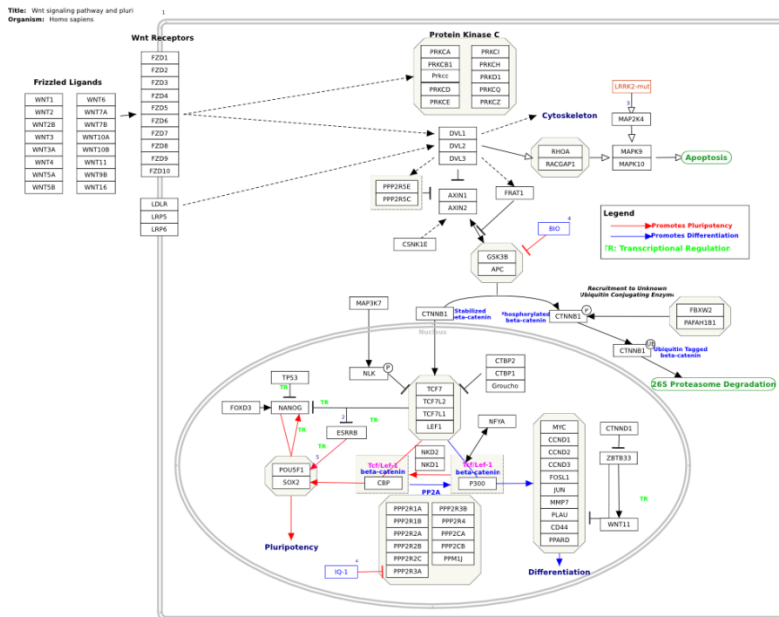
При ПК таргети на някои от четирите микроРНКи са 97 гена. Съотношението на шансовете е 2.121, а комбинираният резултат е 24.462. Важни засегнати гени са: *RPS6KA5*, *JUN*, *SIRPA*, *MARCKS*, *MAPK14*, *MAPK8*, *STAT3*, *NFKBIA*, *FOXO3*, *NTRK3*, *BDNF*, *GRB2*, *GSK3B*, *TSC2*, *MAPK10*, *STAT1* и *DOCK3*.



Фигура 4.48. BDNF сигнален път (195).

4.8.8. Wnt сигнален път.

МикроРНКите от Клъстер 2 АК контролират експресията на 95 гена от общо 102 гена, участващи в Wnt сигналия път. Важни иРНКи, които се таргетират са: *APC*, *CCND1*, *CCND2*, *CTNNB1*, *FZD10*, *FZD3*, *FZD4*, *JUN*, *MYC*, *NLK*, *AFAH1B1*, *PPP2CA*, *PPP2R2A*, *PPP2R5E*, *PRKCB*, *RACGAP1*, *RHOA*, *SOX2*, *WNT1*, *WNT2*, *WNT4*, *WNT5A*, *WNT6* и *WNT7A*.



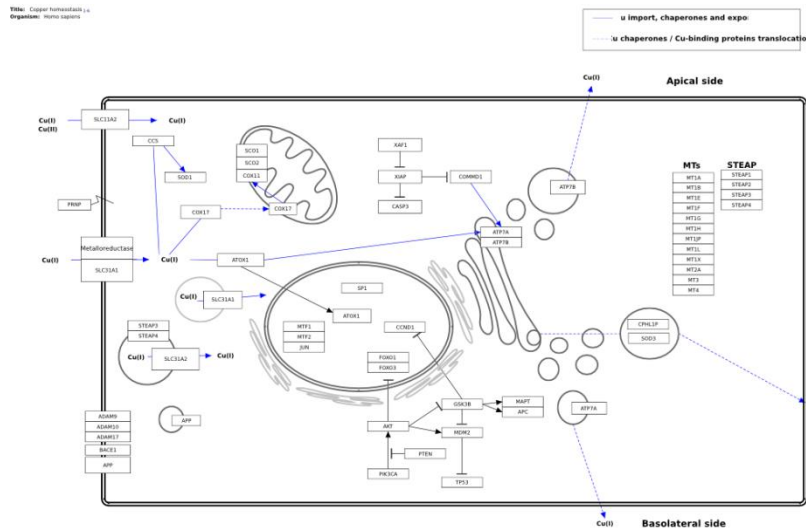
Фигура 4.49. Wnt сигнален път (195).

4.8.9. Клетъчно стареене и автофагия.

При ПК са засегнати повече сигналните пътища, свързани с клетъчния цикъл. Четирите най-централни за микроРНКите от корелационно-базирани мрежи ПК контролират експресията на 180 гена, участващи в клетъчното стареене и автофагия при малигнените заболявания. Важни иРНКи, които се тартегират са: *WEE1*, *YWHAB*, *CCND1*, *CDK6*, *CCND2*, *CDC25A*, *CDKN1A*, *MCM6*, *YWHAG*, *CHEK1*, *YWHAQ*, *RBL1*, *MYC*, *TP53*, *PRIM1*, *KIF4A*, *E2F1*, *MSH6*, *DNMT1*, *POLD3*, *POLE2*, *TOP2A*, *PTEN*, *LAMP2*, *MAP2K3*, *SERPINE1*, *KMT2A*, *INHBA*, *IL6*, *CXCL8*, *JUN*, *BCL2*, *SMAD3*, *CD44* и *MTOR*.

4.8.10. Хомеостаза на меда.

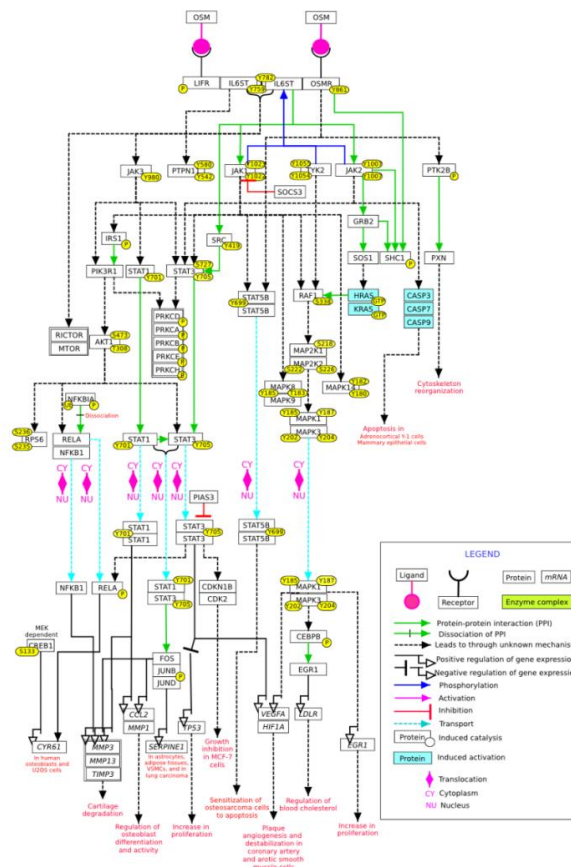
В този път са засегнати общо 40 гена от общо 52 като съотношението на шансовете е 3.417, а комбинираният резултат е 34.188. Важни засегнати гени от четирите микроРНКите при ПК са: *ADAM10*, *PTEN*, *APP*, *CASP3*, *FOXO3*, *MTF1*, *MDM2*, *S31A2*, *PIK3CA*, *XAF1*, *XIAP* и *APP*.



Фигура 4.50. Хомеостаза на медта (195).

4.8.11. Онкостатин М сигнален път.

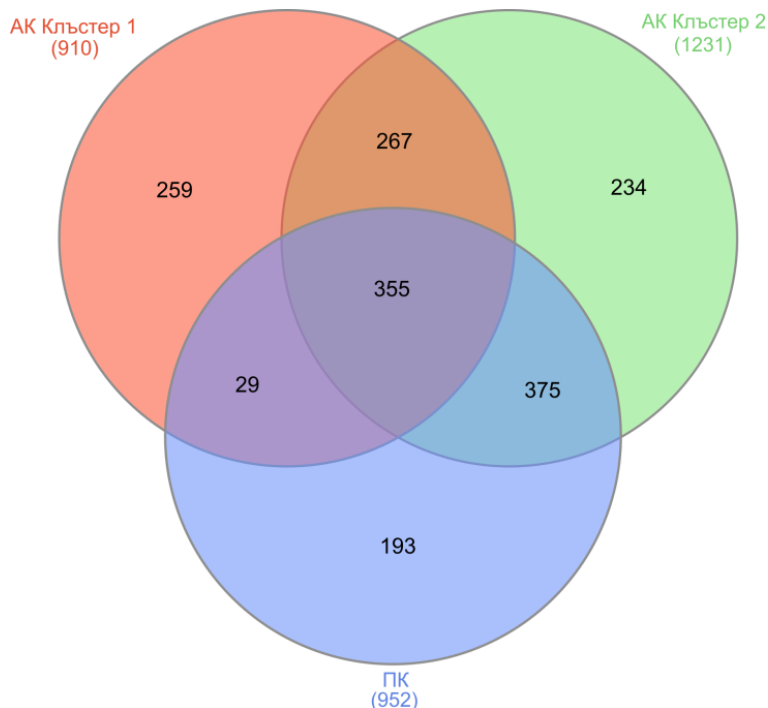
Само два гена, участващи в този сигнален път не са мишени на микроРНКите от АК Клъстер 2. Съотношението на шансовете е 8.382, а комбинираният резултат е 85.557. Шест микроРНКи таргетират гените *CREB1*, *IRS1* и *RICTOR*, а 5 микроРНКи регулират експресията на гените *CASP3*, *MAPK1*, *MAPK8* и *VEGFA*.



Фигура 4.51. Онкостатин М сигнален път (195).

4.9. Избор на мРНКи, регулирани от показалите диференциална експресия микроРНК-и в изследваните АК и ПК и имащи важно клинично значение при РБД.

На фигура 4.52 е представена Вен диаграма показваща общите и специфичните гени, срещащи се сред двайсетте най-силно засегнати сигнални пътя са за двата клъстера на АК и ПК.



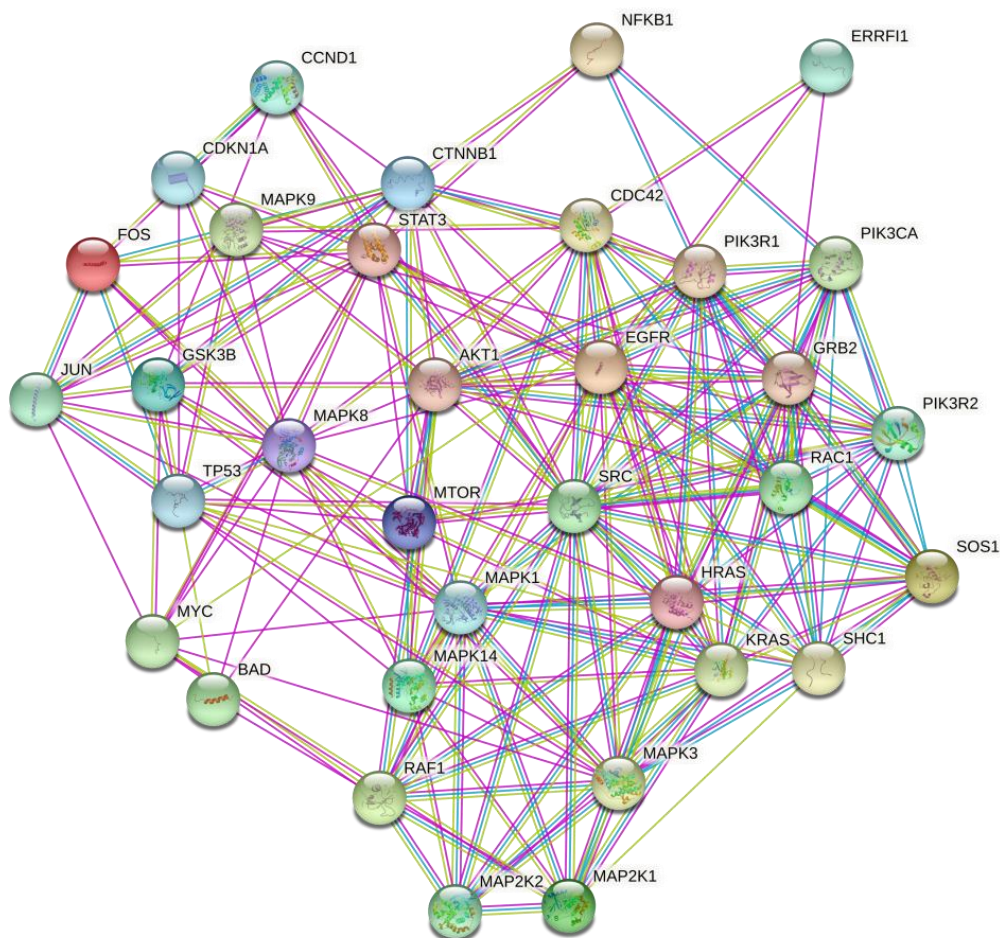
Фигура 4.52. Вен диаграма на общите и специфичните гени за двата клъстера на АК и ПК (217).

При АК Клъстер 1 най-много се обогатяват гените: *MAPK1* (14 сигнални пътя), *AKT1* (13 сигнални пътя), *KRAS* (12 сигнални пътя), *MAPK3* (12 сигнални пътя), *RAF1* (11 сигнални пътя), *MAP2K2* (10 сигнални пътя), *GRB2* (10 сигнални пътя), *RAC1* (10 сигнални пътя), *HRAS* (10 сигнални пътя), *EGFR* (10 сигнални пътя).

При АК Клъстер 2 девет гена най-често се срещат гените –*MAPK1* (17 сигнални пътя), *AKT1* (16 сигнални пътя), *RAF1* (15 сигнални пътя), *MAP2K* (14 сигнални пътя), *PIK3R1* (14 сигнални пътя), *GRB2* (14 сигнални пътя), *HRAS* (14 сигнални пътя), *PIK3R2* (13 сигнални пътя), *MAP2K2* (12 сигнални пътя), *RAC1* (12 сигнални пътя), *CCND1* (12 сигнални пътя), *MAPK3* (12 сигнални пътя), *TP53* (12 сигнални пътя), *KRAS* (12 сигнални пътя), *SOS1* (12 сигнални пътя), *NFKB1* (11 сигнални пътя), *CTNNB1* (11 сигнални пътя), *MAPK8* (11 сигнални пътя), *MYC* (11 сигнални пътя), *JUN* (10 сигнални пътя), *MAPK9* (10 сигнални пътя), *MAPK14* (10 сигнални пътя), *STAT3* (9 сигнални пътя), *SHC1* (9 сигнални пътя), *PIK3CA* (9 сигнални пътя), *SRC* (9 сигнални пътя), *CDC42* (9 сигнални пътя), *MTOR* (9 сигнални пътя), *GSK3B* (9 сигнални пътя), *EGFR* (9 сигнални пътя), *BAD* (9 сигнални пътя), *FOS* (9 сигнални пътя), *CDKN1A* (9 сигнални пътя).

Белтък-белтъчните взаимодействия бяха определени с помощта веб базираната платформа на STRING (<https://string-db.org>) (224). При АК белтъците, участващи в белтък-

белтъчната мрежа са 36 като се разглеждат общо и двата клъстера със среден коефициент на клъстериране 0.519 и р-стойност 1.0×10^{-17} . Мрежата е представена на Фигура 4. 53.

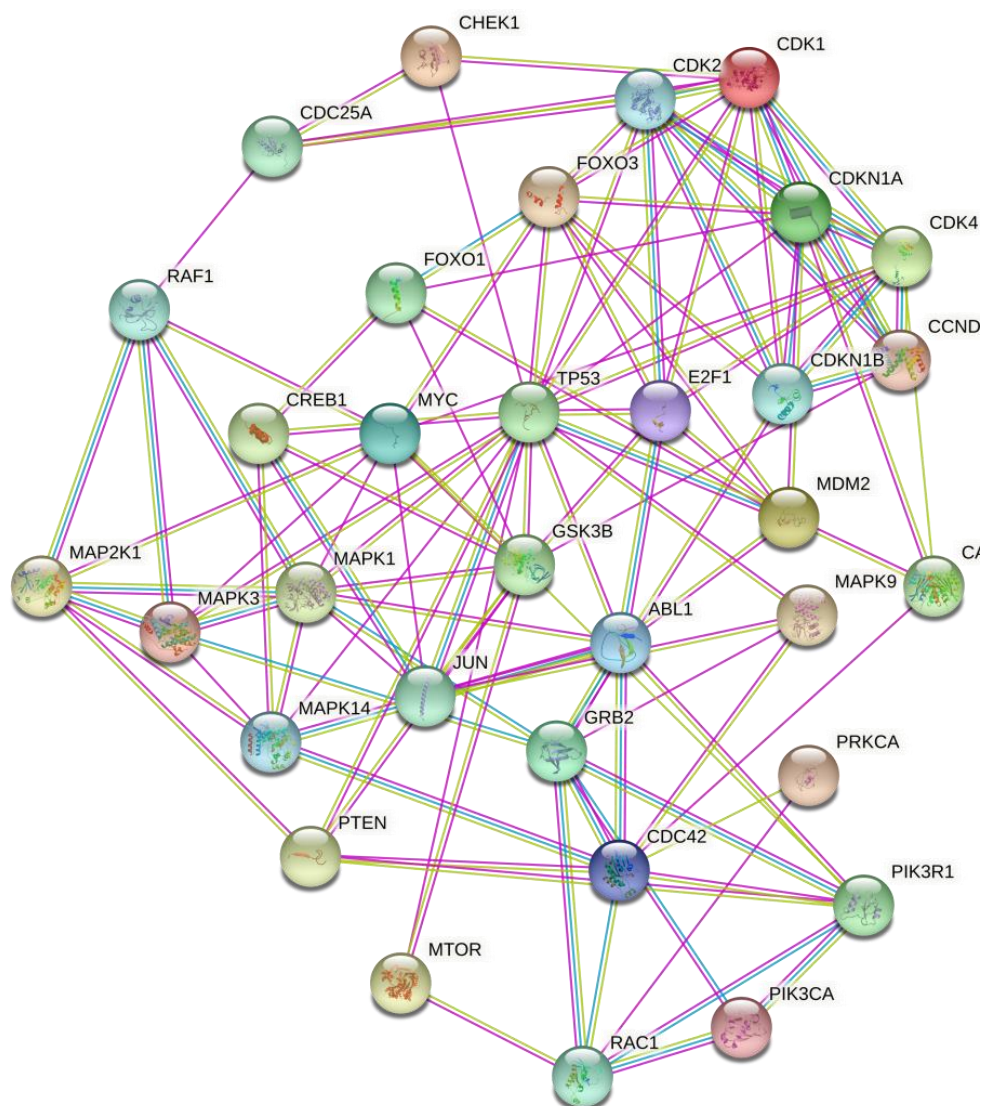


Фигура 4.53. Взаимодействие на при гените, обогатяващи се предимно в 20 най-силно повлияни от микроРНКите сигнални пътя при АК (224).

При ПК, гените, които преобладават в 20-те най-засегнати сигнални пътя са *CDKN1A* (11 сигнални пътя), *CCND1* (11 сигнални пътя), *RAF1* (11 сигнални пътя), *TP53* (11 сигнални пътя), *MDM2* (11 сигнални пътя), *MAPK1* (11 сигнални пътя), *JUN* (10 сигнални пътя), *PTEN* (10 сигнални пътя), *MYC* (9 сигнални пътя), *E2F1* (9 сигнални пътя), *ABL1* (9 сигнални пътя), *RAC1* (9 сигнални пътя), *MAP2K1* (9 сигнални пътя), *GSK3B* (9 сигнални пътя), *CREB1* (9 сигнални пътя), *MTOR* (9 сигнални пътя), *CDKN1B* (8 сигнални пътя), *CDK2* (8 сигнални пътя), *CDK1* (8 сигнални пътя), *PIK3R1* (8 сигнални пътя), *GRB2* (8 сигнални пътя), *FOXO1* (8 сигнални пътя), *CHEK1* (7 сигнални пътя), *CDC25A* (7 сигнални пътя), *CDK4* (7 сигнални пътя), *PRKCA* (7 сигнални пътя), *FOXO3* (7 сигнални пътя), *PIK3CA* (7 сигнални пътя), *CDC42* (7 сигнални пътя), *MAPK3* (7 сигнални пътя), *MAPK9* (7 сигнални пътя), *MAPK14* (7 сигнални пътя), *CASP3* (7 сигнални пътя).

Мрежата на 33-те белтъчните молекули, участващи в образуването на белтък – белтъчната мрежата и обогатяващи се предимно в 20 най-силно повлияни от

микроРНКите сигнални пътя, е представена на Фигура 4. 54.. Връзките помежду им са 166 като средният коефициент на клъстериране е 0.465, а р-стойността е $< 1.0 \times 10^{-14}$.



Фигура 4.54. Взаимодействие на при гените, обогатяващи се предимно в 20 най-силно повлияни от микроРНКите сигнални пътя при ПК (224).

4.10. РЕЗУЛТАТИ ОТ ЕКСПРЕСИОННИЯ АНАЛИЗ на *PD-L1*, *EGFR*, *STAT3*, *PTEN*, *MINA53* и *mTOR* при НДРБД.

В последните години се наблюдава значителен напредък на имуноонкологията и навлизането на имунотерапията като възможност за лечение на белодробния карцином. Поради тази причина избрахме да проследим експресията на *PD-L1* в свежи тъкани на пациенти с НДРБД. Също така, след внимателен анализ на белтък-белтъчните взаимодействия и резултатите от онтологичния анализ, избрахме още 5 гена *EGFR*, *STAT3*, *PTEN*, *MINA53* и *mTOR*, чиято експресия да изследваме при АК и ПК.

Нивата на експресия на подбраните шест иРНКи бяха изследвани и анализирани при 50 пациента с АК и 50 пациента с ПК. На Таблица 4.11 и Таблица 4.12 и са

представени брой и процент пациенти, средна стойност на относителната експресия съответно в групите показващи намалена, без промяна и повишени нива на експресия при АК и ПК.

Таблица 4.11. Относителна експресия на изследваните пациенти с АК. Представени са средните RQ стойности при пациенти, които показват понижени, без промяна и повишени в експресионни нива.

мРНК	Пациенти с понижени експресионни нива n (%) Средно RQ±SD	Пациенти с нормални нива на експресия n (%) Средно RQ±SD	Пациенти с повишени нива на експресия n (%) Средно RQ±SD	Значимост p, χ^2
PD-L1	26 (52) 0.192 ± 0.027	17 (34) 1.257 ± 0.089	7 (14) 7.845 ± 2.175	0.002
STAT3	13 (26) 0.191 ± 0.043	24 (48) 1.139 ± 0.101	13 (26) 12.410 ± 6.780	0.088
PTEN	25 (50) 0.209 ± 0.030	14 (28) 0.989 ± 0.146	11 (22) 19.323 ± 6.911	0.058
EGFR	15 (30) 0.229 ± 0.037	22 (44) 0.928 ± 0.094	13 (26) 5.531 ± 0.905	0.422
MINA53	10 (20) 0.321 ± 0.056	19 (38) 1.286 ± 0.092	21 (42) 13.975 ± 3.239	0.048
mTOR	8 (18) 0.321 ± 0.056	19 (40) 1.286 ± 0.092	25 (42) 13.975 ± 3.239	0.038

* Брой на пациенти с АК= 50

Таблица 4.12. Относителна експресия на изследваните пациенти с ПК. Представени са средните RQ стойности при пациенти, които показват понижени, без промяна и повишени в експресионни нива.

мРНК	Пациенти с понижени експресионни нива n (%) Средно RQ±SD	Пациенти с нормални нива на експресия n (%) Средно RQ±SD	Пациенти с повишени нива на експресия n (%) Средно RQ±SD	Значимост p, χ^2
PD-L1	19 (38) 0.219 ± 0.034	19 (38) 1.111 ± 0.107	12 (24) 6.293 ± 1.582	0.416
STAT3	22 (44) 0.271 ± 0.023	20 (40) 1.012 ± 0.094	8 (16) 3.284 ± 0.383	0.015
PTEN	24 (42) 0.223 ± 0.025	13 (22) 1.079 ± 0.073	13 (22) 5.193 ± 0.984	0.089
EGFR	14 (28) 0.166 ± 0.051	18 (36) 1.048 ± 0.112	18 (36) 5.599 ± 0.896	0.469
MINA53	20 (40) 0.192 ± 0.039	12 (24) 1.369 ± 0.111	18 (36) 5.763 ± 0.914	0.050
mTOR	16 (34) 0.192 ± 0.039	14 (22) 1.369 ± 0.111	20 (40) 5.763 ± 0.914	0.049

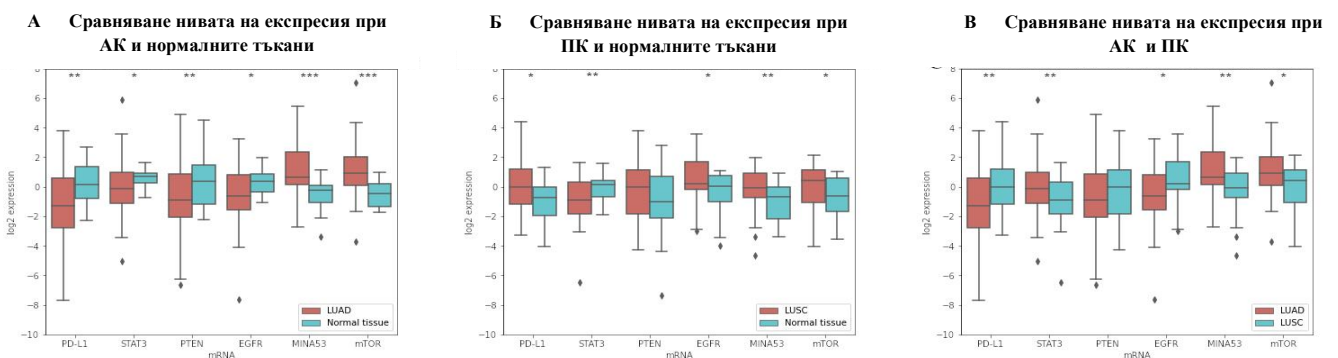
* Брой на пациенти с ПК= 50

За да се определи, кои от изследваните иРНКи са статистически значимо променени в туморна белодробна тъкан в сравнение с нормална белодробна тъкан, в зависимост от разпределението на данните, се проведе съответния параметричен или непараметричен статистически тест (Independent T-тест или Mann–Whitney unpaired тест). Бяха използвани RQ стойностите за всяка от изследваните иРНКи.

При АК *PD-L1* ($p = 0.002$), *STAT3* ($p = 0.032$), *PTEN* ($p = 0.019$), *EGFR* ($p = 0.038$) са с намалена експресия, а *MINA53* ($p = 0.001$) и *mTOR* ($p = 0.001$) са с повишена експресия, сравнение с нормалните тъкани. Преобразуваните експресионни нива са представени чрез бокс плот на Фигура 4.55А.

При ПК експресията на *PTEN* не е статистически значима спрямо нормалните тъкани. Докато експресията на *PD-L1* ($p = 0.033$), *EGFR* ($p = 0.046$), *MINA53* ($p = 0.013$) и *mTOR* ($p = 0.029$) е значимо повишена, а експресията на *STAT3* ($p = 0.003$) значимо намалена. Преобразуваните експресионни нива са представени чрез бокс плот на Фигура 4.55Б.

При сравняване на експресионните нива на изследваните иРНКи при АК и ПК само *PTEN* ($p = 0.226$) не показва статистически значима разлика в експресията между двата подтипа белодробен карцином. Повишени при АК сравнение с ПК беше *MINA53* ($p = 0.014$), а *PD-L1* ($p = 0.017$), *STAT3* ($p = 0.017$), *EGFR* ($p = 0.024$), и *mTOR* ($p = 0.029$) бяха със значимо намалена експресия при АК. Преобразуваните експресионни нива са представени чрез бокс плот на Фигура 4.55В.



Фигура 4.55. (А) Сравнение нивата на експресия на изследваните иРНКи при АК и прилежащите нормални тъкани; (Б) Сравнение нивата на експресия изследваните иРНКи при ПК и прилежащите нормални тъкани; (В) Сравнение нивата на експресия при АК и ПК; По x-оста са иРНКите, по y-оста log₂ от експресионните нива. Беше използван Independent T-тест или Mann–Whitney unpaired тест, според разпределението на данните; $p \geq 0.05$ е незначимо, *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, LUAD – АК, LUSC – ПК

4.10.1. Корелационен анализ между изследваните иРНКи и микроРНКи.

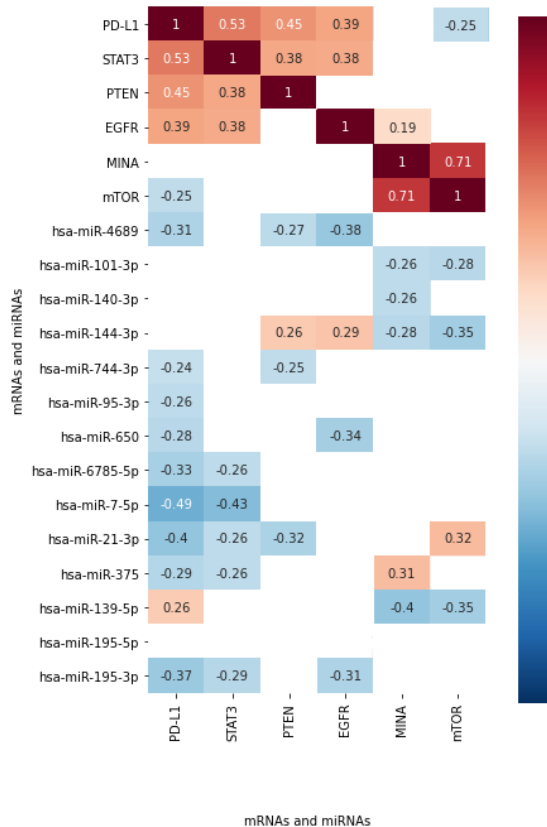
С цел да се анализира дали са свързани и каква е силата на зависимост между нивата на относителна експресия на от една страна между самите иРНКи и от друга страна връзката между изследваните микроРНКи и иРНКи се използва корелационен анализ на Spearman.

На следващата Фигура 4.56 и Фигура 4.67 са представени съответно матрица с корелационните коефициенти на всички възможни комбинации между иРНКите микроРНКите. В червената скала са отбелязани положителните корелации, докато в

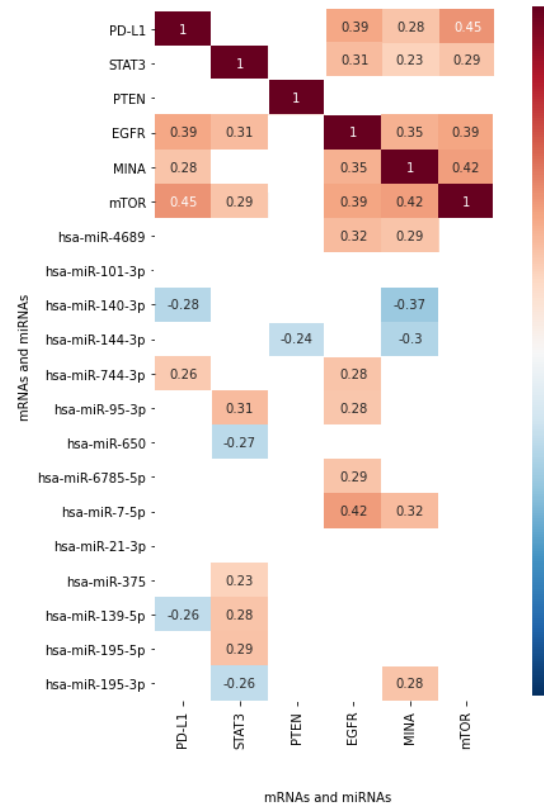
синята скала са отбелязани отрицателните корелации. Визуализирани са само статистически значимите корелации с $p \leq 0.05$.

При АК е налице силна положителна корелация на *PD-L1* с *STAT3*, *PTEN*, *EGFR* и отрицателна с *mTOR*. Експресията на *STAT3* корелира също с *PTEN*, а експресията на *MINA53* с *mTOR*. Връзката на *PD-L1* с miR-744-3p, miR-95-3p, miR-650, miR-6785-5p, miR-7-5p, miR-21-3p, miR-375 и miR-195-3p е отрицателна. Докато *PD-L1* положително корелира с miR-139-5p. Експресията на *STAT3* отрицателно корелира с miR-6785-5p, miR-7-5p, miR-21-3p, miR-375 и miR-195-3p. Нивата на *EGFR* имат отрицателна зависимост с miR-4689, miR-650, miR-195-3p и положителна корелация с miR-144-3p. *PTEN* корелира с miR-4689, miR-144-3p, miR-744-3p и miR-21-3p. Връзката на *MINA53* с miR-101-3p, miR-140-3p, miR-144-3p и miR-139-5p е отрицателна, а с miR-375, ниска положителна връзка. Експресията на *mTOR* корелира с miR-101-3p, miR-144-3p, miR-139-5p и miR-21-3p.

При ПК експресията на *PD-L1* корелира положително с *EGFR*, *MINA53* и *mTOR*. Експресията на *EGFR* корелира също с *PD-L1*, *STAT3*, *MINA53* и *mTOR*. *MINA53* с *mTOR* също както при АК корелират, но коефициента на корелация е по-малък. Връзката на *PD-L1* с miR-140-3p и miR-139-5p е отрицателна, докато за разлика от АК с miR-744-3p е положителна. Експресията на *STAT3* корелира отрицателно с miR-650 и miR-195-3p. Докато коефициентът на корелация е положителен на *STAT3* с miR-95-3p, miR-375, miR-139-5p и miR-195-5p. Нивата на *EGFR* имат положителна зависимост с miR-744-3p, miR-95-3p, miR-6785-5p и miR-7-5p. Връзката на *MINA53* с miR-140-3p, miR-144-3p е отрицателна, а с miR-7-5p и miR-195-3p положителна връзка. Експресията на корелира с miR-101-3p, miR-144-3p, miR-139-5p и miR-21-3p. Относителната експресия *mTOR* не корелира с нито една от изследваните микроРНК при ПК.



Фигура 4.56. Матрица с корелационните коефициенти на възможните комбинации между микроРНКи и мРНКи при АК. Визуализирани са само статистически значимите корелации с $p \leq 0.05$, без приложена корекция.



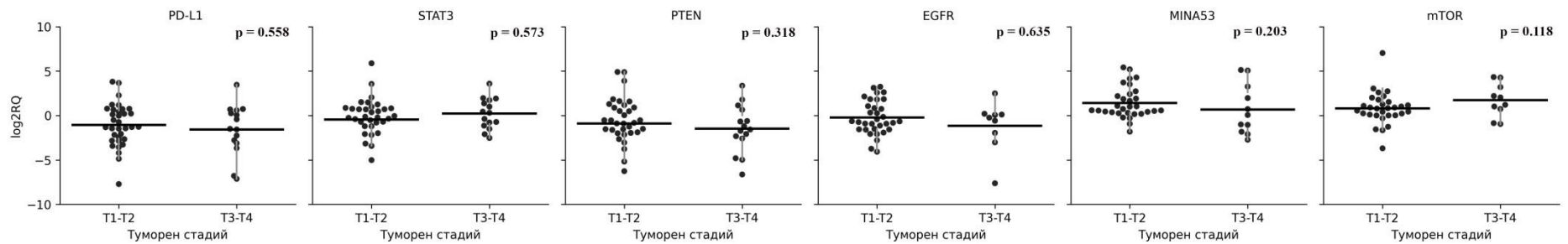
Фигура 4.57. Матрица с корелационните коефициенти на възможните комбинации между микроРНКи и мРНКи при ПК. Визуализирани са само статистически значимите корелации с $p \leq 0.05$, без приложена корекция.

4.10.2. Асоциация между експресионните нива на иРНКите и кликопатологичните характеристики на пациентите.

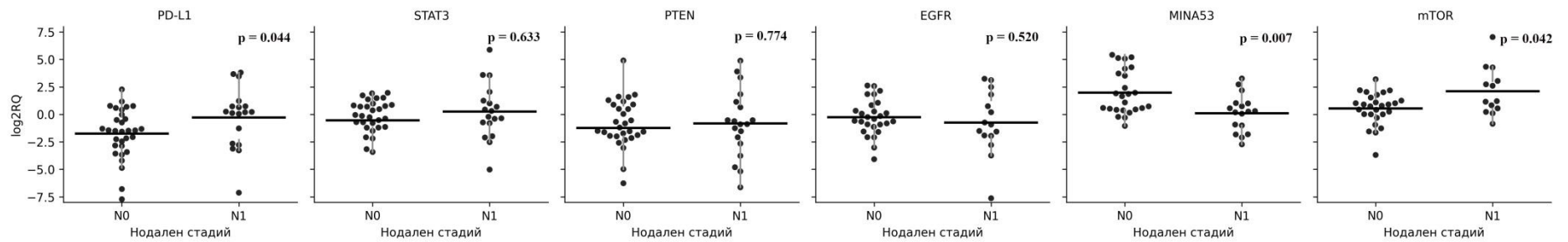
Асоциацията между получените RQ данни на изследваните микроРНКи и групирането, базирано на кликопатологичните характеристики на аденокарциномите е представено на Фигури 4.58 - 4.66.

Получените данни показаха статистически значима асоциация между експресионните нива на изследваните иРНКи с различни кликопатологичните характеристики. При АК иРНКи не показаха асоциация с туморния стадий, локализацията на тумора, пола и годините на пациентите. Експресията на *PD-L1* ($p=0.014$) и *mTOR* ($p=0.041$) се асоциира статистически значимо нодалния стадий и значимо по-висока експресия при наличието на метастази. Слабо диференцираните тумори се характеризират със значимо по-висока експресия на *PD-L1* ($p=0.050$), *PTEN* ($p=0.004$) и *EGFR* ($p=0.006$). Нивата на *STAT3* също са значимо по-високи при дистални метастази ($p=0.045$). *MINA53* се асоциира с лимфните ($p=0.007$) и далечни ($p=0.019$) метастази и диференциацията на тумора ($p=0.016$). При АК с преобладаваща солидна морфология се наблюдава и значително по-висока експресия на *PD-L1* ($p=0.040$), *EGFR* ($p=0.020$) и *PTEN* ($p=0.018$).

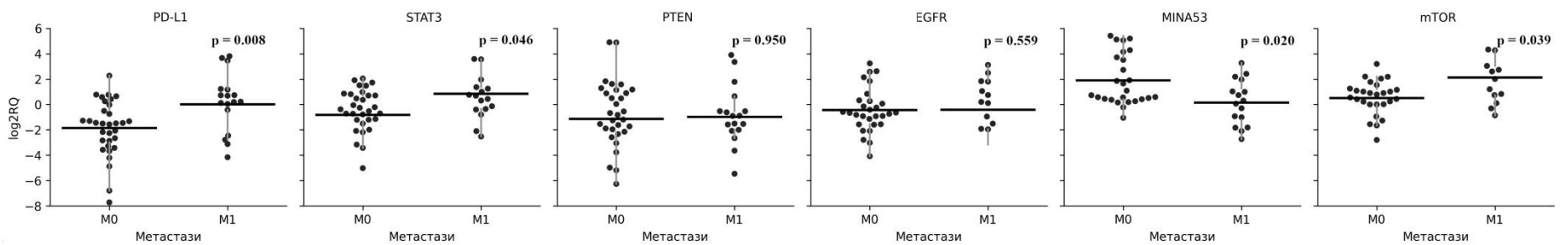
При ПК иРНКи не показаха асоциация също с туморния стадий, локализацията на тумора, пола и годините на пациентите. Експресията на *PD-L1* ($p=0.003$) и *EGFR* ($p=0.010$) са значимо по-високи при ниско диференцираните тумори и при пациентите, при които се откриват далечни метастази. Експресията на *STAT3* е по-ниска при добре диференцираните тумори ($p=0.0001$). Нивата на *PTEN* са по-ниски при пациенти без далечни метастази ($p=0.007$). *MINA53* е с по-висока експресия при пациентите, при които не се откриват дистални метастази и при умерено и ниско диференцираните тумори Фигури 4.67 - 4.74.



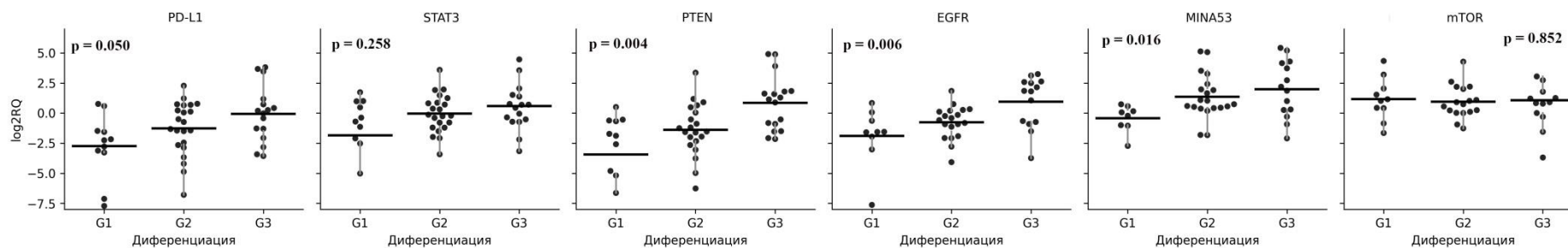
Фигура 4.58. Разпределение на броя пациенти спрямо туморния стадий, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



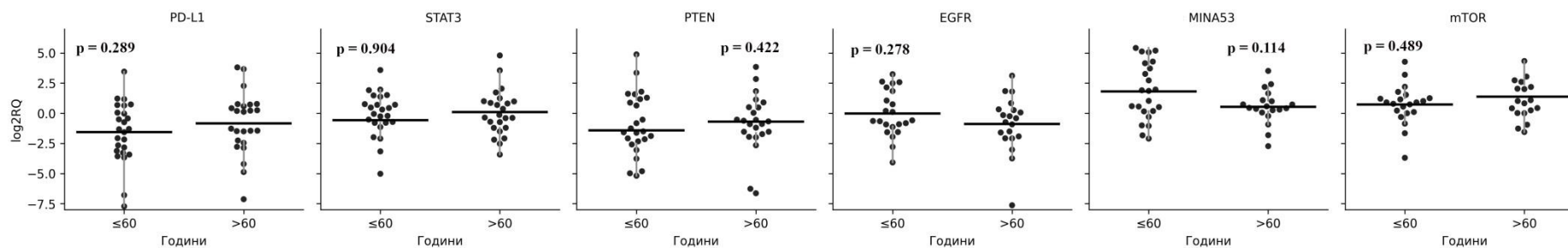
Фигура 4.59. Разпределение на броя пациенти спрямо нодалния стадий, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



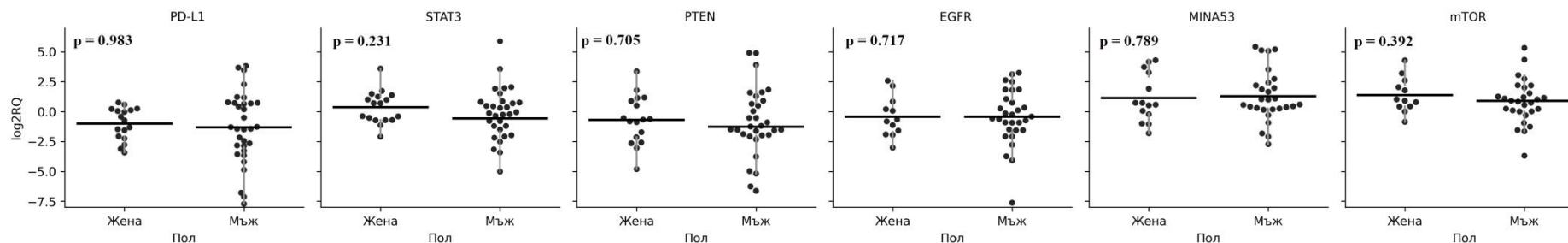
Фигура 4.60. Разпределение на броя пациенти спрямо метастазите, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо) при АК.



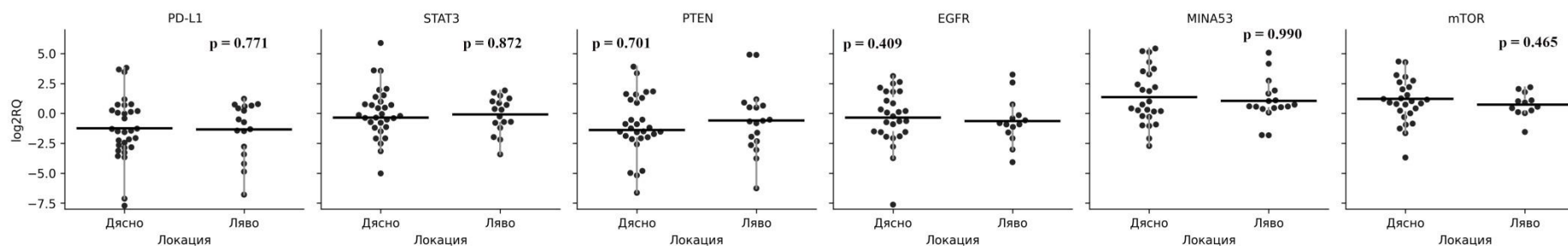
Фигура 4.61. Разпределение на броя пациенти спрямо диференциацията на АК, експресията на иРНКите и резултати от проведения статистически тест (Kruskal–Wallis test; $p \geq 0.05$ е незначимо).



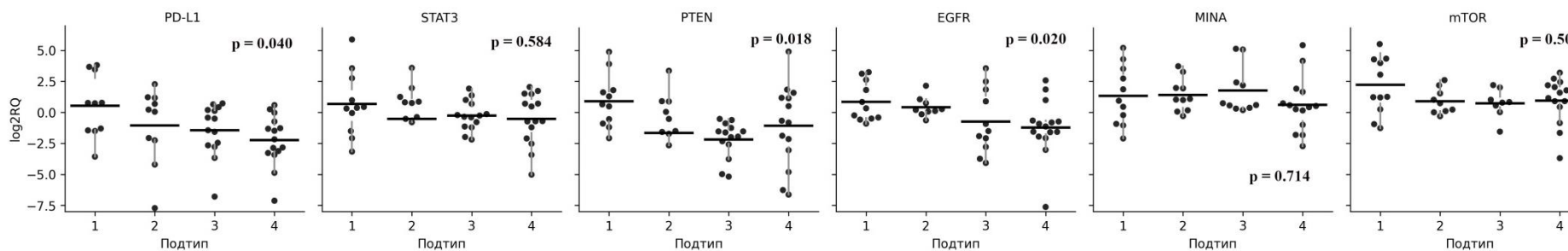
Фигура 4.62. Разпределение на броя пациенти спрямо възрастта на пациентите с АК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо).



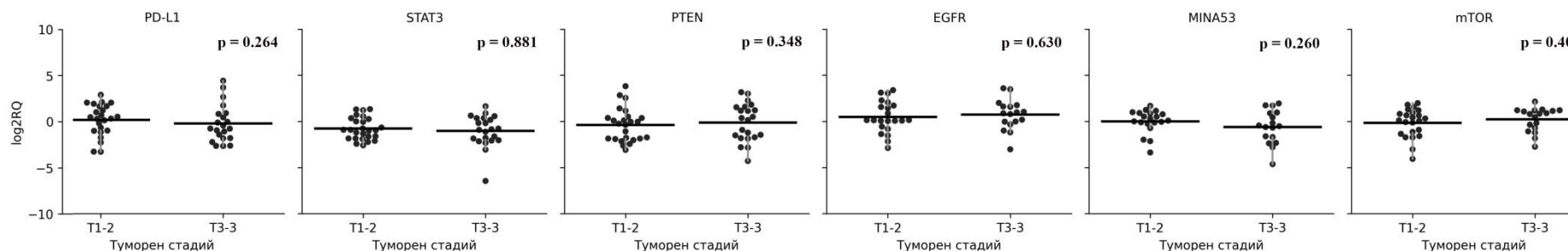
Фигура 4.63. Разпределение на броя пациенти спрямо пола на пациентите с АК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо).



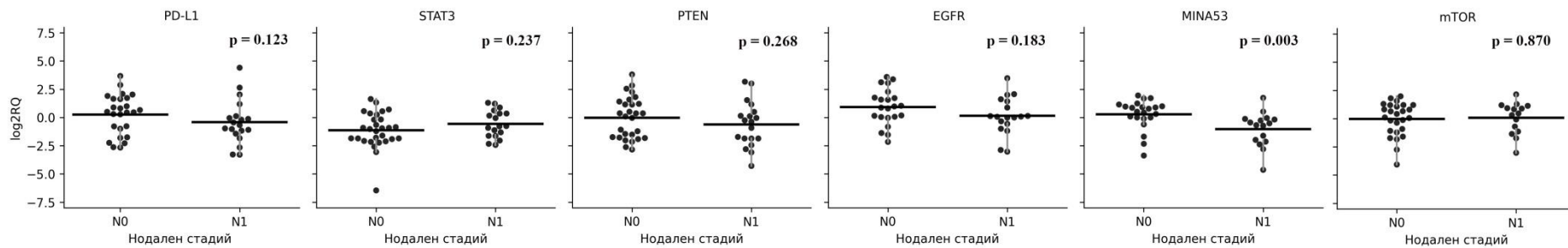
Фигура 4.64. Разпределение на броя пациенти спрямо локализацията на АК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо).



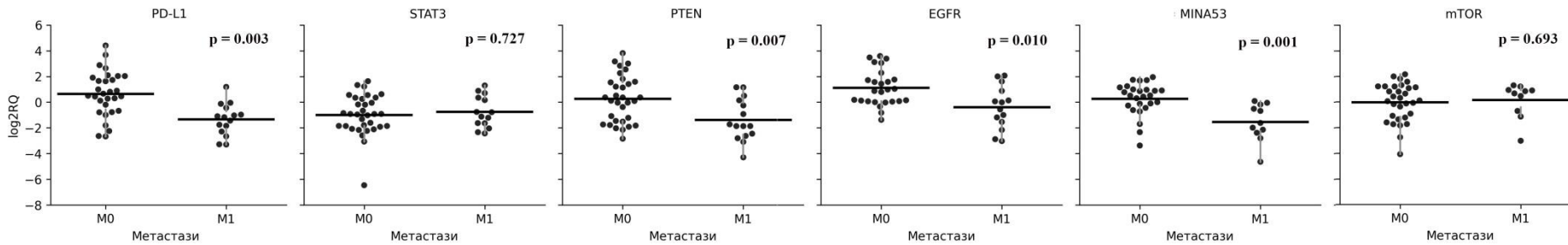
Фигура 4.65 Разпределение на броя пациенти спрямо морфологичния вид АК, експресията на иРНКите и резултати от проведения статистически тест (Kruskal-Wallis test; $p \geq 0.05$ е незначимо).



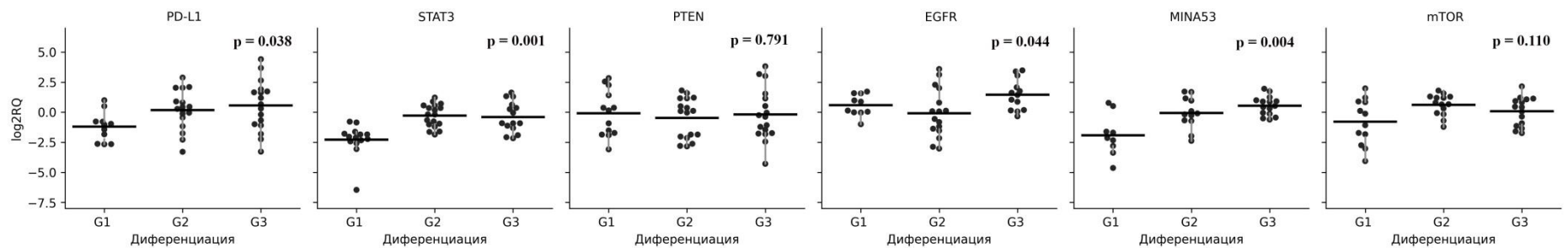
Фигура 4.67. Разпределение на броя пациенти спрямо туморния стадий на ПК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо).



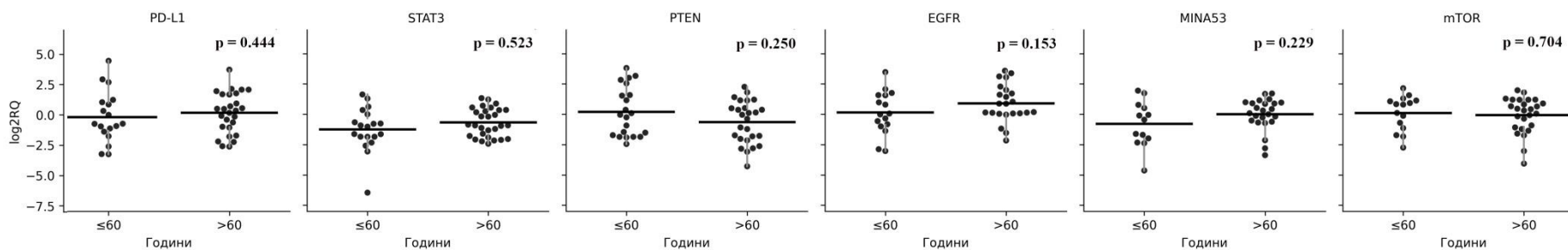
Фигура 4.68. Разпределение на броя пациенти спрямо нодалния стадий на ПК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо).



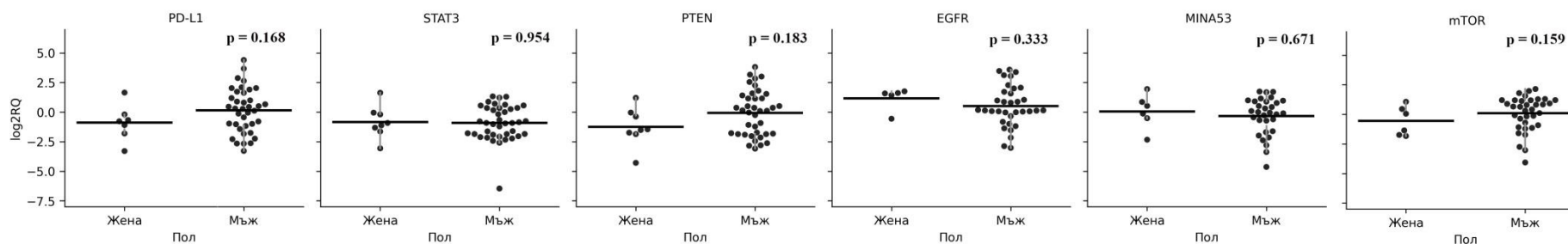
Фигура 4.69. Разпределение на броя пациенти спрямо наличието на метастази при, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U test; $p \geq 0.05$ е незначимо).



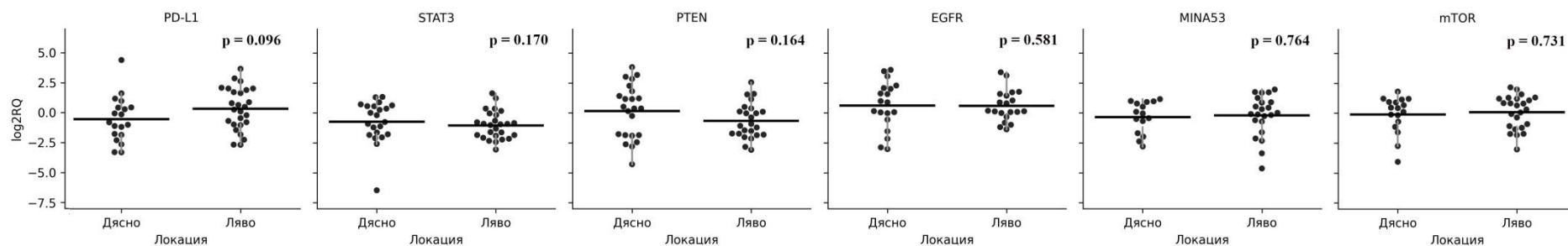
Фигура 4.70. Разпределение на броя пациенти спрямо диференциацията на ПК, експресията на иРНКите и резултати от проведения статистически тест (Kruskal-Wallis test; $p \geq 0.05$ е незначимо).



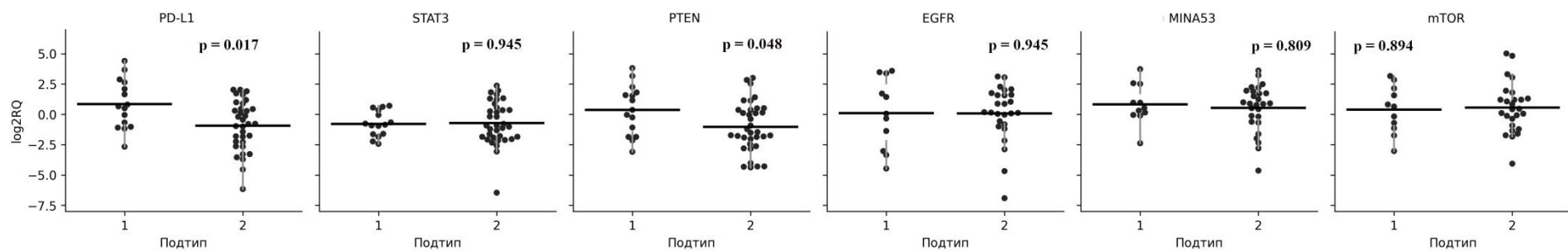
Фигура 4.71. Разпределение на броя пациенти спрямо годините, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U тест; $p \geq 0.05$ е незначимо) при ПК.



Фигура 4.72. Разпределение на броя пациенти спрямо диференциацията на ПК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U тест; $p \geq 0.05$ е незначимо).



Фигура 4.73. Разпределение на броя пациенти спрямо локализацията на ПК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U тест; $p \geq 0.05$ е незначимо).



Фигура 4.74. Разпределение на броя пациенти спрямо морфологичния тип ПК, експресията на иРНКите и резултати от проведения статистически тест (Mann-Whitney U тест; $p \geq 0.05$ е незначимо).

4.11. Резултати от контролираното самообучение (Supervised Learning) за разграничаване на двата подтипа НДБРД.

След почистването, нормализирането, скалирането и разделяне на данните следва тяхното моделиране. Като избрането на подходящ алгоритъм за машинно самообучение и настройката на конфигурациите (хиперпараметрите) на всеки алгоритъм е от съществено значение за разграничаване на двата подтипа белодробен карцином.

4.11.1. Избор на алгоритъм за класификационен анализ.

Върху хетерогенните данни за експресията на 14 микроРНК се тестват няколко алгоритъма за класификация. Всеки един от моделите бива трениран за предсказване на един от двата класа: клас 1 – АК и клас 2 – ПК.

4.11.1.1. Трениране и тестване на модели.

След като данните са в подходящ вид за провеждане на машинно самообучение, те биват разделени на две части. 70% от данните се използват за трениране на модела, а 30% се използват за тестване. При разделяне, за да са сравними данните за трениране и тестване, се прилага стратификация по вид на туморите (АК или ПК). Приложени са четири вида алгоритми:

- Логистична регресия (Logistic regression; LR);
- Дърво на решенията (Decision tree; DT);
- Случайна гора (Random forest; RF);
- Метод на опорните вектори (Support Vector Machines; SVM);

Всеки един от четирите алгоритъма е резултат от различна моделираща функция. Тя определя начина, по който работи всеки алгоритъм и резултатите, които се получават. За да се избегне прекаленото нагаждане на към данните е необходимо да се избере правилната конфигурация (хиперпараметри). За тази цел върху данните за трениране се извършва кросвалидация и регуларизация.

4.11.1.2. Logistic regression.

Алгоритъм, при който променливата има две възможни стойности и използва само линейни функции (198). Моделът пресмята вероятността пациента да принадлежи към една от две категории (АК или ПК), в зависимост от експресията на микроРНКите.

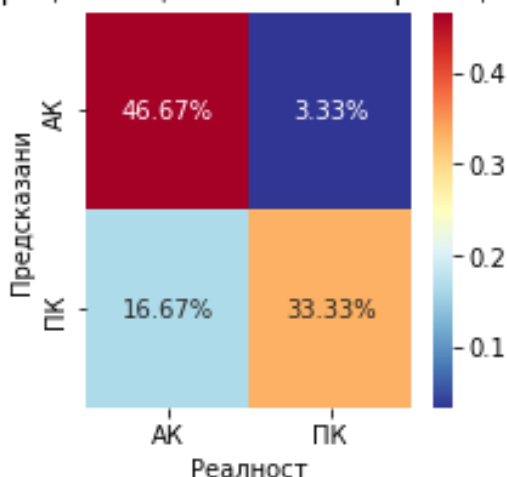
4.11.1.2.1. Конфигуриране на алгоритъма.

За толеранс на оптимизацията е избрана стойността $C=1$ и 10 максимални итерации. Добавена е регуларизация за предотвратяване прекомерно нагаждане на резултатите от модел. За да са максимално по-точни резултатите е приложена комбинация два типа регуларизация – L1 (абсолютна стойностна норма) и L2 (Евклидова норма) с тежест 0.5 или т.нар. Elastic Net регуларизация се използва широко в мултиомични анализи.

4.11.1.2.2. Резултати след извършено трениране на модела.

Фигура 4.75 показва резултатите получени след трениране и тестване на модела с логистичната регресия. Верните и грешните предположения според алгоритъма са систематизирани в матрица за оценка на класификацията.

Матрица за оценка на класификацията



Фигура 4.75. Матрицата за оценка на логистичната регресия.

Матрицата представя процентът на верните предположения след тестване на модела. За ПК 33.33% (10 пациента) за и 46.67% от АК (14 пациента) са правилно класифицирани от логистичната регресия. Неправилно класифицираните от логистичната регресия са 3.33% от АК (1 пациент) и 16.67% (5 пациента) от ПК, погрешно са класифицирани като АК.

4.11.1.3. Decision tree.

Дърво за вземане на решение или Decision tree (DT) е съставено от следните елементи: корен, много възли, клонови и листа. Коренният възел прави разделянето въз основа на стойността на атрибута на класа, вътрешният възел взема атрибут за по-нататъшна класификация, клоните вземат решение за разделяне на възлите на листови възли, накрая, листовите възли ни дават крайния резултат (198). Сложността на времето на DT зависи от броя на записите и атрибутите на данните от обучението. Ако DT е твърде голямо, е трудно да получите желаните резултати (225). Поради тази причина е от съществено значение правилната му настройка. DT е подходящ за работа алгоритъм, когато проблемът е прост и данните са малко (226). Въпреки че алгоритъмът е лесен за разбиране, той има някои проблеми, като например проблема с прекомерното нагаждане към данните и прекалено добрите резултати при работа с небалансирани набори от данни (227). Но DT е способен да картографира както линейни, така и нелинейни връзки (228).

4.11.1.3.1. Конфигуриране на алгоритъма

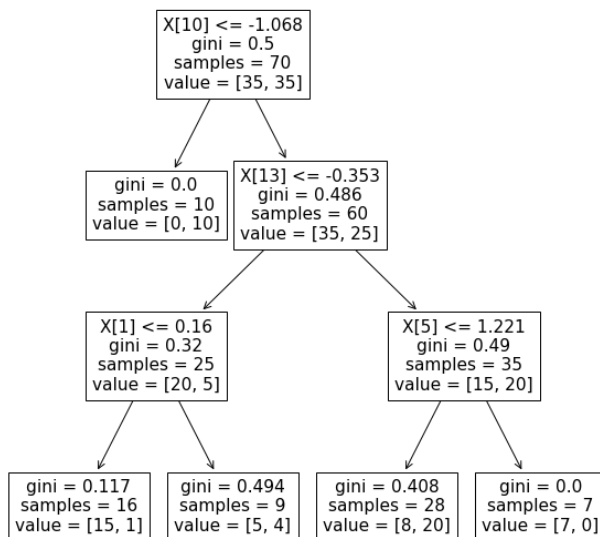
За правилната конфигурация на дърво на решенията е важно да бъдат определени точно: максимална дълбочина на дървото и минималният брой проби, които трябва да се намират в листов възел. Точка на разделяне на всякаква дълбочина ще се разглежда само ако останат минимален брой обучителни проби във всеки от левите и десните клонове.

Максимална дълбочина на дървото в случая е 3, а минималният брой проби, които трябва да се намират в листов възел са 7.

4.11.1.3.2. Резултатите след трениране на модела.

Моделът е съставен от набор от "въпроси", организирани йерархично във формата на дърво и представен схематично на фигура 4.76. Въпросите обикновено се наричат условие. Всеки нелистен възел съдържа условие, а всеки листен възел съдържа предсказание. Изводът на модела на дърво на решенията се изчислява чрез насочване на пример от корена (в горната част) към един от листните възли (в долната част) в съответствие с условията. Стойността на достигнатия листов възел е прогнозата на дървото на решенията. Gini Index е резултат, който оценява колко точно е разделението между класифицираните групи. Колкото по-добре са разделени двата класа, толкова по-малка е неговата стойност (198).

На Фигура 4.77 е показана матрицата на вероятностите след трениране и тестване на модела. Матрицата представя процентът на верните предположения, които са в 26.67% – за ПК (8 пациента) и 20% за АК (6 пациента). 30% от ПК (9 пациента) неправилно са класифицирани като АК от модела, а 23.33% от АК (7 пациента) неправилно са класифицирани като ПК.



Фигура 4.76. Визуализация на Дърво на решенията.



Фигура 4.77. Матрицата за оценка на Дърво на решенията.

4.11.1.4. Random forest (RF).

През последното десетилетие случайна гора или Random forest (RF) алгоритмите се прилагат успешно за различни проблеми в областта на изчислителната биология. Това са клас модели за машинно самообучение, състоящи се от съчетание от дървета на решенията, които се учат независимо (bagging) (229). RF определя консенсусния резултат, като комбинира прогнозите на отделните дървета на решенията. Той е непараметричен, интерпретируем, ефективен и има висока точност на прогнозиране при работа с малък размер на извадката, данни с голяма размерност и сложни структури (230). RF и неговите варианти се прилагат за различни биоинформатични проблеми, като например анализ на генна експресия, анализ на белтъци чрез мас-

спектрометрия, откриване на биомаркери, аотиране на последователности и прогнозиране на белтъчно-белтъчно взаимодействия (229).

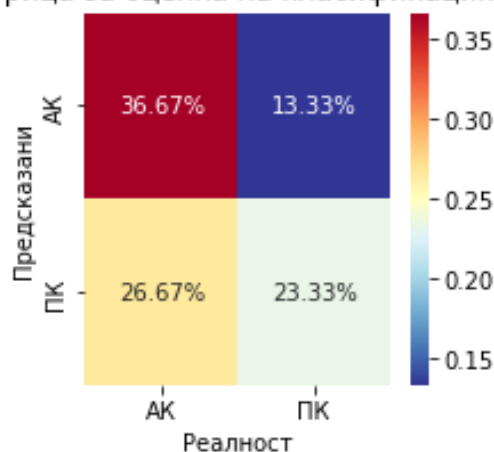
4.11.1.4.1. Конфигуриране на алгоритъма.

Най-важните настройки са броят на дърветата в гората, броят на характеристиките, които се вземат предвид за разделяне във всеки листов възел и максималната дълбочина на дърветата. В случая броят на дърветата в гората е 70 и максималната дълбочина на дърветата е 50 и минималният брой проби, които трябва да се намират в листов възел са 7.

4.11.1.4.2. Резултатите след трениране на модела.

Фигура 4.78 е показана матрицата на вероятностите след трениране и тестване на молела. Общият брой на верните предположения е в 23.33% – за ПК (7 пациента) и 36.67% за АК (11 пациента), а за грешните предположения е 13.33% – за фалшиви положителните резултати при ПК и 26.67% (8 пациента) за фалшиво отрицателните резултати при АК (4 пациента).

Матрица за оценка на класификацията



Фигура 4.78. Матрицата за оценка на случайната гора.

4.11.1.5. Метод на опорните вектори (Support Vector Machines; SVM).

SVM алгоритмите за класификация са въведени за първи път от Vapnik (231). Днес SVM са често използван клас алгоритми за машинно самообучение (232). Те са особено подходящи и добри в случаите, когато броят на наличните данни за обучение е ограничен (233,234). Причината е, че когато данните продължават да се увеличават, обучението става по-сложно и отнема много време. Методът на опорните вектори представя данните като точки в n -мерно пространство, използвайки специфични функции на ядрото. Данните са проектират в пространството по такъв начин, че да бъдат линейно разделими. При работа с два класа се търси начин да се начертае линия (разделителна хиперравнина), която да разделя данните от двата класа. Тази хипер равнина трябва да се избере по такъв начин, че да се намира възможно най-далеч от примерите и на двата класа (232). Докладвано, че те работят добре, когато данните са представени като вектори на непрекъснати променливи и има по-голяма точност в сравнение с другите алгоритми (228). Поради тези характеристики този метод до голяма степен се използва в проблемите на класификацията в областта на

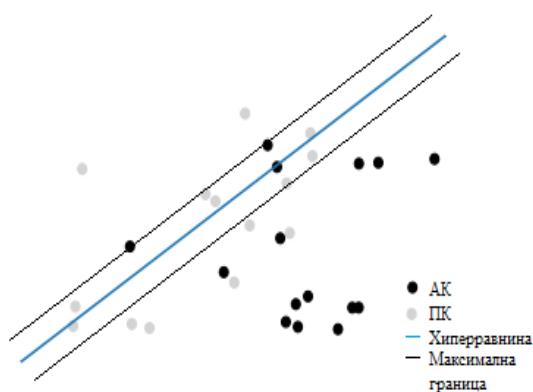
биоинформатиката и по-специално за геномна класификация или под типизиране на ракови заболявания (233,235).

4.11.1.5.1. Конфигуриране на алгоритъма.

Параметърът на регуларизация (лямбда) служи за степен на важност, която се придава на грешните класификации. SVM представлява квадратична оптимизационна задача, която търси максимизиране на маргина между двата класа и минимизиране на количеството погрешни класификации. Използваната лямбда стойност е 1.5 и се използва L1 регуларизация.

4.11.1.5.2. Резултатите след трениране на модела.

На Фигура 4.80 е изобразена геометричната илюстрация на метода на опорните вектори. На Фигура 4.81 е показана матрицата на вероятностите след трениране и тестване на модела. Общият брой на верните предположения е в 30% – за ПК (9 пациента) и 50% за АК (15 пациента), а за грешните предположения е 20% – за фалшиви положителните резултати за ПК (6 пациента).



Фигура 4.80. Геометрична илюстрация на SVM класификатор.



Фигура 4.81. Матрицата за оценка на SVM.

4.11.2. Оценка на моделите.

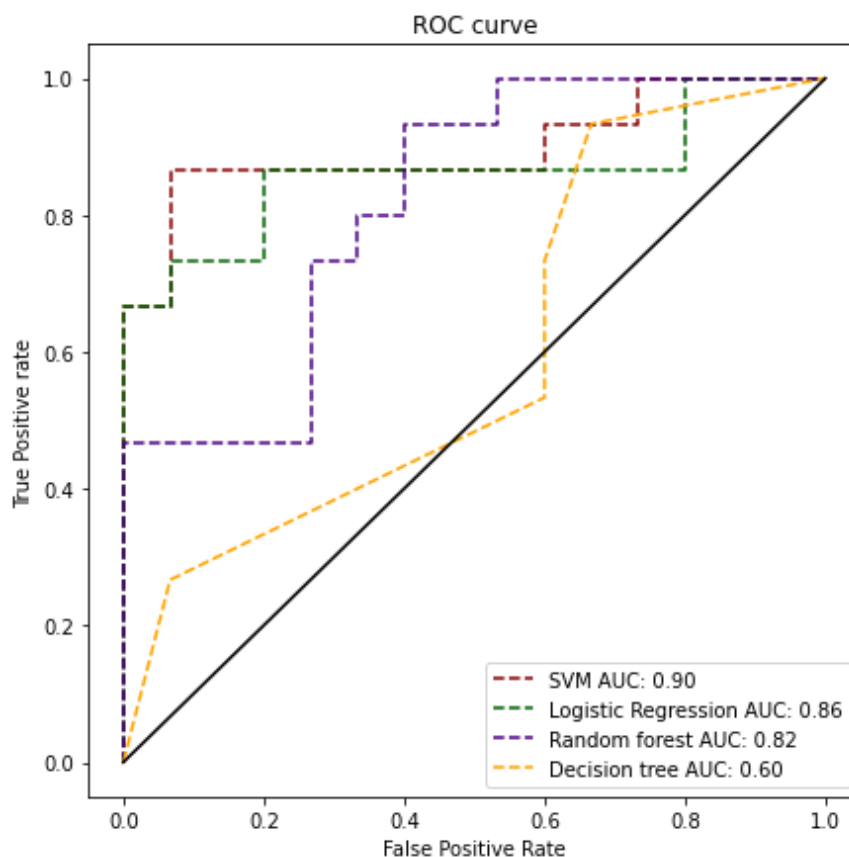
За да се оцени един модел е необходима повече от една метрика. На базата на матриците за оценка на класификацията се изчисляват и основните метрики за оценка на алгоритмите - точност (accuracy), прецизност (precision), чувствителност (recall), хармонична средна стойност (F1-Score), положителна прогностична стойност (PPV) и отрицателна прогностична стойност (NPV). На Таблица 4.13. са показани основните метрики за оценка на всеки от приложените модели.

Алгоритъм	Точност на модела (accuracy)	Прецизност на модела (precision)	Чувствителност на модела (recall)	Средна хармонична стойност (F1-score)
Logistic regression	0.80	0.91	0.67	0.77
Decision Tree	0.47	0.47	0.53	0.50
Random forest	0.60	0.64	0.46	0.54
SVM	0.80	0.99	0.60	0.75

Logistic regression и SVM са с най-висока точност – 0.80, а Decision Tree с най-ниска точност - 0.47. Хармоничната средна стойност между прецизност и чувствителност при Random forest е 0.54 и 0.50 при Decision Tree. Най-висока хармонична средна стойност има логистичната регресия - 0.77, последвана от SVM с 0.75.

Моделите логистичната регресия и SVM се открояват като по-добри класификатори спрямо останалите. Високият резултат, получен при тях, трябва да се провери, чрез различни методи, за да избегнат недостоверни резултати, поради прекомерно нагаждане на алгоритмите към данните.

Един от широко използваните методи за оценка на бинарна класификация е ROC крива. При нея се съпоставят TPR (True Positive Rate) и FPR (False Positive Rate) при различни прагове. На Фигура 4.82 е представена ROC крива за логистичната регресия, дърво вземащо решения, случайна гора и . Черната линия е случайно взет класификатор. Колкото повече се отдалечава от нея кривата на определен модел и се доближава до горния ляв ъгъл на графиката, толкова по-добър е той. Площта под кривата (AUC) определя до каква степен класификаторът може да разграничава единия от другия клас при разпределяне на обектите. Стойностите на площта под кривата са в интервал между 0 и 1. Ако AUC е 1.0, тогава моделът е определил всички случаи правилно. Площта под кривата при е най-голяма при SVM е 0.90, следвана от линейната регресия с AUC = 0.86, при дърво вземащо решения AUC = 0.82 и при случайна гора AUC =0.60.



Фигура 4.82. ROC крива на логистичната регресия, дърво вземащи решения, случайна гора и SVM.

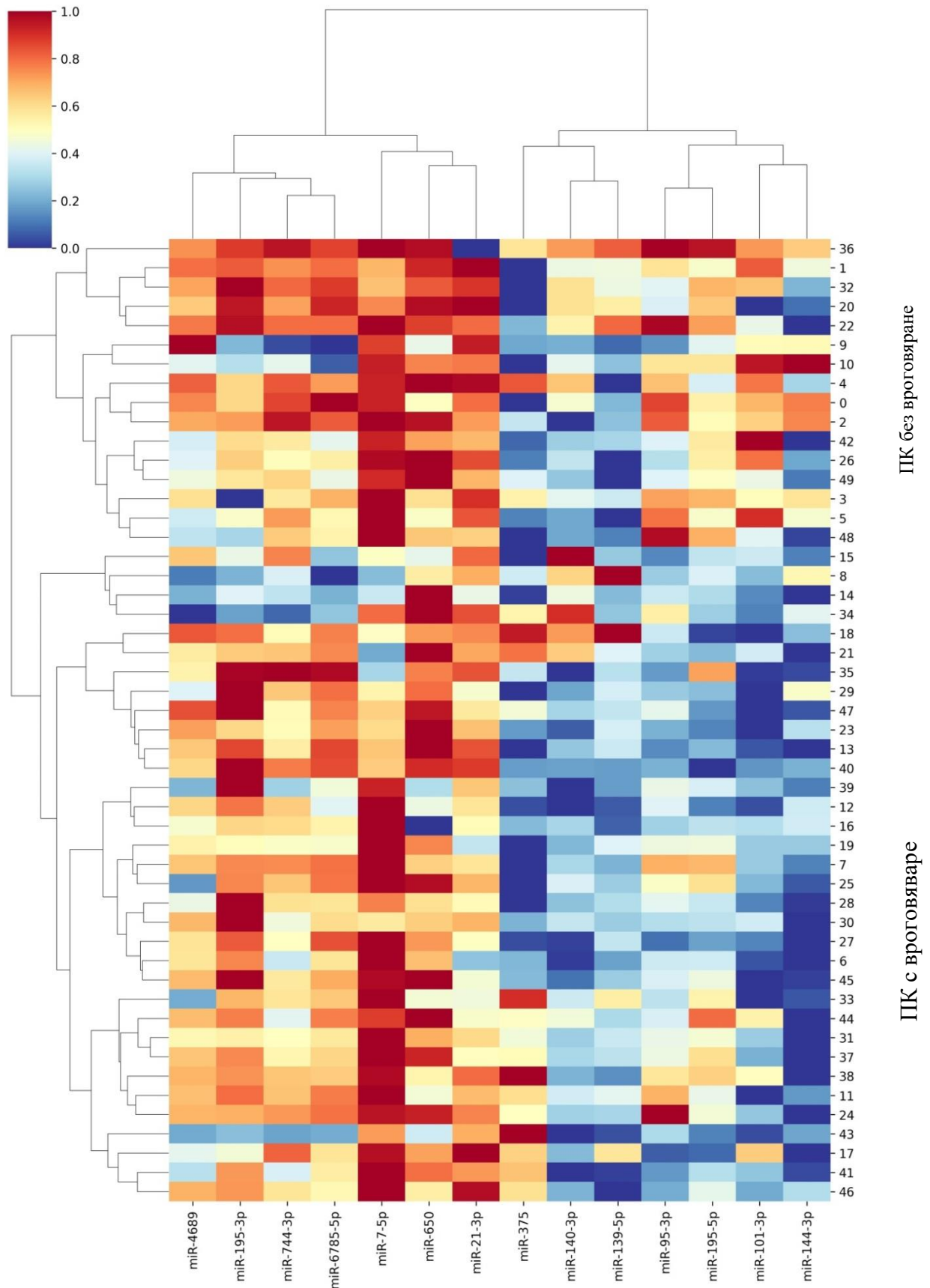
4.12. Резултати от неконтролирано самообучение (Unsupervised Learning) за разграничаване на морфологичните подтипове при АК и ПК.

4.12.1. Йерархичен клъстерен анализ.

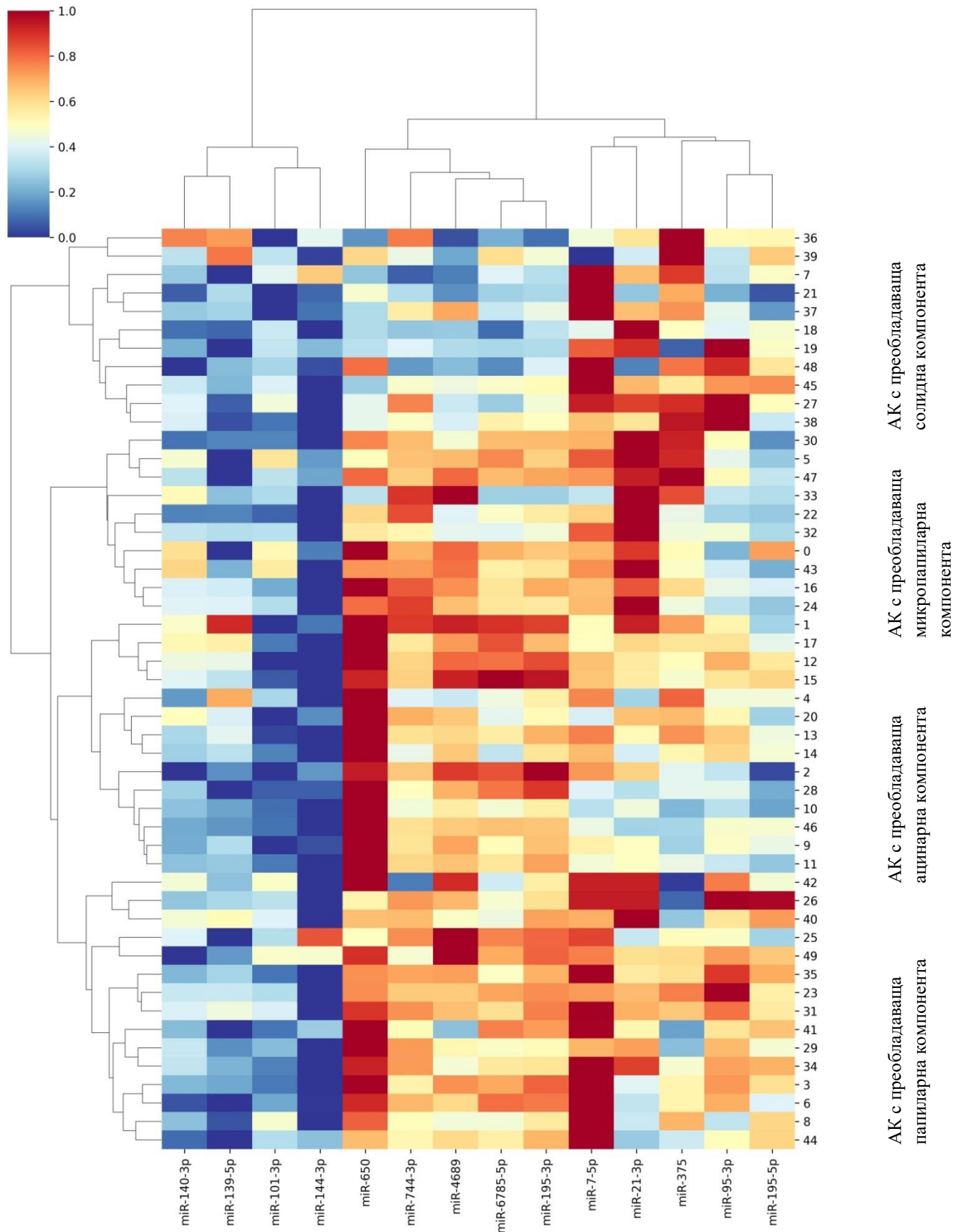
За да се определи дали има разграничаване на морфологичните подтипове при АК и ПК въз основа на микроРНК експресионните им профили бе проведен йерархичен клъстерен анализ и метод на k-най-близките съседи.

При клъстерния анализ се установи, че има добро разделяне на пациентите спрямо морфологичния тип на аденокарциномите и плоскоклетъчните карциноми (Фигури 4.83 и 4.84). При АК има добро разделяне на смесени АК с преобладаване на солидна компонента, АК с преобладаване на микропапиларна компонента, АК с преобладаване на ацинарната компонента и АК с преобладаване на папиларна компонента. Допълнителното разделяне на подклъстери е на база дали туморите се характеризират със слузообразуване или не, диференциацията и TNM стадия.

При ПК се формират два големи клъстера като разделянето е според това дали туморите са без или с вроговяване. Допълнителното разделяне на подклъстери е на основа на TNM стадия и диференциацията на туморите.



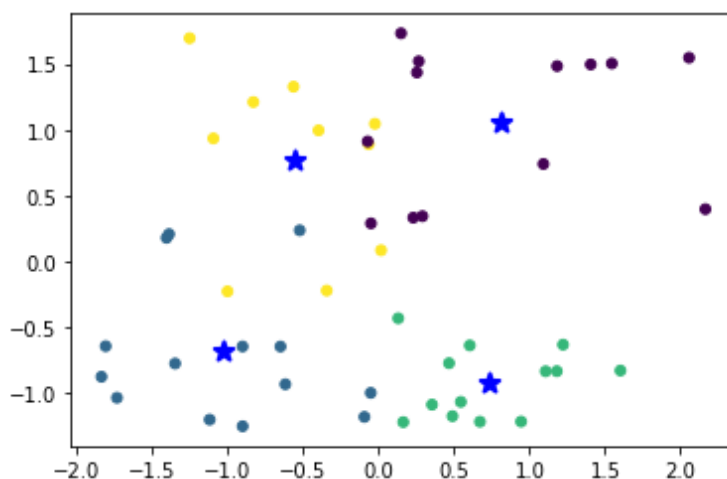
Фигура 4.83. Йерархичен клъстерен анализ въз основа на диференциално експресираните микроРНК-и при туморите от пациенти, диагностицирани с ПК. Използвано е Евклидовото разстояние при клъстерирането, правило на Ward; червено – повишена експресия, синьо – намалена експресия.



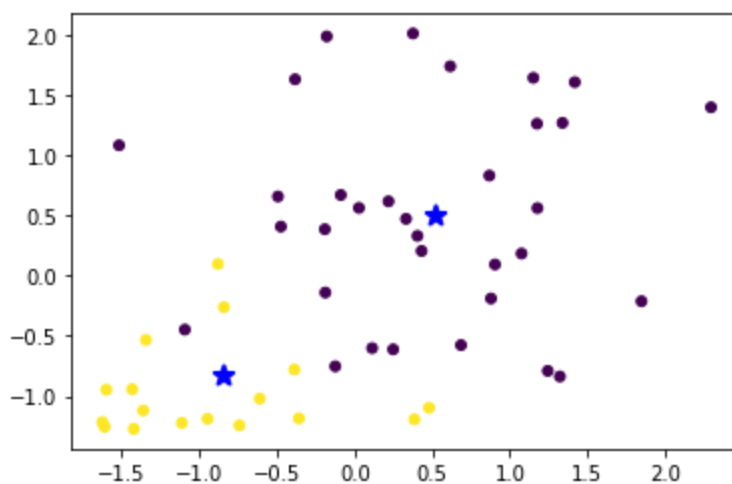
Фигура 4.84. Йерархичен клъстерен анализ въз основа на диференциално експресираните микроРНК-и при туморите от пациенти, диагностицирани с АК. Използвано е Евклидовото разстояние при клъстерирането, правило на Ward; червено – повишена експресия, синьо – намалена експресия.

4.12.2. Метод на k-най-близките съседи.

Проведохме и методът на k-най-близките съседи, за да проверим дали клъстерирането се запазва. И при този анализ наблюдавахме разделяне на два отделни клъстера при ПК и четири клъстера при АК. На Фигура 4.85. и Фигура 4.86. са представени получените резултати при АК и ПК. Със синя звезда са отбелязани центроидите на всеки клъстер.



Фигура 4.85. Методът на k-най-близките съседи въз основа на диференциално експресиранияте микроРНК-и при туморите от пациенти, диагностицирани с АК; светло синьо – солидни, жълто – микропапиларни; светло синьо – смесени; лилаво – ацинарни; синя звезда - центроид



Фигура 4.86. Методът на k-най-близките съседи въз основа на диференциално експресиранияте микроРНК-и при туморите от пациенти, диагностицирани с ПК; жълто – без кератинизация, лилаво – с кератинизация; синя звезда - центроид

4.13. Резултати от проведеното таргетно секвениране от ново поколение.

С цел да се определи соматичния мутационен спектър на пациенти с различните хистологични подтипове на НДБК се проведе таргетно секвениране от ново поколение. Общия брой проби, които бяха включени в анализа бяха 24 туморни проби (12 АК и 12 ПК) от свежо-замразена тъкан, отделена по време на хирургична интервенция. Средната възраст на пациентите, които бяха включени е 60 години (възрастов обхват: 39 – 76 години). Кохортата съставлява 4 жени и 20 мъже, всичките диагностицирани с напреднал плоскоклетъчен карцином на ларинкса. Разпределението според Т стадия при АК беше както следва: Т2 – 7 пациента, Т3 – 3 пациента и Т1 – 2 пациента. Лимфни метастази се откриваха при 7 пациента с АК. Далечни метастази са диагностицирани при 4 пациента с АК. Разпределението на според АК подтипове - 3 АК с преобладаваща солидна компонента, 3 АК с преобладаваща микропапиларна компонента, 3 АК с преобладаваща ацинарна и 3 АК с преобладаваща папиларна. Разпределението според Т стадия при ПК беше както следва: Т3 – 6 пациента, Т2 – 5 пациента и Т4 – 1 пациент. Лимфни метастази се откриваха при 7 пациента с ПК. Далечни метастази са диагностицирани при 5 пациента с ПК. Разпределението на според ПК подтипове - 6 без вроговяване и 6 с вроговяване.

Таргетното секвениране се осъществи посредством TSCAP панел на платформата MiSeq, Illumina (САЩ). От получените данни при АК се установиха 48 варианта в 15 гена като при всички изследвани пациенти се откриват повече от един вариант. При ПК се откриват 42 замени в 16 гена. От получените данни се установиха 12 непубликувани варианта и 80 публикувани мутации в общо 21 ракови асоциирани гени. Новооткритите варианти са в гените *TP53*, *RBI*, *SMO*, *FBXW7*, *STK11*, *GNAS*, *FGFR2* и *ATM* и определени като патогенни посредством положителна предикция за патогенност от повече от четири предиктивни програми. Тези варианти са определени с потенциално патогенен ефект.

В допълнение, бяха установени три полиморфизма в *TP53*, *JAK3* и *EGFR* гените. Вариантът в *TP53* гена е често срещаният полиморфизъм rs1042522, c.215C>G (p.Pro72Arg), имаш отношение към отговора от химиотерапия (236). При 33% от изследваните пациенти (n=4) миссенс вариант rs1042522(C;C) е открит в хомозиготно състояние, докато при 58% (n=8) варианта беше rs1042522(C;G) хетерозиготно състояние при АК. При един пациент се откри честият вариант rs1042522(G;G). При ПК 83% (n=4) вариант rs1042522(C;C) е открит в хомозиготно състояние и при 7% в вариантът rs1042522(C;G) е в хетерозиготно състояние. Вторият неутрален вариант, rs3213409 в гена *JAK3*, c.2164G>A (p.Val722Ile) се наблюдава при двама от пациента с ПК и при един пациент с АК. Синонимният вариант в *EGFR* гена rs1050171 се открива при 75% (n=8) пациента с АК и при всички пациенти с ПК.

При АК всички наблюдавани варианти включват 33 мутации с промяна на смисъла (миссенс) варианта, 9 варианта, променящи рамката на четене, и 1 вариант, засягащ сплайс място. При ПК всички наблюдавани варианти включват 33 миссенс варианта, 5 варианта, променящи рамката на четене, 1 малка три нуклеотидна делеция, непроменяща рамката на четене и 2 безсмислени (nonsense) мутация.

От всички изследвани 12 ПК пациента, 83% носят мутации в *TP53* гена (n=10) като при трима пациента (101T, 130T, 135T) паралелно се откриват по два патогенни варианта в *TP53* гена. Само при един пациент не се откри мутация в *TP53* гена.

От изследваните 12 АК тумора, 75% носят мутации в TP53 гена (n=9), като при четирима пациента (121T, 133T, 134T и 178T) се откриват по два патогенни варианта в същия ген. При трима пациента с АК не се откриха мутации в TP53 гена. Втория най-често мутирал ген е KRAS и при ПК с 42%, следван от мутации в STK11 (33%), KIT (33%), CDH1 (25%) и FBXW7 (17%). В 50% от АК тумори се откриват мутации в ATM и KRAS гените, следвани от GNAS (33%) и FGFR2 (17%).

Таблица 4.14. и Таблица 4.15 и Фигури 4.89 и 4.90 представя всички мутации и новооткрити варианти, които бяха установени при АК и ПК. Използвани база данни: <https://cancer.sanger.ac.uk/cosmic>, <https://www.ncbi.nlm.nih.gov/snp/> и литературни източници.

Таблица 4.14. Генетични варианти открити при АК.

Пациент	Ген	Екзон/ Инtron	Тип на нуклеотидната замяна	Генетична промяна	АК замяна	Реф./База данни
100T	BRAF	15	MS	c.1799T>A	p.Val600Glu	COSM476
100T	KRAS	2	MS	c.35G>T	p.Gly12Val	COSM520
120T	RB1	6	MS	c.602C>G	p.Ala201Gly	
120T	APC	16	FS	c.4060delT	p.Ser1355fs	COSM29331
121T	FGFR2	7	MS	c.887A>G	p.Lys296Arg	(237)
121T	GNAS	8	MS	c.2527T>C	p.Cys843Arg	(238)
121T	TP53	5	MS	c.535C>A	p.His179Asn	COSM44151
121T	TP53	4	MS	c.296C>G	p.Ser99Cys	rs1567556019
133T	FGFR2	7	MS	c.887A>G	p.Lys296Arg	(237)
133T	GNAS	8	MS	c.2527T>C	p.Cys843Arg	(238)
133T	STK11	6	FS	c.837delC	p.Pro281fs	COSM20871
133T	TP53	6	FS	c.599delA	p.Asn200fs	COSM255157
133T	TP53	4	MS	c.296C>G	p.Ser99Cys	rs1567556019
134T	KIT	10	MS	c.1621A>C	p.Met541Leu	COSM28026
134T	KRAS	2	MS	c.35G>A	p.Gly12Asp	COSM521
134T	KRAS	2	MS	c.34G>T	p.Gly12Cys	COSM516
134T	STK11	1	MS	c.223A>G	p.Arg75Gly	rs1599915199
134T	TP53	7	MS	c.707A>G	p.Tyr236Cys	COSM10731
134T	TP53	5	FS	c.465474del CGCGTCCGC	p.Arg156fs	
137T	CDKN2A	2	MS	c.178G>A	p.Ala60Thr	
137T	SMARCB1	2	FS	c.208delA	p.Thr72fs	COSM 3309692
137T	TP53	6	FS	c.599delA	p.Asn200fs	COSM255157
177T	GNAS	8	MS	c.2527T>C	p.Cys843Arg	COSM9233128
177T	KRAS	2	MS	c.34G>T	p.Gly12Cys	COSM516
177T	KRAS	2	MS	c.32C>T	p.Ala11Val	COSM511
177T	TP53	5	MS	c.469G>T	p.Val157Phe	COSM10670
178T	KRAS	2	MS	c.34G>T	p.Gly12Cys	COSM516
178T	KRAS	2	MS	c.32C>T	p.Ala11Val	COSM511
178T	TP53	5	MS	c.469G>T	p.Val157Phe	COSM10670
178T	FGFR3	9	FS	c.1206delC	p.Lys405fs	

178T	TP53	5	FS	c.424430del CTGTGC	p.Pro142fs	
68T	JAK3	16	MS	c.2164G>A	p.Val722Ile	COSM34213
68T	TP53	7	MS	c.743G>T	p.Arg248Leu	COSM16468
8T	GNAS	8	MS	c.2527T>C	p.Cys843Arg	COSM9233128
8T	KRAS	2	MS	c.34G>T	p.Gly12Cys	COSM516
8T	STK11	1	MS	c.142A>G	p.Lys48Glu	rs766776431
90T	ATM	56	MS	c.641-26885A>G	p.Tyr2755His	rs1591594384
90T	CDH1	9	MS	c.1277C>T	p.Thr426Ile	rs876658901
90T	TP53	7	MS	c.731G>A	p.Gly244Asp	COSM10883
95T	KRAS	2	MS	c.35G>T	p.Gly12Val	COSM520
95T	STK11	6	FS	c.837delC	p.Pro281fs	COSM20871
95T	TP53	7	MS	c.731G>A	p.Gly244Asp	COSM10883

Таблица 4.15. Генетични варианти открити при ПК.

Пациент	Ген	Екзон/ Инtron	Тип на нуклеотидната замяна	Генетична промяна	АК замяна	Реф./База данни
101T	CDH1	9	MS	c.1277C>T	p.Thr426Ile	rs876658901
101T	TP53	8	MS	c.824G>A	p.Cys275Tyr	COSM10893
101T	TP53	4	MS	c.332T>C	p.Leu111Pro	COSM44045
130T	TP53	8	SG	c.811G>T	p.Glu271Ter	COSM43750
130T	TP53	5	MS	c.488A>G	p.Tyr163Cys	COSM10808
131T	GNA11	5	FS	c.618delG	p.Gly208fs	COSM1392334
131T	TP53	5	FS	c.465_474del CCGCGTCCGC	p.Arg156fs	
131T	ATM	9	MS	c.1229T>C	p.Val410Ala	COSM21825
131T	KRAS	2	MS	c.34G>T	p.Gly12Cys	COSM516
135T	FGFR3	9	FS	c.1205_1206insC	p.Lys405fs	rs775260779
135T	TP53	8	MS	c.824G>A	p.Cys275Tyr	COSM10893
135T	TP53	4	MS	c.332T>G	p.Leu111Arg	COSM44570
136T	SMO	3	MS	c.628G>T	p.Val210Leu	
136T	KRAS	2	MS	c.35G>T	p.Gly12Val	COSM520
136T	TP53	4	MS	c.332T>G	p.Leu111Arg	COSM44570
136T	STK11	6	FS	c.837delC	p.Pro281fs	COSM20871
152T	FGFR2	7	MS	c.887A>G	p.Lys296Arg	
152T	KRAS	2	MS	c.35G>T	p.Gly12Val	COSM520
179T	CDH1	9	MS	c.1277C>T	p.Thr426Ile	rs876658901
40T	KRAS	2	MS	c.35G>T	p.Gly12Val	COSM520
40T	TP53	10	MS	c.1010G>T	p.Arg337Leu	COSM220135
40T	TP53	6	SG	c.592G>T	p.Glu198Ter	COSM44241
47T	FBXW7	11	MS	c.1654A>G	p.Ile552Val	
47T	TP53	7	MS	c.747G>T	p.Arg249Ser	rs28934571

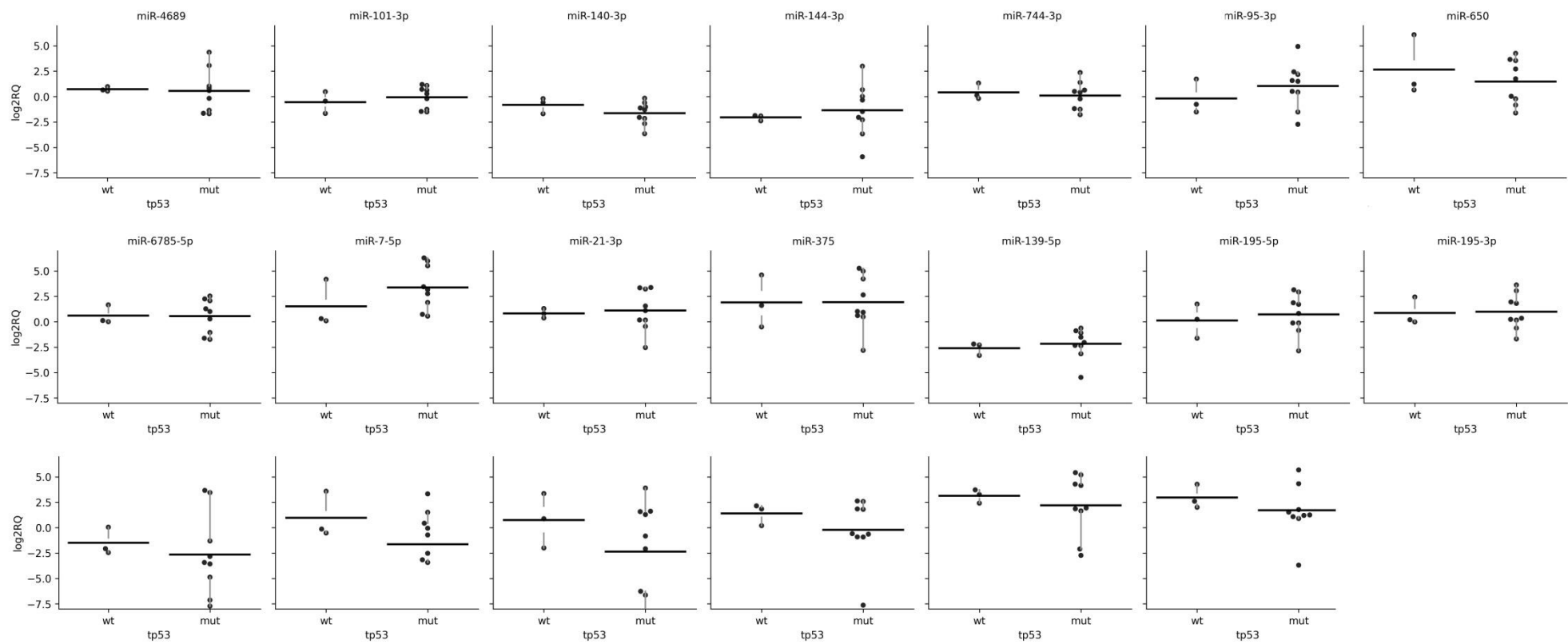
47T	STK11	1	MS	c.170A>G	p.Glu57Gly	
94T	KRAS	2	MS	c.35G>T	p.Gly12Val	rs121913529
94T	CDH1	9	MS	c.1277C>T	p.Thr426Ile	rs876658901
94T	TP53	7	MS	c.743G>T	p.Arg248Leu	rs11540652
94T	JAK3	16	MS	c.2164G>A	p.Val722Ile	rs3213409
98T	PIK3CA	14	MS	c.2132A>G	p.Lys711Arg	
98T	NOTCH1	26	inframe deletion	c.4732_4734delGTG	p.Val1578del	COSM13047
98T	TP53	7	MS	c.707A>G	p.Tyr236Cys	COSM10731
98T	STK11	1	MS	c.175T>C	p.Ser59Pro	rs2080672150
98T	JAK3	16	MS	c.2164G>A	p.Val722Ile	COSM34213
99T	FBXW7	11	MS	c.1654A>G	p.Ile552Val	
99T	TP53	4	MS	c.332T>C	p.Leu111Pro	COSM44045
99T	STK11	6	FS	c.837delC	p.Pro281fs	COSM12924
99T	GNAS	8	MS	c.2527T>C	p.Cys843Arg	

4.14. Връзка между нивата на експресия на микроРНКите и иРНКите и наличието на мутации.

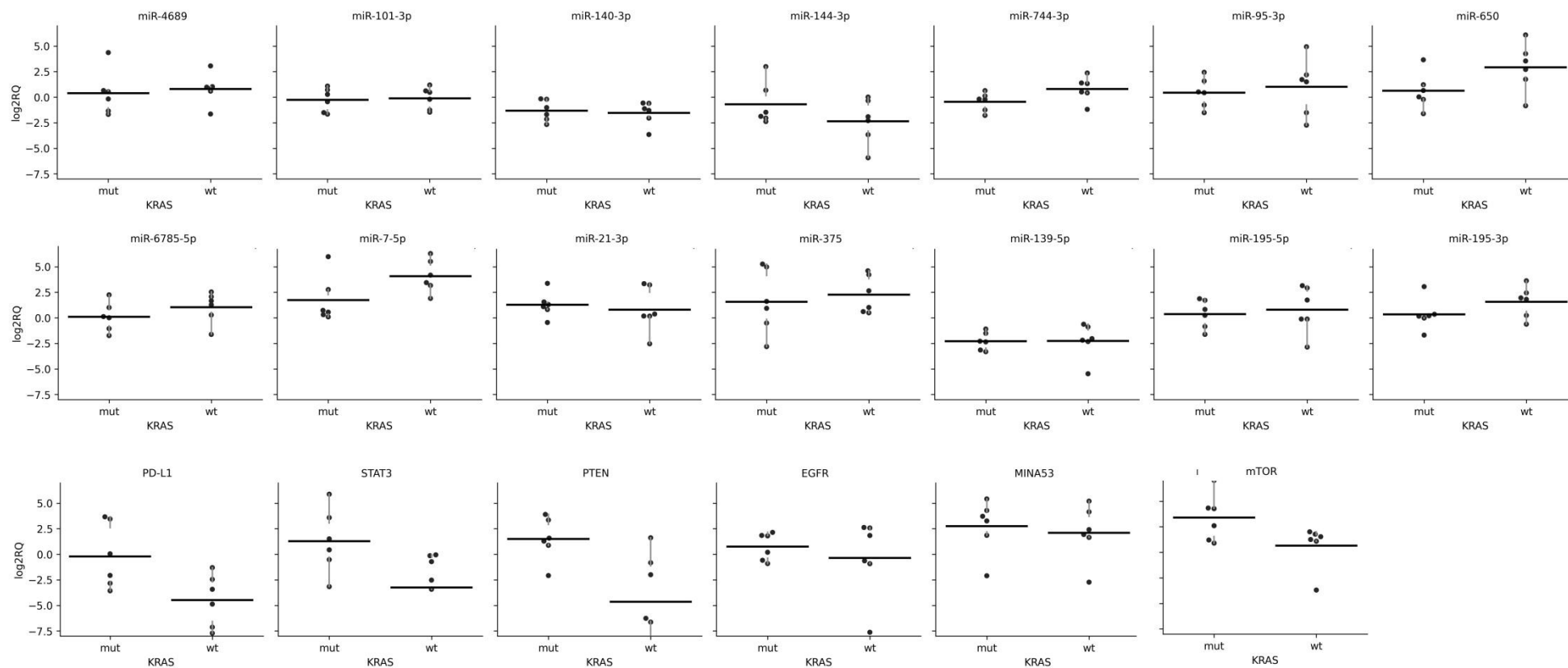
С цел да се изследва връзката между експресията на включените в дисертационния труд микроРНКи, мРНКи и мутационния статус на изследваните белодробни тумори се сравниха нивата на експресия между двете групи пациенти: положителни (пациенти, при които се откриват мутации в даден ген) и отрицателни (пациенти, при които не се откриват мутации в даден ген).

Поради малкия брой изследвани пациенти, статистически анализ не бе възможен. Въпреки това се наблюдаваха тенденции, които за да се потвърдят или отхвърлят, трябва да се валидират в по-голяма извадка.

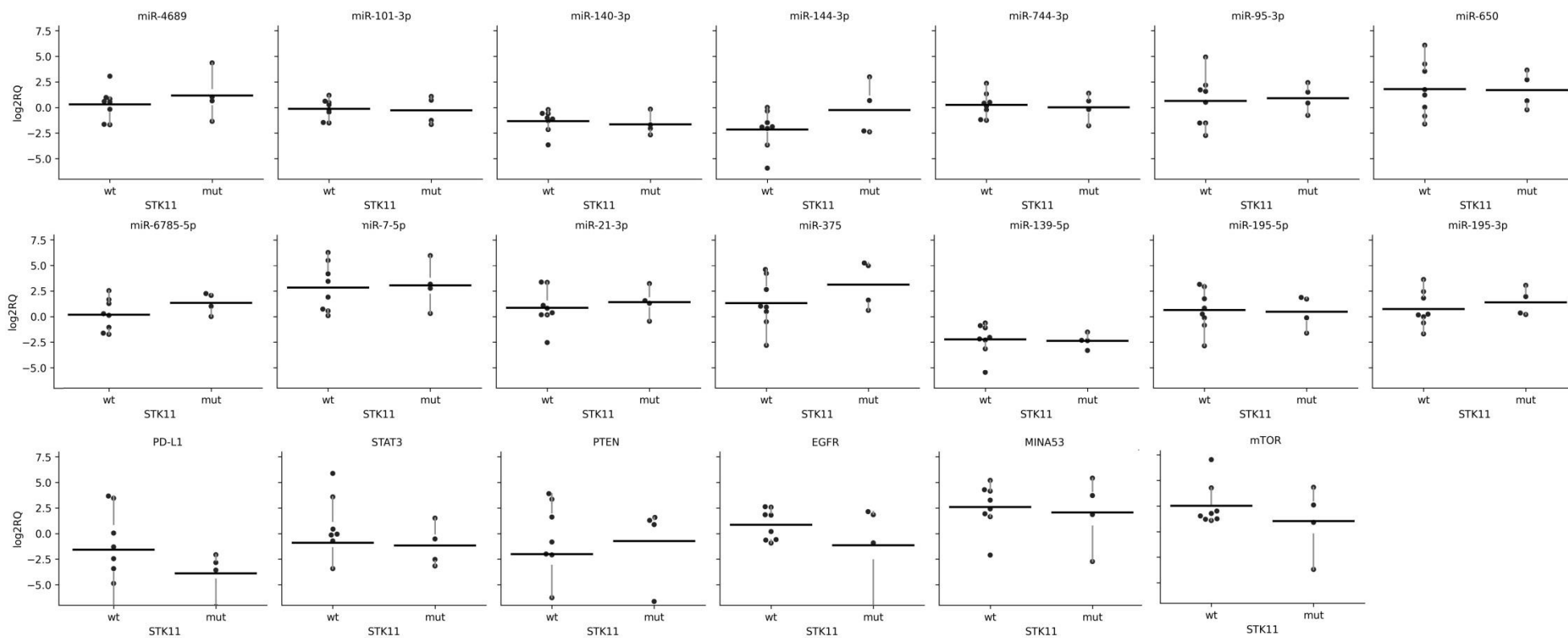
При пациентите с АК носещи патогенна мутация в *TP53* средното ниво на експресия на *miR-650*, *PD-L1*, *STAT3*, *EGFR*, *PTEN* е по-ниско в сравнение с тези пациенти, при които липсва мутация в *TP53* гена. Докато средното ниво на експресия на *miR-7-5p* е по-високо, при положителните по отношение на мутации в *TP53* гена пациенти (Фигура 4.87). Относителната експресия на *miR-650* и *miR-7-5p* беше видимо по-висока при *KRAS* негативните пациенти в сравнение в *KRAS* положителните пациенти, а експресията *miR-144-3p*, *PD-L1*, *STAT3*, *PTEN* и *mTOR* значително по-висока при пациентите с мутация (Фигура 4.88). При *STK11* положителните пациенти средното ниво на експресия на *miR-144-3p*, *miR-6785-5p*, *miR-375*, *miR-195-3p* и *PTEN* е по-високо, сравнение с негативните пациенти (Фигура 4.89).



Фигура 4.87. Експресия на изследваните микроРНКи и иРНКи според мутационния статус на *TP53* гена при АК; wt - пациенти, при който не се откриват мутации, mut - пациенти, при който се откриват мутации

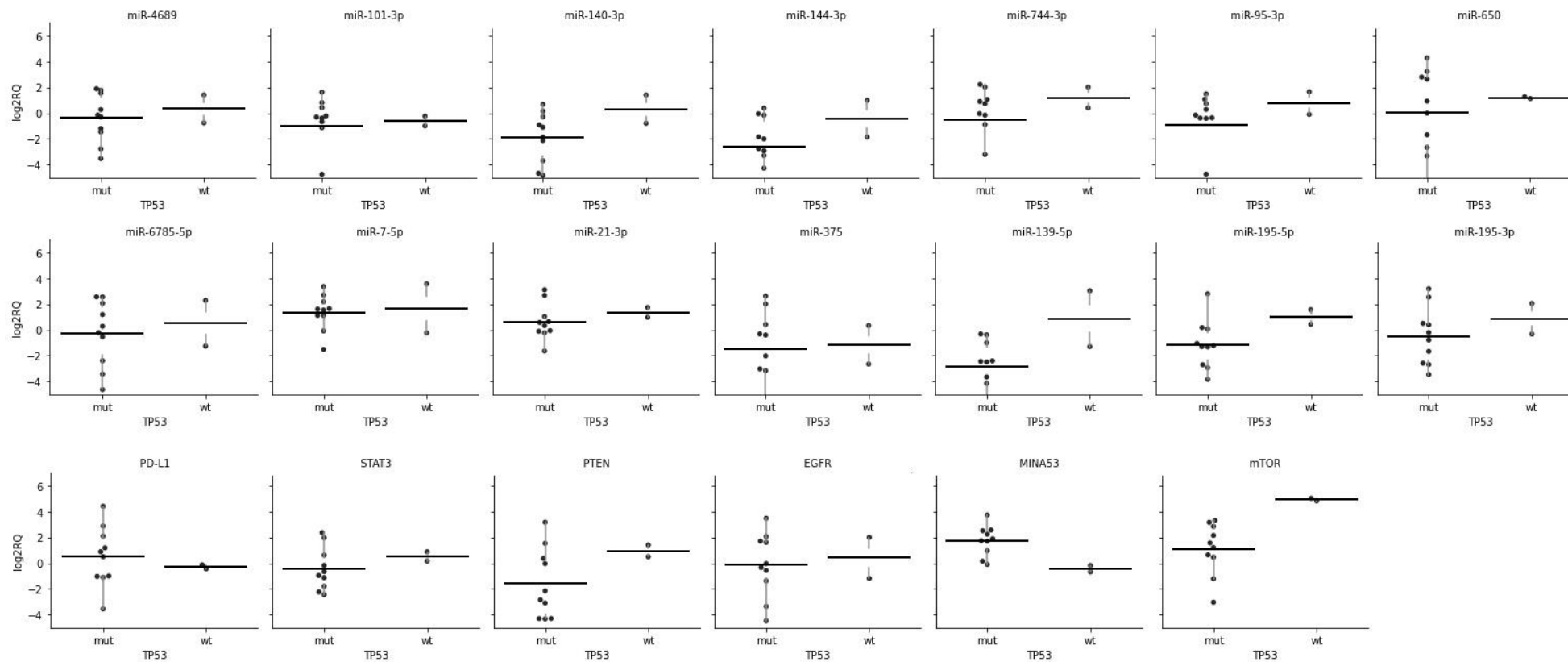


Фигура 4.88. Експресия на изследваните микроРНКи и иРНКи според мутационния статус на *KRAS* гена при АК; wt - пациенти, при който не се откриват мутации, mut - пациенти, при който се откриват мутации

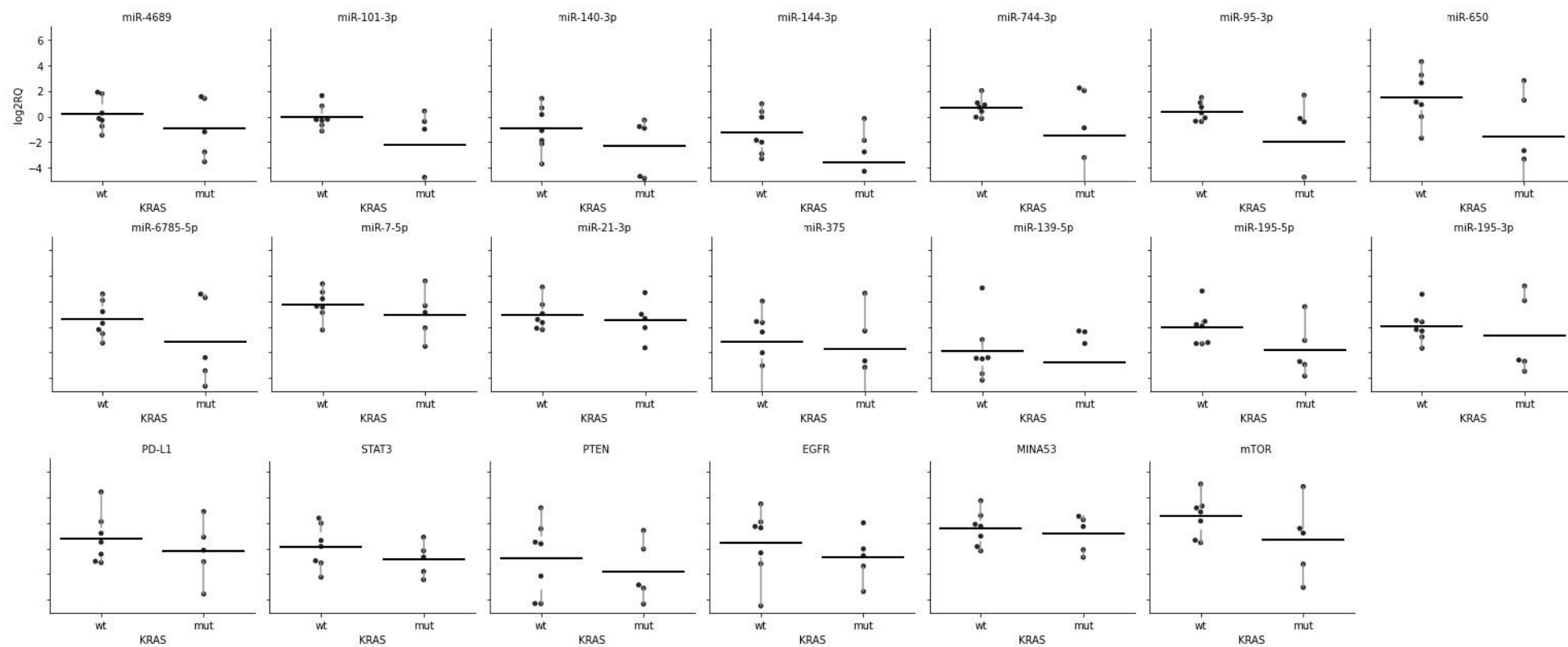


Фигура 4.89. Експресия на изследваните микроРНКи и иРНКи според мутационния статус на *STK11* гена при АК; wt - пациенти, при които не се откриват мутации, mut - пациенти, при които се откриват мутации.

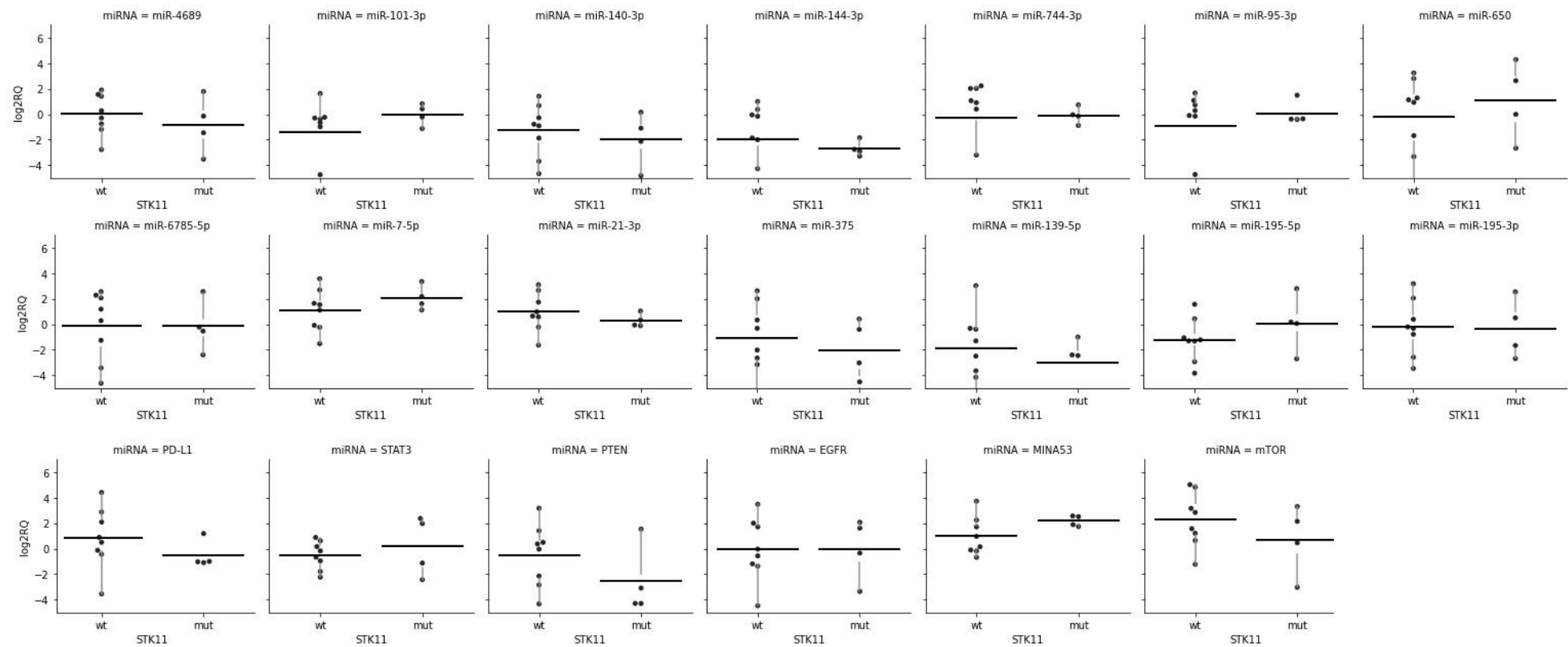
При пациентите с ПК не носещи патогенна мутация в *TP53* средното ниво на експресия на miR-140-3p, miR-144-3p, miR-195-3p, miR-95-3p, miR-195-5p, miR-139-5p, е по-високо в сравнение с тези пациенти, при които се открива мутация в *TP53* гена. Докато средното ниво на експресия на *MINA53* е по-високо при положителните пациенти (Фигура 4.90). Относителната експресия на всички изследвани маркери без miR-21-3p и miR-375 беше видимо по-висока при *KRAS* негативните пациенти в сравнение в *KRAS* положителните пациенти (Фигура 4.91). При *STK11* положителните пациенти средното ниво на експресия на miR-101-3p, miR-650, miR-7-5p, miR-195-5p и *MINA53* е по-високо сравнение със средното ниво на експресия при негативните за мутация в *STK11* гена пациенти (Фигура 4.92).



Фигура 4.90. Експресия на изследваните микроРНК и иРНК според мутационния статус на *TP53* гена при ПК; wt - пациенти, при които не се откриват мутации, mut - пациенти, при които се откриват мутации



Фигура 4.91. Експресия на изследваните микроРНКи и иРНКи според мутационния статус на *KRAS* гена при ПК; wt - пациенти, при който не се откриват мутации, mut - пациенти, при който се откриват мутации



Фигура 4.92. Експресия на изследваните микроРНК и иРНК според мутационния статус на *STK11* гена при ПК; wt - пациенти, при които не се откриват мутации, mut - пациенти, при които се откриват мутации

ЧАСТ 5

ОБСЪЖДАНЕ НА ПОЛУЧЕНИТЕ РЕЗУЛТАТИ

5.1. Обсъждане на резултатите от микроРНК експресионното профилиране чрез микрочипове в туморни и нормални белодробни тъкани.

Ракът на белия дроб е водещата причина за смъртност в света (2). НДРБД съставлява приблизително 85% от всички случаи на рак на белия дроб. Основните хистологични подтипове НДБК – АК и ПК, могат да бъдат допълнително разделени въз основа на техните движещи мутации, профили на експресия и метилиране (239), въпреки че точните механизми на регулиране на генната експресия все още не са установени (240). МикроРНКите участват в редица клетъчни процеси включително имунния отговор и туморната прогресия и могат да бъдат тумор-супресори или онкогени (241). Неотдавна беше показана ролята на микроРНКите в регулирането на специфичните профили на генната експресия при 15 типа епителни тумора от TCGA, включително РБД (242,243).

Аберантната експресия на микроРНКите се счита за важен диагностичен и прогностичен биомаркер при НДРБД (244). Все повече доказателства сочат, че различни генетични промени, включително хромозомни пренареждания, геномни амплификации, делеции, генни мутации и епигенетични промени допринасят за дерегулацията на микроРНКите (245,246). Около 80% от гените за микроРНКи са локализирани в интроните на гените, кодиращи белтъци и тяхната експресия корелира с експресията им (247,248). Гените, кодиращи две добре известни онкогенни микроРНКи, miR-21 и miR-205, са локализирани в регионите, които са често амплифицирани в хода на туморогенезата (249). МикроРНКите от miR-17-92 клъстерът са локализирани в третия интрон на C13orf25 на хромозома 13q31.3. Този регион е често амплифициран при НДРБД, което води до увеличената им експресия и подпомагане на туморната прогресия (250,251). Обратно, микроРНКите от семейството let-7 се намират в региони, които обикновено се делетирани при рак на белия дроб, включително let-7g при 3p21.1-21.2, let-7a-2 при 11q23-q24 и let-7c при 21q11.1 (252).

Аберантна експресия или мутации в гените, кодиращи белтъците, участници в биогенезата на микроРНКите, допринасят за дисрегулацията при НДБК. Szubak и сътр., съобщават, че при НДРБД често се наблюдава увеличена експресия на ензимите DICER и DRISHA (253). В допълнение, фосфорилирането на аргонавт 2 (AGO2) от c-Src инхибира свързването на DICER с AGO2. По този начин се получава промяна в експресионните профили на микроРНКите и се подпомага туморогенезата (254).

В допълнение към генетичните изменения епигенетични промени като аберантно метилиране на ДНК и различни хистонови модификации могат да допринесат за промяна в експресионните профили на микроРНКите (255). Намалената експресия на miR-34a, вследствие на метилиране на промотора на гена е често срещано явление при различни ракови заболявания, вкл. НДРБД (256). В резултат на хиперметиране на промотора за гена на miR-124a значително намалява експресия ѝ и се увеличава експресията на таргетните мРНКи - CDK6 и Rb1 (257). По този начин miR-124a активира медиаторите на клетъчния цикъл, в рамките на пътищата CDK-циклин-Rb1 и участва в злокачествената трансформация и туморната прогресия при рака на белия дроб (258). miR-200c има важна роля в регулацията на епително мезенхимния преход (ЕМП) и инхибира инвазивността и метастатичния потенциал на раковите клетки. Хиперметиране на промотора на гена на miR-200c или метилиране

на хистон 3 (H3K4) значително намалява експресията на miR-200a и в резултат и индуцира химиорезистентност и способността на раковите клетки да метастазират (259,260). Обратното, хиперметирането на промотора на miR-196b-5p води до увеличената ѝ експресия при ПК, таргетирайки тумор-супресорите GATA6 и TSPAN12, като по този начин подпомага клетъчната миграция и пролиферация (261).

Известни са повече от 90 различни проучвания, докладващи различия в експресията на микроРНКите в туморната и нормална белодробна тъкан, повече от 60 проучвания, при които е установено диференциална експресия в серум и около 40, при които е изследвана експресията им в плазма (262). Напоследък в литературата се разграничават две стратегии за идентифициране на диференциално експресираните микроРНКи. Първата стратегия е на базата на получените резултати от широко профилиране чрез експресионен микрочипов анализ или РНК секвениране, изследователските групи избират само малък брой специфични микроРНКи, чиято експресия да бъде валидирана в по-голяма извадка чрез RT-qPCR (263). Друг подход е биоинформатичен анализ на свободно достъпните бази данни. Използвайки различни алгоритми учените идентифицират диференциално експресираните микроРНКи и след това да ги валидират (264).

В ранната статия на Zhang и сътр. 2012 г., микрочипов анализ е използван за изследване на нивата на експресия на микроРНКте в 6 не-дребноклетъчни карцинома. За валидиране на получените резултати, учените използват qRT-PCR в независим набор от 54 плоскоклетъчни белодробни карциноми, 51 белодробни аденокарциноми и съответните им нормални тъкани. Доказано е, че четири микроРНКи (miR-205, miR-93, miR-221 и miR-30e) са силно експресирани при ПК и пет микроРНКи (miR-29b, miR-29c, let-7e, miR-100 и miR-125a-5p) са силно експресирани при АК (265).

В проучването на Hamamoto и сътр. 2013 г., учените идентифицират чрез микрочипов анализ специфични микроРНКи, способни да разграничат АК от ПК. След това се изследва експресията на избрани микроРНКи в нови 88 резецирани туморни и прилежащите им нормални тъкани, чрез qRT-PCR. Три микроРНКи (hsa-miR-196b, hsa-miR-205 и hsa-miR-375) се характеризират с добра способност да разграничават АК от ПК, съответно със 76% чувствителност и 80% специфичност (266).

Hu и сътр., 2019 г. изследват микроРНКа експресионните профили на 24 белодробни аденокарцинома и техните нормални тъкани чрез микрочипов анализ. Всички пациенти са от град Сюанвей, провинция Юннан, където замърсяването на въздуха е повишено и честотата на рак на белия дроб е висока. Идентифицирани са 65 микроРНКи с повишени и 90 микроРНКи с понижена експресия. Авторите избират само първите 4 повишени микроРНКи - miR-135b-3p, miR-485-5p, miR-135b-5p, miR-9-5p и първите четири микроРНКи с намалена експресия - miR-144-5p, miR-486-5p, miR-30a-3p и miR-144-3p), които да валидират чрез RT-qPCR в 30 допълнителни АК и техните прилежащи нормални тъкани проби. Установява се, че иРНК-мишените на изследваните от тях микроРНКи, участват в множество биологични процеси, свързани с рака и сигнални пътища, като cGMP-PKG, MAPK, Hippo сигнален път и др (267).

В допълнение към тези проучвания, други авторски колективи предприемат различен подход, за да идентифицират диференциално експресираните микроРНКи, а именно чрез извличане на данни и биоинформационен анализ на наличните публични набори от данни. Hu и сътр., 2017 г., изследват три набора от данни - GSE19945, GSE25508 и GSE51853. Данните от TCGA съдържат резултатите от микрочипов

анализ, проведен на белодробни плоскоклетъчни карциноми, белодробни аденокарциноми, едроклетъчни белодробен карциноми и съответните им нормални тъкани. Установява се, че miR-375, miR-203 и miR-205 са диференциално експресирани микроРНК и могат да отличават ПК от други подтипове на НДРБД. иРНКите, таргетирани от трите микроРНК засягат най-силно Hippo сигналния път (268).

В своето изследване Sherafatian и сътр., 2019 г. използват машинното самообучение и по-скоро дървото на решенията, за анализ на данните от TCGA от микрочипов анализ на АК и ПК. Двамата тествани модела идентифицират две микроРНК, способни да разграничат двата типа НДРБД – miR-944 и miR-205 и две микроРНК разграничаващи белодробни тумори тъкани от нормални тъкани – miR-183 и miR-135b (269).

В много други предишни проучвания в опит да се идентифицират надеждни биомаркери за диагностични и прогностични цели, са изследвани специфичните експресионни профили на микроРНКите при пациенти с НДКБК (262,270–272). Въпреки това, валидирането на характерни микроРНК в тъкан се оказва трудна задача, която често води до противоречиви резултати, които не могат да бъдат повторени в независими проучвания. Причините за това могат да бъдат различни. От една страна са свързани с разнообразие на клиникопатологичните характеристики на пациентите (туморен стадий, пол, възраст, етническа принадлежност), броя на изследваните туморни проби и вътретуморната хетерогенност на включените проби (273,274). Освен това вариабилността между изследванията се допълва от нееднаквите подходи, използвани в преданалитичната фаза за събиране и съхранение на материалите. Трудност за репродукция на резултатите възниква и от различните платформи и методи, използвани за изследвания на експресията на микроРНКите, ендогенните контроли и статистическия анализ на получените резултати (275). МикроРНКите, за които многократно е потвърждавано в различни проучвания, че са диференциално експресирани и са добри класификатори на подтиповете НДРБД, досега включват miR-21, miR-375 и miR-205 (276–278). Докато микроРНКите, идентифицирани за разграничаване на АК и ПК от нормалната тъкан се различават значително в различните проучвания с малко припокриване (269,279).

В настоящото проучване се насочихме към молекулно профилиране на двата основни подтипа на НДРБД - аденокарцином и плоскоклетъчен рак на белия дроб чрез микрочипов експресионен анализ. При АК установихме, че по-голям брой от микроРНКите с установена промяна са с повишена експресия (68 на брой; Таблица 4.2, Раздел 4), отколкото с понижена (39 на брой; Таблица 4.3, Раздел 4). При ПК установихме 240 микроРНКите с диференциална експресия спрямо нормалните тъкани. Повечето от тях са с повишена експресия (144 на брой; Таблица 4.4, Раздел 4) и по малко с намалена експресия (96 на брой; Таблица 4.5, Раздел 4).

МикроРНКите с повишена експресия при АК показват промяна от 2,06 до 26 пъти. Осемте най-експресирани микроРНК според микрочиповия анализ са: miR -182-5p (24.32), miR-183-5p (26.97), miR-31-5p (24.08), miR-625-5p (19.52), и miR-6834-3p (23.66), miR-650 (14.30), miR-7-5p (14.56) и miR-744-3p (11.38). Степента на промяна на микроРНК-ите с намалена експресия варира в границите -43.55 до -2.07. MiR-1343-5p (-33.57), miR-195-3p (-15.55), miR-3188 (-43.55), miR-4433b-3p (-11.77), miR-4689 (-20.42) и miR-144-3p (-6.21) показват най-ниска експресия според микрочиповият анализ.

МикроРНКите с повишена експресия при ПК спрямо нормалната белодробна тъкан показват промяна от 2.01 до 12.83 пъти. Осемте най-експресирани микроРНКи според микрочиповия анализ са: miR-4785-5p (16.17), miR-6772-5p (16.88), miR-6788-5p (17.62), miR-6857-5p (16.76), и miR-6877-5p (25.80), miR-7515 (22.11), miR-7-5p (14.56) и miR-744-3p (11.38). МикроРНК-ите с намалена експресия показват степен на промяна в интервала от -132.12 до -2.04. С най-ниска експресия са: miR-133b (-76.92), miR-195-5p (-74.25), miR-218-5p (-132.12), miR-30a-3p (-126.85), miR-338-3p (-79.36), miR-451a (-98.50) и miR-101-3p (-17.86).

Прекурсорните микроРНК-и се процесират от DICER и се освобождават като дуплекси miR/miR*. След което, дуплексът се свързва в белтъка аргонавт (AGO) и се образува микроРНКиндуцирания комплекс за заглушаване (miRISC). Едната верига на дуплекса на микроРНКа молекулата селективно се свързва с AGO белтъка и определя специфичността на miRISC комплекса. Другата верига на микроРНКа дуплекса се освобождава от комплекса и се разгражда (280). При хипоксичен стрес EGFR рецепторът потиска узряването на специфични тумор-супресорни микроРНКи чрез фосфорилиране на AGO2. Фосфорилираният AGO2 белтък намалява афинитета си за свързване с DICER и води до инхибиране процесирания на микроРНКите от прекурсорни микроРНКи до зрели микроРНКи при НДРБД (281). Напоследък редица проучвания показват, че и двете вериги са функционално активни и участват в регулацията на транслацията. Освен термодинамичната стабилност, крайният 5' нуклеотид, влияе върху избора на водеща верига и различните AGO белтъци (AGO1–4) имат предпочитания към различни нуклеотиди (282–285). Например AGO1 избира за водеща верига веригата с U в 5' края, докато AGO2 предпочита рА или рU 5' като краен нуклеотид (286).

От установените диференциално експресирани микроРНКи повечето (общо 43) спадат към miR* веригите при АК. Шестте микроРНКи са с увеличена експресия и на двете рамена (miR-5p и 3p) - miR-141, miR-21, miR-31, miR-629, miR-6516 и miR-7. При ПК към miR* веригите спадат общо 90 микроРНКи. Три от микроРНКите са с повишена експресия (miR-6743, miR-6858 и miR-6886) и 13 микроРНКи, при които експресията на miR-5p и 3p е значително намалена - miR-101, miR-126, miR-145, miR-181a, miR-195, miR-199a, miR-23, miR-30a, miR-30b, miR-30c, miR-30d, miR-30e и miR-342.

5.1.1. Обсъждане на резултатите от анализа на основните компоненти (РСА) и клъстерния анализ.

Анализът на основните компоненти беше извършен върху матрица с данни за експресията на микроРНКите. При този метод броят на оригиналните променливи е по-голям или равен на броя на главните компоненти. Главните компоненти са ортогонални и чувствителни към относителното мащабиране на променливите. РСА е добър инструмент, когато се използват няколко вариации вместо стотици хиляди данни (287). Чрез намаляването на размерността се повишава възможността за интерпретация и в същото време се минимизира загубата на информация. При анализа на основните компоненти върху условията (тумор/нормална тъкан) се изчислява стандартната

корелация между всеки вектор на състоянието и всеки вектор на главния компонент (собствен вектор).

Въз основа на микроРНК експресионните профили РСА анализът показва добро разграничаване на АК от нормалните белодробни тъкани (Фигура 4.10А, Раздел 4). Компонент 1 на РСА анализът обяснява 69.1%, компонент 2, обяснява 9.45% и компонент 3, обяснява 3.36% от общото вариране. Както е показано на фигура 4.10Б при ПК имаше 68.68% вариация в първия основен компонент и 10.18% вариация във втория основен компонент и 3.64%, което показва ясно разграничение между пробите от плоскоклетъчните туморни тъкани и нормална тъкан.

Една нормална тъкан не показва добре разделяне от туморните тъкани и отпадна от по-нататъшните анализи при АК. За да намалим евентуален уклон в анализа един плоскоклетъчен тумор и една норма също бяха изключени. В нормалните тъкани, които отпаднаха от анализите е възможно наличието на туморни клетки, тъй като РНК е изолирана от свежа тъкан, а не от микродисектирана тъкан. Освен това е възможно да са настъпили премалигнени промени в прилежащите нормални тъкани, тъй като те са отстранени по време на операция и са близко разположени до тумора. Интересно е, че при РСА анализа, тези тъкани се намират между нормалните белодробни тъкани и туморите. Това наблюдение може наистина да опише премалигнените промени, които са настъпили при тях. Бъдещи проучвания, използващи RT-qPCR и лазерна микродисекция биха помогнали да се разгледат тези проблеми по-подробно. За отпадането на един плоскоклетъчен тумор възможното обяснение е, че в туморната тъкан може да се открива по-голям процент нормална тъкан или строма.

След нормализирането на суровите данни, филтрирането и статистическите анализи се проведе йерархично клъстериране при АК, ПК и нормалните тъкани. Някои алгоритми за клъстериране като йерархично клъстериране не разпределят данните в фиксиран брой клъстери, а създават йерархия на групиране на базата на определена мярка за сходство. Така въпреки че всяка проба е получена от различен индивид и има уникален модел на експресия (Фигура 4.16 и Фигура 4.17), нормалните тъкани са по-сходни една с друга и туморните тъкани са по-сходни като група и се групират в два отделни клъстера. При АК и ПК наблюдаваме добро разделяне на двете групи тъкани и се потвърждават резултатите от РСА анализа.

От ранните етапи на тяхното откриване е известно, че много микроРНКите се експресират тъканно-специфично (288). Следователно не е изненадващо да се открият значително различни профили на микроРНКа експресията между отделните морфологични подтипове АК и ПК. Въпреки това получените резултати от йерархично клъстериране показват, че експресионните им профили имат значителна вариабилност между различните хистопатологични подтипове АК и ПК. Пациентите, диагностицирани в напреднал стадии, нискодиференцирани диференцирани тумори, с наличие на нодални метастази и при двата карцинома се характеризират с различен набор от микроРНКи с повишена експресия. Те се групират отделно от пациентите, диагностицирани в ранен туморен стадий, добре диференцирани или умерено диференцирани тумори, без нодални метастази.

5.1.2. Валидация на избраните микроРНКи чрез RT-qPCR.

В настоящия дисертационен труд се използват два метода за определяне на експресията на микроРНКите. Микрочиповият анализ ни дава глобален поглед върху промяната на експресията и е метод за идентифициране на микроРНКи като потенциални биомаркери. Докато RT-qPCR е метод, чрез който се изследват истинските промени на предикативните и прогностични микроРНКи. И двата метода имат своите предимства и недостатъци. Досега в научната литература няма специфични насоки за анализ и нормализиране на експресията на микроРНКите.

Микрочиповият анализ позволява едновременното проследяване на експресия на всички 2549 зрели микроРНКи. Този метод, базиращ се на хибридикация дава възможност данните да се нормализират чрез прилагане на различни статистически процедури към глобалните хибридикационни сигнали, вместо да се разчита на специфични гени (214). Чувствителността на микрочипове е по-ниска от РНК секвенирането и RT-qPCR, тъй като при този метод няма стъпка на амплификация (289). Измерването също е ограничено от чувствителността на флуоресцентния четец и е възможно получаването на фалшиво негативни или положителни резултати поради неправилна корекция на фоновия шум, крос-хибридикация и последващо нормализиране на суровите данни (290,291). Също така няма единен стандарт за всяка стъпка от анализа, което може да доведе до различни резултати при един и същи експеримент в различни лаборатории, или дори в една лаборатория, но по различно време. Поради тази причина е необходимо резултатите, получени чрез този метод на скрининг за биомаркери, да се валидират чрез по-чувствителен метод (292).

RT-qPCR е „златният стандарт“ за количествено определяне на РНК (293). Характеризира се с по-малък технически шум и по-голяма възпроизводимост от микрочипове. Въпреки това, структурата на микроРНКите молекулите повлиява получаването на надеждни резултати. Комбинацията от малката дължина на зрели микроРНК (~22 нуклеотида) и хетерогенно съдържание на GC, представлява предизвикателство за синтеза на кДНК и дизайна на праймери. Това води до значителна разлика в температурите на топене на различни микроРНК молекули. Също така последователността, която представлява интерес, присъства не само в зрялата микроРНК, но също и в прекурсорите на микроРНКите (290,294). В допълнение членовете на едно семейство микроРНК (например, семейство let-7) обикновено се различават само по един нуклеотид, предимно в 3' края на последователността (295). Въпреки това, точната и надеждна интерпретация на резултатите от RT-qPCR, зависи до голяма степен от използването на подходящи референтни гени за нормализиране, с цел да се елиминират или минимизират небιологичните вариации между тестовите проби (296). В литературата най-често използваните стабилни ендогенни контроли са *RNU6B* и гена за *5S PHK* (297). За да се постигнат точни, възпроизводими и биологично релевантни RT-qPCR данни за микроРНК експресията, в настоящото изследване използвахме като референтен ген *RNU6B*. Експресията на избраните 14 микроРНКи от микрочиповете изследвахме в нови 100 пациента чрез RT-qPCR.

След нормализиране на данните от микрочиповете и RT-qPCR, установихме че експресията на 8 микроРНКи при АК - 3 с повишена експресия и 5 с намалена

експресия, се валидира. При ПК експресията на 10 микроРНК се валидира чрез RT-qPCR - 3 с повишена експресия и 7 с намалена експресия Фигура 4.18, Раздел 4. Въпреки че резултатите от микрочиповете не се валидираха от RT-qPCR за 6 микроРНК при АК и 4 микроРНК при ПК, те имат своето биологично значение и поради тази причина не ги изключихме от проведените анализи.

5.1.3. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ ОТ RT-qPCR за микроРНКите и асоциация с кринико-патологичните характеристики.

5.1.3.1. miR-4689

В научната литература няма много данни за ролята на miR-4689 при онкологичните заболявания. Генът *MIR4689* е локализиран на късото рамо на хромозома 1 (1p36.31) в близост до *tp73*, *NPHP4* и *CHD5* гените. Аберации на късото рамо на хромозома 1 са често срещано явление при много видове ракови заболявания, включително рак на белия дроб, колоректален карцином, рак на стомаха, невробластом, рак на черения дроб, рак на панкреаса и гърдата (298). Според цитогенетичните проучвания 1p36 съдържа различни структурни пренареждания при НДКБД, а микросателитни анализи показват честа загуба на хетерозиготност (LOH) в 1p36 при всички основни подтипове на белодробния рак (299,300).

Според получените резултати, експресията на miR-4689 е значително по-висока при АК с преобладаваща ацинарна компонента и значително по-ниска при АК с преобладаваща солидна компонента (Фигура 4.29, Раздел 4). При ПК експресията ѝ се асоциира с Т-статуса като при Т3-Т4, тя е значително по-висока (Фигура 4.30, Раздел 4).

5.1.3.2. miR-6785

Друга микроРНК, за експресията на която в белодробна тъкан няма данни, е miR-6785-5p. Генът *MIR6785* е локализиран на дългото рамо на хромозома 17 (17q25.1). Експресията на miR-6785-5p е по-висока при напреднал стадий ПК (Фигура 4.30, Раздел 4) и ацинарни АК (Фигура 4.29, Раздел 4). Чести хромозомни промени обхващащи 17q25 са наблюдавани от Sunraweavong и сътр., както при АК, така и при ПК (301). Като този специфичен регион според тях може да съдържа онкогени със значително биологично значение в туморогенезата на НДРБД (301). Също така наблюдават, че пациентите, чиито тумори съдържат амплификация на локуса 17q25 имат по-лоша обща преживяемост от останалите пациенти, допълнително подчертавайки клиничното значение на този регион (301). Chen и сътр., показват че при третиране на клетки от стомашен карцином с розмариновата киселина се повишава химиочувствителността на резистентни клетки от към 5-флуороурацил чрез понижаване на miR-6785-5p и увеличаване на експресията на FOXO4 (302).

5.1.3.3. miR-101-3p

Генът *MIR101* е локализиран на късото рамо на хромозома 1 (1p31.3). Анализът на данните от TCGA показва, че раковите клетки на белия дроб експресират по-малко miR-101-3p, в сравнение с нормалните тъкани като ниската/средната експресия на miR-101-3p се свързва с намалена обща преживяемост (303).

MiR-101-3p, като тумор-супресорна микроРНК, е докладвана при редица онкологични заболявания (304–306). Тя инхибира растежа и индуцира апоптозата на раковите клетки на белия дроб чрез потискане на PTEN/АКТ сигналния път (307). Експресията ѝ е тясно свързана с образуването на туморните метастази, миграцията и инвазията на рак на белия дроб (307). При орофарингеален рак, miRNA-101 инхибира растежа и метастазите на тумора като инхибира експресията на ZEB1 (308). Li и сътр., съобщават, че miR-101-3p инхибира метастазирането и е обратно пропорционално свързана с прогнозата при пациенти с рак на гърдата (309). Освен това е наблюдавано, че свръхекспресията на miR-101-3p потиска инвазия на раковите клетки (310). Друго скорошно проучване при назофарингеален карцином разкрива, че при повишената експресия на miR-101-3p е възможно да се инхибира появата на белодробните метастази (311). Подобно на резултатите от предишни проучвания, в изследваната от нас извадка, miR-101-3p е значително намалена при двата типа и с относително по-ниска експресия при АК, сравнение с ПК (Фигура 4.21, Раздел 4). MiR-101-3p показва значителна асоциация с туморния стадий и при двата подтипа НДРБД, а при АК се асоциира и с наличието на далечни метастази (Фигура 4.24, Раздел 4). Спрямо морфологичния тип АК експресията ѝ е относително по-ниска при туморите с преобладаваща ацинарна компонента, сравнение с туморите с преобладаваща солидна компонента (Фигура 4.29, Раздел 4). При ПК без вроговяване нейната експресия е относително по-висока сравнение с карциномите, характеризиращи се с кератинизация (Фигура 4.37, Раздел 4).

Експресията на miRNA-101 е свързана е с чувствителността на раковите клетки към лъче- и химиотерапия. Xin Meng и сътр. установяват, че miR-101-3p инхибира на BIRC5 и може да сенсibiliзира клетките на АК към облъчване (312). BIRC5 е силно експресиран при АК в сравнение с нормалните белодробни тъкани и се счита за лош прогностичен фактор (312). Според проведен биоинформатичен анализ и луциферазен анализ, BIRC5 е идентифициран като директна цел на miR-101-3p и нивата на експресия на BIRC5 и miR-101-3p са в отрицателна корелация (312). Тяхната експресия има противоположни ефекти върху преживяемостта на пациентите. След облъчване, възстановяването на експресията на BIRC5 и намалената експресия на miR-101-3p значително инхибират пролиферацията и увеличават апоптозата на аденокарциномните клетки (312). Повишената експресия на miR-101-3p може да обърне резистентността към темозоломид чрез инхибиране на гликоген синтаза киназа 3β при глиобластом (313). В друго проучване miR-101-3p инхибира автофагията и засилва индуцираната от цисплатин апоптоза при хепатомни клетки (314). Това предполага, че регулирането на експресията на miR-101-3p и неговата иРНК-мишена BIRC5 може да помогне за разработването на нови радиосенсибилизатори и нови стратегии за лечение на пациенти с РБД.

Освен това, някои дълги некодиращи РНК (lncRNAs) могат да инхибират експресията на miR-101. По този начин индиректно се повишава експресията на крайния продукт от транслацията - белтъците, които иначе биха били регулирани от тази микроРНК. Една такава некодираща РНК молекула е MALAT1 (315). Тя има важна роля при рака на белия дроб и е един от основните регулатори на метастатичния потенциал и туморна инвазия (316). Експресията на MALAT1 е значително повишена при РБД и подпомага химиорезистентността, чрез EZH2. При колоректален рак

експресията ѝ се свързва с лоша прогноза при лечение с химиотерапия, базирана на оксалиплатина (317).

Chen и сътр., установяват чрез имунопреципитация на РНК и луциферазен анализ взаимодействието между MALAT1 и miR-101, между miR-101 и SOX9, директна мишена на miR-101, в клетките на рак на белия дроб (318). MALAT1 споделя почти същото място на свързване на miR-101 с SOX9, което предполага, че MALAT1 може да се конкурира с SOX9 за свързване на miR-101 като по този начин инхибира експресията на miR-101 и насърчава експресията на SOX9 (318). От своя страна SOX9 е транскрипционен фактор на MALAT1 и се свързва с промотора, стимулирайки експресията на тази некодираща РНК молекула (319,320). Освен това SOX9 активира Wnt сигналния път и подпомага прогресията и химиорезистентността на раковите клетки (321,322). Обратната връзка MALAT1-miR-101-SOX9 може потенциално да действа като ефективен терапевтичен кандидат, комбиниран с традиционна цисплатина базирана химиотерапия при НДРБД (318).

5.1.3.4. miR-140-3p

Генът *MIR140* е локализиран на дългото рамо на хромозома 16 (16q22.1) в интрон на WW домена на E3 убиквитин протеин лигаза 2 (WWP2) (248). Към регулаторна последователност нагоре от гена за pre-miR-140 се свързват транскрипционните фактори, NFAT3 и SMAD3 като по този начин се регулира транскрипцията на miR-140 (323). В сравнение с нормалните тъкани, експресията ѝ е понижена при различни ракови заболявания, включително на рак на белия дроб (323). miR-140-3p функционира като туморен-супресор при плоскоклетъчен рак на белия дроб, чрез регулация на BRD9 (324). Също така инхибира клетъчния растеж, инвазия, метастазирание и стимулира апоптозата на туморните клетки като таргетира ATR8A1 (325). ATR8A1 е свръхекспесиран при НДБРД и насърчава експресията на MMP-9 и Vimentin и потиска експресията на E-cadherin. Това води до увеличена способност за инвазия/миграция на не-дребноклетъчните ракови клетки (325).

В нашето проучване, в резултатите от микрочиповия анализ и валидационния RT-qPCR анализ, установихме, че експресионните нива на miR-140-3p са понижени при АК и ПК (Фигура 4.21, Раздел 4). MiR-140-3p е една от микроРНК-а класификаторите, които могат да разграничат ПК от нормалните белодробни тъкани (Фигура 4.20, Раздел 4). Нивата на miR-140-3p имат асоциация с туморния стадий при ПК (Фигура 4.30, Раздел 4).

Song и сътр., показват, че свръхекспресията на miR-140 инхибира клетъчната пролиферация при остеосарком и в клетъчни линии на рак на дебелото черво (326). Повишената експресия на miR-140-3p отслабва резистентността към цисплатина и потиска свойствата, характерни за стволови клетки при АК клетките като повлиява Wnt/ β -catenin сигналния път (327). Yang и сътр., съобщават, че miR-140-5p е значително понижена при хепатоцелуарен карцином и клетъчни линии като свръхекспресия ѝ спира туморния растеж чрез таргетиране на TGF- β и FGF9 (328). Kong и сътр., демонстрират, че miR-140 потиска растежа на тумора и метастазите при недребноклетъчен рак на белия дроб чрез IGF1R (329). Jude и др. показват, че miR-140-3p модулира експресията на CD38 по повърхността на клетките на гладката мускулатура в дихателните пътища клетки като директно се свързва с 3'-

нетранслиращата област на CD38 и индиректно, включва активиране на p38, MAPK и NF-κB (330).

Напоследък се докладват различни кръгови РНК молекули и дълги не кодиращи РНКи, които могат да повлияят на експресията на miR-140-3p, като по този начин намаляват инхибирането на таргетните гени (331). Експресията на circ-AASDH е повишена при АК и корелира с размера на тумора, клиничния стадий и по-лошата прогноза (332). Wang и сътр., показват, че circ-AASDH регулира експресията на E2F7 като таргетира miR-140-3p и допринася за злокачествената прогресия на АК (332).

Относителните експресия на miR-140-3p е по-висока при аденокарциномите с преобладаваща солидна компонента сравнение с другите морфологични типове (Фигура 4.29, Раздел 4). Интересно е да се отбележи, че miR-140-3p е силно експресирана във В-клетките. Различните ѝ нива могат да показват липса на В-клетъчна инфилтрация и да свидетелстват за развитието на различни потискащи механизми в имунната микросреда при различните подтипове АК.

5.1.3.5. miR-144-3p

Генът *MIR144* е локализиран на дългото рамо на хромозома 17 (17q11.2). Двете вериги на miR-144 (miR-144-5p и miR-144-3p) показват намалена експресия в раковите тъкани. Според резултатите, получени от RT-qPCR експресията на miR-144-3p е понижена както при пациенти с АК (Фигура 4.19, Раздел 4), така и при пациенти с ПК (Фигура 4.20, Раздел 4), като при АК тя е значително по-ниска (Фигура 4.21, Раздел 4). Експресията на miR-144-3p се свързва с туморния стадий както в АК, така и ПК ((Фигури 4.22 и 4.30, Раздел 4)). При солидните АК нивото ѝ на експресия е относително по-високо, докато при ацинарните АК, то е най-ниско (Фигура 4.29, Раздел 4). При ПК с кератинизация нивото ѝ е по-ниско от карциномите, при които липсва вроговяване (Фигура 4.37, Раздел 4).

След анализ на данни от TCGA се установява, че ниска експресия на miR-144-3p и повишена експресията на хистон метилтрансфераза EZH2 (333) дава лоша прогноза при пациенти с ПК и допринася за прогресията на белодробния аденокарцином (334). EZH2 е свръхекспресирана в АК клетки и подпомага инвазивността и метастазирането на туморите (335).

miR-144-3p може ефективно да потисне TGF-β1-индуцираната инвазия и адхезия на рак на белия дроб, чрез регулиране на Src-Akt-Erk сигналния път (336). При АК, но не и при пациенти с ПК, серумното ниво на IL-1β е свързано с miR-144-3p и се предполага, че може да повлияе на miR-144-3p на транскрипционно ниво (337). И двата маркера са независими рискови фактори за прогнозата при АК. Следователно, въпреки че IL-1β и miR-144-3p могат да медиират стимулираната от възпаление туморогенеза при пациенти с АК, точният механизъм и тяхното взаимодействие остава неясен (337).

miR-144-3p регулира цисплатиновата резистентност на раковите клетки на белия дроб чрез Nrf2 (338). Последният е критичен транскрипционен фактор на различни антиоксидантни и цитопротективни гени и е свързан с резистентността към противоракови лекарства, вкл. цисплатина (339,340). Активира експресията на гените, свързвайки се с ARE елементи в промоторните им области (341). Освен това, свръхекспресията на Nrf2 може да регулира експресията на miR-144-3p чрез свързване с ARE последователността в промоторната област в *MIR144* гена (338). X-Inactive

Specific Transscript (XIST) е некодираща РНКи, която е свързана с резистентност към цисплатин и таргетира miR-144-3p и понижава нивата ѝ (89).

5.1.3.6. miR-744-3p

Генът *MIR744* е локализиран на късото рамо на хромозома 17 (17p12). Променената експресия на miR-744 е докладвана при множество видове рак. При рака на простатата miR-744 подпомага прогресията на рака като активира Wnt/ β -катенин сигналния път, свързвайки се с SFRP1, GSK3 β , TLE3, NKD1 и LKB1 (342,343). Това подсказва за онкогенната роля на тази микроРНКа. При глиобластом, колоректален рак, хепатоцелуларен карцином и рак на яйчниците miR-744 има тумор-супресорна роля като инхибира пролиферацията, инвазията и предизвиква клетъчна смърт (344–347).

При изследваните пациенти с ПК наблюдавахме значително намалена експресия на miR-744 и повишена експресия при АК в сравнение с нормалните тъкани (Фигура 4.21, Раздел 4). Също така експресията на miR-744-5p е значително намалена при пациентите с метастази (Фигури 4.24 и 4.32, Раздел 4). Тази констатация може да подсказва различната роля на miR-744 в канцерогенезата на двата вида НДКБК. Малко са проучванията, които изследват ролята на miR-744-5p при НДКБК, като основно са на клетъчни линии.

Chen и сътр., установяват значително намалена експресия на miR-744-5p в клетъчни линии на НДКБК, а *in vitro* изследванията показват, че свръхекспресията ѝ може да инхибира клетъчната пролиферация, образуването на колонии и клетъчната инвазия (348). Wang и сътр. илюстрират канцерогенната роля на XIST, чрез таргетиране на miR-744 и увеличаване на експресията на RING1 (349). Като по този начин miR-744 активира Wnt/ β -катенин сигналния път (349). Sui и сътр. установяват, че MAFG-AS1 и MAFG таргетират miR-744-5p и повишената им експресия може да подпомогне туморния растеж при АК (350).

Zhu и сътр. чрез машинно самообучение анализират данните от TCGA на 456 първични НДКБК (набор за валидиране) и 87 формалин-фиксираны парафинови тъкани от НДКБК (набор за обучение) (351). Те стигат до заключението, че повишената експресия на miR-744 е свързано с по-лоша прогноза. MiR-744 увеличава инвазивната способност на клетъчните линии на белодробния рак, докато нокдаунът на miR-744 намалява пролиферацията и миграцията на клетките (351). Освен това miR-744 може пряко и непряко да регулира експресията на c-FOS чрез MAPK сигналния път и да допринесе за канцерогенезата. Те изследват *in vitro* ефекта на miR-744 за повишаване на химиор,- и радиорезистентността на клетките от рак на белия дроб (351).

5.1.3.7. miR-95-3p

Генът *MIR95* е локализиран на късото рамо на хромозома 6 (4p16.1), в интрон 13 на *ABLIM2* (Actin Binding LIM Protein Family Member 2) гена. Установено е, че експресията на miR-95 е променена при рак на стомаха, остеосарком, рак на простатата, рак на дебелото черво, рак на черния дроб, НДРБД и корелира с епително-мезенхимата транзиция при глиобластом (352–356).

В изследваните пациенти експресията на miR-95-3p е повишена при АК, докато при ПК тя е намалена (Фигура 4.21, Раздел 4). Асоциира се с наличието на метастази и при двата типа НДБК (Фигури 4.24 и 4.32, Раздел 4). При АК с микропапиларна

компонента експресията ѝ е относително по-ниска в сравнение с другите морфологични подтипове (Фигура 4.29, Раздел 4). При плоскоклетъчните тумори без кератинизация, локализиращи в десния бронх, при които се откриват лимфни и далечни метастази, експресията ѝ е относително по-висока (Фигури 4.37, 4.36, 4.31, 4.32, Раздел 4)..

Канцерогенезата на НДРБД е многоетапен процес, включващ редица промени в генната експресия и физиологичната структура (357,358). За да метастазират от първичното огнище към съседна тъкан, раковите клетки реорганизируют цитоскелетните си структури. Тъй като актин-свързващите белтъци играят важна роля в полимеризацията на актина, много изследвания са фокусирани върху функциите им в регулирането на метастатичния потенциал на раковите клетки. По време на метастатичните процеси нивото на актин-свързващите протеини се контролира динамично. Намалената експресия на *ABLIM2* води до понижаване на експресията на miR-95-3p в мозъчни метастатични клетки. Това предполага, че промените в нивата на актин-свързващия протеин могат да повлияят на експресията на гените, свързани с метастазите, включително циклин D1, чрез регулация на miR-95 (359).

Освен това miR-95 индуцира пролиферацията, химио- или радиорезистентност чрез директно свързване към *sorting nexin 1 (SNX1)* и сфинголипид фосфатаза 1 (*SGPP1*) при НДРБД (360,361). Доказано е, че SNX1 е тумор-супресор при НДКБК, а miR-95 упражнява своя туморопромотиращ ефект чрез таргетиране на SNX1 (361). Висока експресия на сфингозин-1-фосфат рецептора (S1P) е доказано, че потиска апоптозата, предизвикана от йонизиращо лъчение, чрез активиране на PI3K-Akt сигнализацията. Въпреки това S1P може да бъде дефосфорилиран от SGPP1, получава се сфингозин, за който е известно, че насърчава клетъчната смърт. Следователно намаленото количество SGPP1 дава на клетките резистентен фенотип към химиотерапия и лъчетерапия, чрез регулиране на PI3K-Akt сигналния път (360,362).

5.1.3.8. miR-650

Генът *MIR650* е локализиран на дългото рамо на хромозома 22 (22q11.22) в близост до гена за вариабилната верига на Ig λ (363). Променената експресия на miR-650 е докладвана при рак на стомаха, глиобластом, хронична лимфоцитна левкемия, хепатоцелуларен карцином и се счита, че има предиктивна и прогностична стойност (240,364–366). Интересно е, че miR-650 може да има различна роля при различните видове рак. Например при хроничната лимфоцитна левкемия експресията на miR-650 е свързана с благоприятна прогноза, докато при стомашен карцином miR-650 подпомага прогресията и метастазирането (240,364).

В изследваната от нас извадка, в сравнение със съседни нетуморни тъкани, miR-650 е значително повишена при АК и ПК като значително по-висока експресия се наблюдава при АК в сравнение с ПК (Фигура 4.21, Раздел 4). Асоциира се с аденокарциномите с преобладаваща ацинарна компонента, без нодални и далечни метастази (Фигури 4.29, 4.23 и 4.24, Раздел 4).

В проучването на Tang и сътр., мултивариантните анализи показват, че високата експресия на miR-650 е независим прогностичен фактор за преживяемостта при НДРБД (367). Също така, нивото на miR-650 в тъканите на АК корелира с отговора на пациентите към доцетаксел. Заглушаването на miR-650 може да повиши *in vitro* чувствителността на АК клетките към доцетаксел, докато повишената експресия на

miR-650 намалява чувствителността на клетки както *in vitro*, така и *in vivo* (367). Освен това потискането на miR-650 засилва каспаза-3-зависимата апоптоза и регулира експресията на Bcl-2/Bax чрез потискане към ING4 (240). Инхибиторът на растежа 4 (ING4) е пряка цел на miR-650. Експресията на ING4 е повишена в химиочувствителните АК клетки, а експресията му е обратно пропорционална на miR-650. Също така по този начин се стимулира клетъчната пролиферация и инвазия чрез Wnt-1/ β -catenin сигналният път (367).

5.1.3.9. miR-7-5p

Генът *MIR7* е локализиран на дългото рамо на хромозома 9 (9q21.32). Относителното ниво на miR-7-5p при НДБК е увеличено спрямо нормалните белодробни тъкани (Фигури 4.19 и 4.20, Раздел 4). MiR-7-5p свързва различни сигнални пътища, чрез обратни или директни връзки, участва в диференциацията, регулирането на клетъчната пролиферацията, апоптозата, миграцията, ремоделирането на хроматина и резистентността към множество лекарства (368,369). MiR-7 действа главно като тумор-супресор при различни видове рак, но различни колективи докладват онкогенната ѝ роля при рак на белия дроб (368,370). Повишени нива на експресия на miR-7 в белодробни тумори, в сравнение с нормалните тъкани и метастази са открити в други предишни проучвания (371–373). Двойната функция на miR-7 като онкогени или тумор-супресор, зависи от редица фактори, като например вида на рака и клетките, местоположението на тъканта, разположението на клетките и др. (374)

В изследването на Chou и сътр., EGFR повишава експресията на miR-7 чрез Ras/ERK/Мус път и потиска транскрипционния репресор ERF (370). Интересно е, че EGFR е директна мишена на miR-7 и аномалията в този фино настроен механизъм за регулиране, може да допринесе за онкогенната трансформация. PI3K/АКТ сигналният път също регулира експресията на miR-7, но точният механизъм трябва не е напълно установен (370). Hong и сътр., идентифицира ремоделиращия хроматина протеин SMARCD1 като мишена на miR-7 при РБД. Това води до химиорезистентност на раковите клетки, чрез повлияване на асоциацията на SMARCD1 с p53 (375).

Друга директна мишена на miR-7-5p е PARP1 ген. PARP1 е ядрен ензим, който играе решаваща роля в много биологични процеси, включително в поправката на ДНК и генната транскрипция (376). PARP1 се активира при увреждане на ДНК и катализира полимеризацията на ADP-рибозните единици, което води до събиране на белтъците, отговорни за ДНК поправката към местата, където има на увреждане в структурата на ДНК (377). Съществуват два основни пътя за поправка на двойноверижни скъсвания в ДНК: нехомоложна рекомбинация (NHEJ) и хомоложна рекомбинация (HR). PARP1 има ключова роля при хомоложна рекомбинация, а тя е от решаващо значение за поддържане на геномна цялост и оцеляване на клетката в отговор на химиотерапия (378). Lai и сътр., установяват, че понижаването на експресията на miR-7-5p води до увеличаване експресията на PARP1, което от своя страна засилва хомоложна рекомбинация и причинява химиорезистентност (379).

В изследваните от нас пациенти експресията на miR-7 е по-висока при пациенти без метастази в лимфните възли и показва връзка с наличието на далечни метастази при АК (Фигури 4.23 и 4.24, Раздел 4). Cheng наблюдава, че ниското ниво на експресия на miR-7 е свързано с по-лоша туморна диференциация, напреднал туморен стадий, по-

висока честота на метастази в лимфните възли, напреднал стадий на p-TNM и по-кратка 5-годишна обща преживяемост (380). Оста miR-7/TGF- β играе критична роля в метастазирането на белодробния рак, предизвикано от киселинното рН на туморната микросреда (381). Su и сътр., показват, че ниското рН на туморната микросреда индуцира метастазиране на рак на белия дроб, чрез инхибиране на miR-7-5p. TGF- β 2 е директна мишена на miR-7-5p. Намалената експресия на miR-7-5p ще засили експресията на TGF- β 2, което от своя страна увеличава метастатичния потенциал на белодробния рак (381).

В нашата група пациенти с АК и ПК бяха преобладават карциноми с туморен стадий I-II, като повечето от тях бяха без лимфни възли и далечни метастази. Туморната хетерогенност може също да обясни наблюдаваната експресия на miR-7. В изследваните АК с преобладаваща солидна и папиларна. и компонента (Фигура 4.29, Раздел 4) и при ПК без кератинизация, експресията на miR-7-5p е повишена (Фигура 4.37, Раздел 4).

5.1.3.10. miR-21-3p

miR-21 (miR-21) е една от най-изучаваните досега микроРНКи, имаща плейотропен характер в различни биологични процеси. Според предишни изследвания, miR-21 е докладвана като една от най-често свръхекспресирани микроРНКи при рак на шийката на матката, ларингеален плоскоклетъчен карцином, рак на гърдата и рак на дебелото черво, както и в тъканите на рак на белия дроб (382–384). Нивото на експресия на miR-21 се свързва при НДРБД с тютюнопушене, метастази в лимфните възли и по-лоша прогноза (385,386).

Генът *MIR21* е локализиран на дългото рамо на хромозома 17 (17q23.1), в интрон на *TMEM49* гена. Промоторът е разположен в 10-ия интрон на *TMEM49* гена и има няколко места за свързване с различни транскрипционни активатори и репресори - AP-1, съставен от Fos и Jun семейството, PU.1, C/EBP α , NFI, SRF, p53, p65 и STAT3. MiR-21-3p и -5p се генерират от един и същ прекурсор (pre-miR-21) като TGF- β , BMP EGFR и Notch сигналните пътища повлияват биогенезата на miR-21 (387). Въпреки това, ролята и биологичната функция на съпътстващата верига, miR-21-3p, не са напълно разбрани (387). Съобщава, че miR-21-3p подпомага пролиферацията и инвазивната способност на раковите клетки и повишава резистентността към цисплатин чрез таргетиране на RBPMS, RCBTB1 и ZNF608 (388). Процесите, повлияни от тези гени, включват регулиране на транскрипционната активност чрез SMAD протеини (RBPMS), клетъчен цикъл и транскрипционен контрол чрез хроматиново ремоделиране (RCBTB1), транскрипционна регулация (ZNF608). Последните проучвания съобщават, че свръхекспресията на miR-21-3p може да предизвика микросателитна нестабилност, насочена към MSH2, ключов ген за възстановяване на несъответствие (389). Също така miR-21 има роля в поляризацията на имунните клетки, белодробното възпаление, мускулната контракция, ремоделиране на тъканите, епително-мезенхимната преход (ЕМП) и патологичната фиброза, вкл. белодробната фиброза (387).

Резултати от микрочиповият анализ показват повишена експресия и на двете вериги на miR-21 (Фигура 4.10, Раздел 4). MiR-21-3p беше избран за валидиране чрез RT-qPCR, тъй като беше значително повишена при АК и ПК, включени в микрочиповия анализ. Нивата на експресия на miR-21-3p бяха повишени както при АК,

така при ПК в сравнение с пробите от нормална белодробна тъкан (Фигури 4.19 и 4.20, Раздел 4). въпреки че разликата в нивата на експресия не бяха толкова изразени за при двата карцинома. При АК с преобладаваща микропапиларна компонента (Фигура 4.29, Раздел 4) и ПК с кератинизация (Фигура 4.37, Раздел 4), експресията ѝ е значително повишена, в сравнение с другите морфологични подтипове АК и ПК. При последните експресията на miR-21-3p е значително намалена при ранните туморни стадии, като се наблюдава значително повишаване на нивата ѝ при прогресия (Фигури 4.22, 4.23 и 4.24, Раздел 4).

5.1.3.11. miR-375

Генът *MIR375* е локализиран на дългото рамо на хромозома 2 (2q35). Публикуваните данни показват двойната роля на miR-375 като онкогена или тумор супресорна микроРНК, в зависимост от биологично поведение на туморите (390). При микрочиповия и последвалия RT-qPCR анализ експресията на miR-375 беше повишена при АК и намалена при ПК (Фигура 4.18, Раздел 4). При микрочиповия анализ експресията в аденокарциномните клетки беше повишена с 3,23 FC (Таблица 4.2, Раздел 4), докато при ПК – понижена с 11,7 FC (Таблица 4.5, Раздел 4).

Много предишни проучвания съобщават, че експресията на miR-375 е значително повишена при аденокарцином и дребноклетъчен белодробен карцином, но е намалена при плоскоклетъчен карцином (268,391,392). Експресията на miR-375 се асоциира с нодалния стадий и наличието на отдалечени метастази при АК (Фигури 4.23 и 4.24, Раздел 4). Нивата ѝ са повишени при АК с преобладаваща солидна компонента, сравнение с микропапиларните, ацинарните и папиларните АК (Фигура 4.29, Раздел 4). При ПК, miR-375 се асоциира с туморния стадий като при прогресия, експресията ѝ е повишена (Фигура 4.30, Раздел 4). Намалената експресия на miR-375 е свързана с прогресията на тумора и лошата прогноза при повечето ракови заболявания (390). Дългата не кодираща РНК ROR1-AS1 подпомага метастичния потенциал на клетките и индуцира епителен-мезенхимален преход, чрез таргетиране на miR-375 (393). За разлика от това, свръхекспресията на miR-375 се свързва с локален рецидив при рак на гърдата и рак на простатата (394,395). Освен това miR-375 инхибира индуцираната от IFN- γ експресия на PD-L1 и това е предпоставка за Т-клетките да създадат антитуморна имунна среда (396).

miR-375 играе роля не само в карциногенезата на белия дроб, но има и важна функция като регулатор на ключови патофизиологични механизми като обмяната на въглехидратите (397) и хепатит Б вирусните инфекции (398). Участва в мултиалергичните заболявания индуцирани от интерлевкин 13 (IL-13) (399) и в развитието на диабет (400–402) като се свързва с 3'-UTR на инсулиноподобния растежен фактор 1 рецептор (IGF-1R) (403).

5.1.3.12. miR-139-5p

Генът *MIR139* е локализиран на дългото рамо на хромозома 11 (11q13.4), в интрона на *PDE2A* гена. Интронните микроРНК в повечето случаи се експресират по координиран начин заедно с гените, в които не кодиращи последователности са локализирани (404). Baskerville и сътр., при систематичен анализ на различни тъкани, установяват че miR-139 е една от микроРНКите, показваща най-силна корелация с експресията на *PDE2A* гена (404). Watanabe и сътр., чрез имунопреципитация,

установяват че транскрипцията на miR-139 е заглушена заедно с *PDE2A* ген, чрез метилиране на хистон 3 (H3K27me3) или други хистонови модификации (405). Освен това високо експресиранията некодираща РНК LINC00324 се свързва с miR-139, повишава експресията на IGF1R и стимулира клетъчната пролиферация и инвазия (406).

В изследваните от нас карциноми, експресията на miR-139-5p беше значително намалена както при АК, така и при ПК (Фигури 4.19 и 4.20, Раздел 4). Намалената експресия на miR-139-5p е свързана с АК с преобладаваща микропапиларна компонента (Фигура 4.29, Раздел 4) и наличие на метастази в лимфни възли (Фигура 4.23, Раздел 4) и далечни метастази (Фигура 4.24, Раздел 4), а при ПК се асоциира с туморния стадий (Фигура 4.30, Раздел 4). MiR-139-5p има ролята на тумор-супресорна микроРНК при появата и развитието на НДРБД като получените от нас резултати допълнително подкрепят участието на miR-139-5p в развитието на това злокачествено заболяване.

Диференциалната експресия на miR-139-5p играе важна роля в туморогенезата, метастазите и рецидивите, което предполага, че потенциално може да се използва като предиктивен и прогностичен биомаркер (407). Механизмите на действие на miR-139-5p включват инхибиране на клетъчната пролиферация, миграция и инвазия, чрез таргетна таргетна на MAD2L1, CXCR4 и HDGF (408–410). В друго проучване е показано, че miR-139-5p играе основна роля при рак на белия дроб чрез инхибиране на клетъчната пролиферация, метастази и апоптоза чрез насочване към с-Met и матричните металопроотеинази MMP-7 и MMP-9 (411). Наскоро беше анализирана ролята на miR-139-5p по отношение на чувствителността към цисплатина в клетките на недребноклетъчен рак на белия дроб. miR-139-5p повишава чувствителността към цисплатина чрез свързване към НохВ2 и неговото последващо разграждане (412). Vach и сътр., установяват, че експресията на miR-139-5p е значително понижена в gefitinib-резистентни НДРБД, което предполага, че дисрегулацията на miR-139-5p участва в развитието на резистентен към лекарства рак на белия дроб (413).

5.1.3.13. miR-195-5p и miR-195-3p

Генът *MIR195* е локализиран на късото рамо на хромозома 17 (17p13.1), в интрон 7 на *AKO98506* гена, кодиращ неизвестен функционален протеин (414). Pre-miR-195 дава началото две вериги - miR-195-5p и miR-195-3p. Нарушения в регулацията на miR-195 могат да допринесат за патогенезата на множество заболявания. По-специално, miR-195 упражнява своя ефект чрез свързване към иРНК на множество онкогени или тумор-супресори при няколко вида рак, включително рак на гърдата, рак на белия дроб, рак на простатата, рак на стомаха, колоректален рак, хепатоцелуларен рак и рак на пикочния мехур (415–418). Участва в широк спектър от ракови процеси, включително пролиферация, метастази, инвазия, апоптоза, ангиогенеза и химиочувствителност (419). Механизмите, чрез които miR-195 медира своето действие са изключително сложни и участват в различни сигнални пътища, като Rb-E2F, PI3K/AKT, NF-κB, MAPK/REK, Wnt/β-catenin и Hippo/YAP (419).

В настоящото проучване miR-195-5p беше значително понижено както в АК, така и в ПК (Фигури 4.19 и 4.20, Раздел 4). Докато miR-195-3p нивата са нормални или леко повишени спрямо нормалните белодробни тъкани и ролята на двете вериги е различна при АК и ПК.

При АК експресията на miR-195-5p се асоциира с туморния стадий (Фигура 4.22, Раздел 4) и наличието на далечни метастази (Фигура 4.24 Раздел 4). При ПК експресията на miR-195-5p се асоциира с наличието на далечни метастази (Фигура 4.32, Раздел 4). Нивата са й относително по-високи при АК с преобладаваща солидна и папиларна компонента (Фигура 4.29, Раздел 4) и при ПК, при които липсва кератинизация (Фигура 4.37, Раздел 4).

При АК нивата на miR-195-3p са по-ниски при наличието на метастази – далечни (Фигура 4.24, Раздел 4) и лимфни (Фигура 4.23, Раздел 4), а при ПК експресията й е по-висока при напреднал туморен стадий (Фигура 4.30, Раздел 4) и при туморите, при които има кератинизация (Фигура 4.37, Раздел 4).. Тези разлики в експресията предполагат, че miR-195-3p може да участва в появата на АК и да инхибира пролиферацията. Относително по-високи нива се наблюдават при АК с преобладаваща ацинарна компонента, докато при АК със солидна и микропапиларна компонента, нивата й са по-ниски (Фигура 4.29, Раздел 4).

Повишени нива на miR-195-3p се детектират при бъбречния карцином и се свързва с прогресията на заболяването (420). Понижаването на експресията на miR-195-3p в тъканите на рак на шийката на матката е свързано с неблагоприятна прогноза за пациентите (421). Имунологичният компонент на туморната микросреда действа като нож с две остриета, който може да инхибира или насърчава развитието на тумора (422). Инфилтрираните имунни клетки влияят върху растежа и прогресията на тумора, както и резултата от лечението и прогнозата на пациента (423). Резултатите от анализа на данните от TCGA проведен от Yi Lao и сътр., установяват, че miR-195-3p участва в развитието на АК чрез фагозомния път. Експресията й е свързана с имунна инфилтрация и може да стимулира функцията на В-клетките, дендритните клетки, еозинофилите, незрелите дендритни клетки, макрофагите, мастоцитите, НК-клетките, плазмоцитни, дендритни и фоликуларни хелперни Т-клетки (424). Ниската експресия на miR-195-3p се асоциира с лоша преживяемост при АК, което може да бъде обещаващ прогностичен биомаркер (424).

MiR-195-5p наскоро беше докладвана като туморен супресорна микроРНК при НДРБК и по-ниската й експресия може да действа като диагностичен биомаркер (425–427). MiRNA-195-5p потиска растежа и метастазите на рак на белия дроб чрез регулиране на директната си цел FOXP1 (Forkhead box k1) (428). Функционалните анализи показват, че свръхекспресията на miR-195-5p инхибира инвазията, пролиферацията, миграцията на клетки, чрез директно насочване към CPNE1. Въпреки това механизмите, чрез които miR-195 потиска туморогенезата при НДРБК, не са напълно изяснени. MiR-195-5p регулира прогресията на клетъчния цикъл, миграцията и инвазията чрез потискане на експресията на TrxR2, CEP55, VEGF, MYB, CHEK1, HDGF или IGF1R (139,429–431). А чрез таргетиране на циклин D3, miR-195-5p е доказано, че причинява спиране на клетъчния цикъл във фазата G1 и Survivin индуцира апоптоза и стареене в NSCLC клетки. Интересно е, че BIRC5 гена, който кодира Survivin, е с променена експресия, както в аденокарцинома, така и в тъканите на плоскоклетъчния карцином и високата експресия на BIRC5 е свързана с лоша преживяемост при АК, но не и в ПК (431).

5.2. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ от корелационно-базираните мрежи и йерархичния клъстерен анализ.

Съчетавайки сложността на инициирането, прогресията, метастазирането и хетерогеността на туморите на РБД със сложната регулаторна функции на микроРНК молекулите, е изключително трудно да се предвиди, проучи и анализира ролята на дадена микроРНК. Към днешна дата е установено, че много микроРНКи имат решаваща роля в регулирането на голямо разнообразие от биологични процеси, като клетъчна пролиферация, развитие, диференциация, апоптоза и метаболизъм (432). Задачата, да се анализират множеството иРНК-цели, които трябва да бъдат функционално характеризирани във връзка една с друга, е изключително трудна. Освен това е възможно регулация на една микроРНК чрез друга микроРНК (микроРНК:микроРНК), чрез директни или индиректни взаимодействия (432,433).

При РБД са засегнати редица микроРНКи, а не една и синергичното взаимодействие на множество микроРНКи осигурява възможност за по-добро разбиране на сложните пост-транскрипционни механизми на регулации (434). С наличието на голям обем данни, получени от експерименти с висока производителност и напредъка на компютърните алгоритми, става възможно да се изследват сложните синергични връзки между микроРНКи (435–438). Тези изчислителни методи могат да бъдат разделени на три различни категории: (1) основани на нуклеотидна последователност (439–441), (2) базирани на корелация (442–444) и (3) базирани на причинно-следствена връзка (445).

В настоящия дисертационен труд, използвайки данните за експресията на микроРНКите при АК и ПК, избрахме изграждането на корелационно-базирани мрежи (Фигури 4.40 и 4.41, Раздел 4). Статистическите и математически подходи на мрежовия анализ, екстраполирани в биологичната област, могат да помогнат, да се разберат по-добре молекулярните взаимодействия и синергизма на микроРНКите. Също така да се разбере структурата и динамиката на сложната вътреклетъчна мрежа от взаимодействия, които допринасят за малигнения фенотип на клетката.

На фигура 4.40 е представена корелационно-базираната мрежа при АК. микроРНКите от Клъстер 1, miR-195-3p, miR-4689, miR-650 и miR-6785-5p, са с повишена експресия при АК и преобладаваща ацинарна компонента. Дентрограмата на микроРНКите при проведения йерархичен клъстерен анализ, представен на фигура 4.84 спомага за анализа на корелационно-базираната мрежа при АК.

Големият Клъстер 2 (Фигури 4.40 и 4.85, Раздел 4) е образуван от два подкластера – 2.1 и 2.2. Подклъстер 2.1 miR-95-3p, miR-195-5p, miR-7-5p и miR-375. Експресията им се асоциира с АК с папиларна и солидна компонента, а Подклъстер 2.2 включва микроРНКите miR-139-5p, miR-144-3p, miR-140-3p и miR-101-3p, характерни за АК с преобладаваща солидна компонента. Освен това повечето синергични двойки микроРНК-микроРНК показват еднакви модели на експресия.

Една от характеристиките на графите е степента на централност. Тъй като тя измерва най-кратките пътища между два възела (микроРНКи), които минават през конкретния възел (микроРНК) и показва колко основна е дадената микроРНК в мрежата. Следователно тази микроРНКа, която е с висок резултат за степента на централност е най-важна в мрежата (Таблица 4.8, Раздел 4). За АК Клъстер 1 основната микроРНКа е miR-195-3p, а за АК Клъстер 2 това е miR-95-3p.

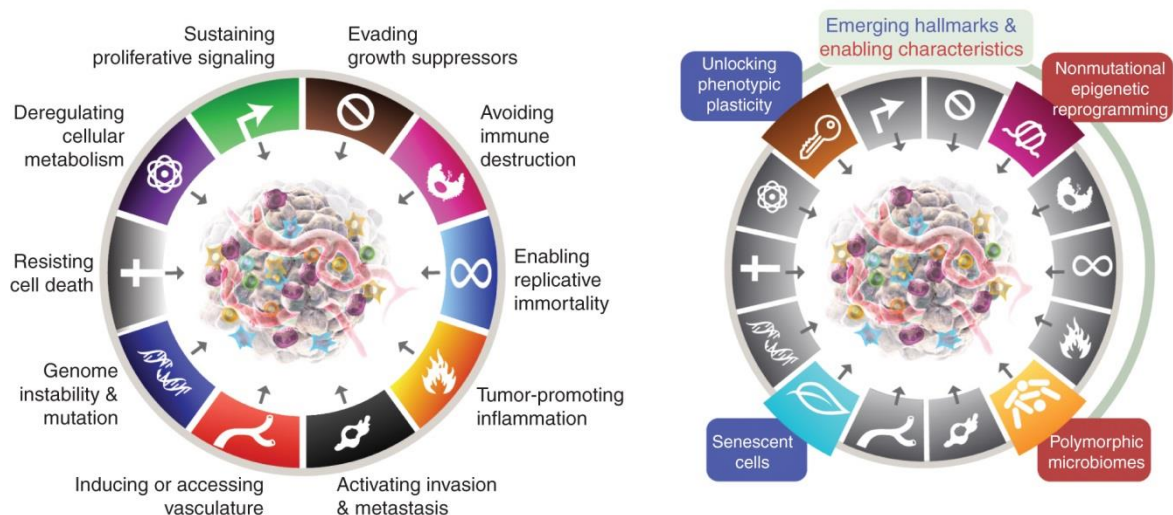
На Фигура 4.41, Раздел 4 е представена корелационно-базираната мрежа при ПК, която не показва ясно изразено клъстериране като при АК. Според дендрограмата на микроРНКите от йерархичния клъстерен анализ на (Фигура 4.83, Раздел 4), микроРНКите се разделят на 2 клъстера, с 2 подкъстера всеки. Първият клъстер включва miR-4689, miR-195-3p, miR-744-3p, miR-6785-5p, miR-7-5p, miR-650, miR-21-3p. Това са микроРНКите, които се асоциират с повишена експресия при ПК с кератинизация или с TNM – статуса. Според корелационно-базираната мрежа miR-21-3p и miR-744-3p имат най-много връзки с останалите микроРНКи, а miR-744-3p е микроРНК молекулата, която има най-висок резултат за степента на централност (Таблица 4. 9, Раздел 4). Доколкото ни е известно, това е първото проучване, което показва значителни промени в експресията на miR-744 в тъкани от белодробен аденокарцином и плоскоклетъчен карцином. Следователно можем да предположим, че нивата на експресия на miR-744-3p и сигналните пътища, които тя регулира, играят важна роля в канцерогенезата на ПК. Въпреки това е необходим допълнителен функционален анализ в увеличен брой проби, за да се установи по-добре биологичната функция и прогностичният потенциал на miR-744 при ПК.

Вторият клъстер от йерархичния анализ включва микроРНКите - miR-375, miR-140-3p, miR-139-5p, miR-95-3p, miR-195-5p, miR-101-3p и miR-144-3p (Фигура 4.83, Раздел 4). Това са микроРНКите, които се асоциират с повишена експресия при ПК без кератинизация или с TNM – статуса. Според корелационно-базираната мрежа микроРНКите с най-много връзки с останалите микроРНКи и най-голямо значение са miR-140-3p и miR-195-5p (Фигура 4.41, Раздел 4).. За да разкрием допълнително потенциалните функции на ключовите микроРНКи при ПК, проведехме анализ на функционалното обогатяване на четирите възлови микроРНКи - miR-21-3p, miR-744-3p, miR-140-3p и miR-195-5p.

5.3. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ проведения онтологичен анализ и анализ на сигналните пътища.

Предполагаме, че изследваните 14 микроРНКи могат както директно, така и косвено да си сътрудничат помежду си, за да регулират своите цели. Поради тази причина изследвахме техните иРНК-мишени. Част от тези иРНК са валидирани в други проучвания при различни малигнени заболявания включително НДРБД. Друга част от тях са установени чрез различни алгоритми и биоинформатични анализи. По тази причина е добре провеждането на допълнително валидиращо проучване на подобрени гени мишени.

Понастоящем осемте отличителни белега на раковите клетки според Hanahan и Weinberg (Фигура 5.1) включват придобити способности за поддържане на пролиферацията, потискане на супресори на растежа, потискане на клетъчната смърт, репликативно безсмъртие, индуциране/достъп до кръвоносни съдове, инвазия и метастазирание, препрограмиране на клетъчния метаболизъм и избягване на имунния отговор (223).



Фигура 5.1. Вляво, понастоящем отличителните белези на рака включват осем отличителни способности и две базови характеристики. В ляво се включват допълнителни отличителни белези и характеристики. Графиката е адаптирана от Hanahan и Weinberg (223).

Въпреки че АК и ПК принадлежат към НДРБД те изглеждат много различни по отношение на прогноза, генната експресия и профили на сигнални пътища. В много проучвания те се разглеждат заедно и се установяват различни повлияни пътища, като е затруднено определянето на най-важните и най-силно засегнатите. За да се разбере молекулната основа на диференциалните биологични механизми на развитие и инвазивност на АК и ПК, както за разработването на по-ефективно лечение, е необходимо те да се разглеждат като отделни клинични единици.

5.3.1. Поддържане на клетъчната пролиферация.

В хода на канцерогенезата хомеостазата на тъканите е нарушена. Много често при раковите клетки се откриват различни по вид мутации в редица растежни фактори, техните рецептори или цитозолни сигнални молекули, участващи в сигналните пътища, водещи до растеж и пролиферация. Тези изменения им дават предимство при оцеляването (446).

Анализът на засегнатите сигнални пътища от микроРНКите от Клъстер 1, асоцииращи се с АК с ацинарна компонента, показва че сигналните пътищата ErbB, RAS, MAPK са значително свръх представени (Фигура 5.1). Рецепторът за епидермален растежен фактор (EGFR), известен също като ErbB1/HER1, е член на семейството на ErbB рецепторни тирозин кинази, което включва също ErbB2 (Neu, HER2), ErbB3 (HER3) и ErbB4 (HER4) (447). Епидермалният растежен фактор (EGF) е един от високоафинитетните лиганди на EGFR. Клетъчната сигнализация чрез EGF/EGFR индуцира растежа, диференциацията, миграцията, адхезията и оцеляването на клетките чрез различни взаимодействащи си сигнални пътища (447). Известно е, че няколко лиганда като епидермален растежен фактор (EGF), трансформиращ растежен фактор алфа, хепарин-свързващ EGF и други, се свързват специфично с EGFR. Други EGFR – асоциирани сигнални пътища са: MAPK сигнален път, PI3Ks, фосфолипаза C (447). Други сигнални пътища, активирани от EGFR, включват сигнализацията чрез FAK, JNK, p38MAPK и ERK5 (447). Като фактор на растежа, диференциацията и

оцеляването на клетките, мутациите и свръхекспресията на EGFR са свързани с онкогенезата на АК с преобладаваща ацинарна компонента. Необходими са последващи валидиращи изследвания за оценка на miR-195-3p, miR-4689, miR-650 и miR-6785-5p като подходящи биомаркери за развитие на таргетна терапия при анализирани белодробни аденокарциноми.

МикроРНКите от Клъстер 2 (АК с преобладаваща папиларна и солидна компонента) също таргетират иРНКи имащи отношение за поддържане на пролиферацията (Фигура 4.42, Раздел 4). Белтъците участват в сигналните пътища на МАРК, лептин, онкостатин М. Напоследък в редица проучвания при НДБРД се открива автокринен лептин, като при АК с по-висока степен на малигненост (АК с микропапиларна и солидна морфология), експресията на лептин е повишена не само в самия тумор, но и в туморната строма (448). Автокринният лептин стимулира пролиферацията на туморните клетки като регулира положително PI3K/AKT/mTOR сигналния път и отрицателно Р53 сигналния път (448). Онкостатин М е член на семейството на цитокините от групата на интерлевкин-6 и играе важна роля във възпалението, имунитета и раковите заболявания (449). Той се секретира главно от Т-лимфоцити, неутрофили и макрофаги. Първоначално се е смятало, че има тумор-супресорен ефект (449). По-късно е открита и неговата роля в прогресията на някои видове рак, подобно на това, което е наблюдавано за IL-6 (450) той стимулира пролиферацията на раковите клетки чрез JAK/STAT сигналния път (451).

По повърхността на клетките може да има два типа инсулинови рецептори (IR), т.е. IR-A (фетална изоформа) и IR-B (зряла изоформа). Докато IR-B участва главно в глюкозната хомеостаза, активирането на IR-A от инсулин или IGF-2 може да доведе до клетъчна пролиферация и туморогенеза (452). Раковите клетки на белия дроб експресират рецептори за инсулин (главно IR-A) и IGF-1R, като инсулинът може да активира IR-A и IGF-1R (453). Още повече, че туморните клетки могат да свръхекспресират IR-A и IGF-1R, а IGF-2 има висок афинитет към IR-A (453,454). По този начин IGF-2, експресиран от ракови клетки, може да упражни автокринен/паракринен ефект, чрез свързване с IR-A и да насърчи още повече пролиферацията им (455).

След свързване с инсулин, инсулиновият рецептор се автофосфорилира и създава докинг места за свързване с IRS белтъците. IGF-1IRS белтъците (IRS1-4) са семейството на адаптери, регулиращи метаболитни и митогенни сигнални пътища (456). Освен, че се свързва с инсулиновия рецептор, IRS1 също се свързва и предава сигнали от рецепторите за пролактин, растежен хормон (GH), лептин, васкуларен ендотелен растежен фактор (VEGF), тропомиозин рецепторна киназа В (TrkB), анапластична лимфомна киназа (ALK), инсулиноподобен растежен фактор (IGF1) и интегрини (457–460). Тъй като тези рецептори индуцират клетъчна пролиферация, оцеляване и миграция, се предполага, че IRS1 може да участва в развитието на рак и метастази (459). В съответствие с това повишената експресия и фосфорилиране на IRS1 са установени в множество клетъчни линии и са свързани с развитието на рак на белия дроб (461). Jiang и сътр., в своето *in vitro* проучване, доказват че инсулинът насърчава пролиферацията, миграцията и инвазията на рак на белия дроб чрез увеличаване на фосфорилирането на IRS1 и активиране на PI3K/AKT сигналния път (462). Четирите микроРНКи при ПК, с най-голямо значение според проведения

мрежови анализ, таргетират иРНКи, участващи в инсулиновия сигнален път (Фигура 4.43, Раздел 4).

5.3.2. Потискане на супресори на растежа.

В допълнение към характерната способност за поддържане на растежа и деленето, раковите клетки трябва също така да заобикалят мощни програми, които регулират отрицателно клетъчната пролиферация. Много от тези програми зависят от действието на тумор-супресорните гени.

От GO термините при АК Клъстер 2 преобладават фокалните адхезии (Фигура 4.42, Раздел 4). Това са големи, динамични белтъчни комплекси, чрез които цитоскелетът на клетката се свързва с извънклетъчния матрикс (463). Те са мултибелтъчни структури, съдържащи интегрин, и образуват механични връзки между вътреклетъчните актинови снопове и извънклетъчния субстрат (464). Регулират се от сложна мрежа от взаимодействия между стотици белтъци, като централна роля имат различни сигнални молекули, включително интегрини и рецептори на растежен фактор, които стриктно регулират клетъчния цикъл и влияят върху оцеляването на туморните клетки (465).

Прави впечатление, че при ПК GO анализът показва че термините са свързани с регулацията на клетъчния цикъл, включващи Rb1 и TP53 сигналните пътища (Фигура 4.43, Раздел 4). Следователно раковите клетки, при които е налице нарушение в тези пътища, следствие на постраскрипционната регулация на белтъците от изследваните микроРНКи, се наблюдава усилена клетъчна пролиферация. RB1 белтъка интегрира сигнали от различни извънклетъчни и вътреклетъчни молекули и в отговор клетката решава да премине към деление (466,467). Докато RB1 трансдуцира сигнали за инхибиране на растежа, произхождащи до голяма степен извън клетката, TP53 сигналния път се активира от вътреклетъчни сигнали (466,468). Такива сигнали може да са степента на увреждане на генома, нивата на нивата на нуклеотидните пулове и глюкозата, сигналите, стимулиращи растежа и наличието на свободни радикали (222).

Известно е, че мутацията на TP53 се наблюдава по-често при ПК, отколкото в АК (469–471). Преобладават безсмислени мутации (nonsense) и резултатът е скъсяване на полипептидната верига (470). Според Wang и сътр., тези мутации водят до повишени нива на експресия на гените, участващи в p53 сигналния път (472). Освен това, P53 белтъка не изпълнява функциите си на тумор-супресор и води до повишен клетъчен растеж и пролиферация в сравнение с АК (473).

5.3.3. Инвазия и метастазиране.

Нарушения във връзката клетка-клетка и клетка-субстрат се асоциира с процеса на епително-мезенхимен преход (ЕМП) и има връзка с миграцията на раковите клетки (463,474). При АК Клъстер 2 повечето от засегнатите белтъците са кадхерини (Фигура 4.42, Раздел 4). Те представляват клас Ca^{2+} зависими трансмембранни гликопротеини, които медируют клетъчната адхезия и играят важна роля в поддържането на клетъчната полярност и адхезия (475). В туморните клетки загубата на този комплекс води до намаляване на взаимодействията клетка-клетка, насърчава отделянето на туморните

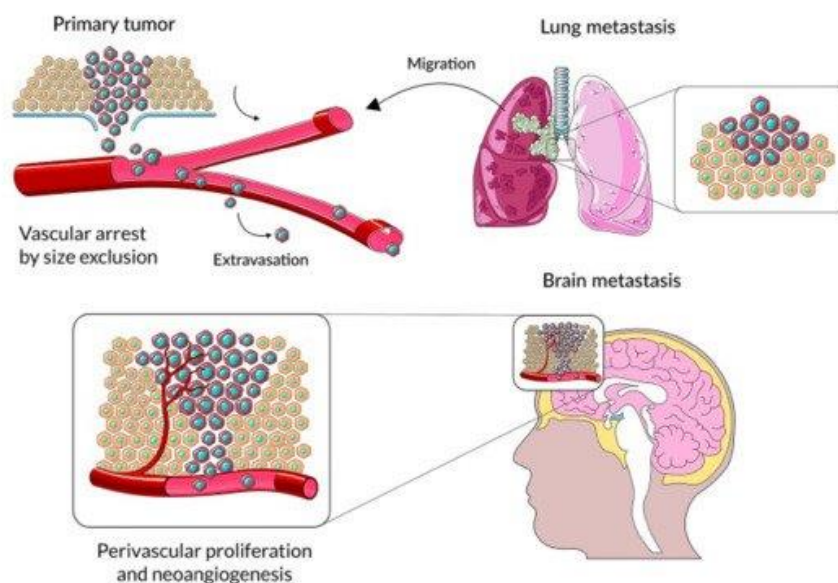
клетки от първичната лезия и увеличава способността им да преминават през базалната мембрана, което благоприятства възникването на метастази (476,477).

Сред кадерините най-важен е епителният E-кадерин (475). Намалената му експресия или промяна в експресията на някой от елементите на комплекса E-кадерин/катенини, води до нарушения в тъканната архитектура и увеличава метастатичния раков потенциал (475). Белтъците от семейството Slug/Snail са репресори на транскрипцията на гена E-кадерин и тяхната експресия се активира от сигналите на фибробластния растежен фактор (FGFR), трансформирания растежен фактор- β (TGF β 1), ErbB1 и ErbB2 (477). Експресията на Slug/Snail води до загуба на клетъчна адхезия и повишена клетъчна миграция, както и до натрупване на β -катенин, който може да функционира независимо или синергично със сигналите на Wnt (478,479).

Действието на β -катенин също се координира от TGF- β сигналният път. Гените, активирани само чрез Wnt сигналният път или от TGF β /Smads и Wnt/ β -катенин, са различни (480). Активностите на транскрипционните фактори Smads и β -катенин синергизират, за да увеличат генната транскрипция и насърчават агресивността на тумора и резистентността към химиотерапия и лъчетерапия (481).

Онтологичният анализ на 11560 иРНК-мишени на микроРНКите от **АК Клъстер 1** показва, че получените GO термините са силно свързани със специфични клетъчни процеси, улесняващи метастазирането в мозъка, клатрин-медирана ендоцитоза, липиден метаболизъм и други (Фигура 4.42, Раздел 4). Въпреки големия напредък през предходните десетилетия в лечението на пациентите с рак на белия дроб към днешна дата прогнозата на пациентите остава лоша. Мозъчните метастази сериозно влошават преживяемостта и качеството на живот на пациентите с РБД като неизвестни генетични и епигенетични фактори са в основата на метастазирането на туморните клетки (482).

Различни проучвания показват, че белодробните метастази проявяват характеристики, подобни на нервните клетки и промените, настъпили в раковите клетки, може да имат силна връзка с мозъчната строма – астроцити и съдове (483,484). Наблюдават се разлики в профилите на генна експресия между мозъчните метастази и първичните тумори. Разлики се наблюдават и при сравняване експресията на мозъчните метастази, регионалните лимфни и други екстракраниални метастази (485). Установява се диференциална експресия на гените, участващи в интравазацията и екстравазацията на туморните клетки, непосредствено свързани със стadiите на метастазирането – инициация, прогресия и вирулентност (485,486). Тези констатации предполагат, че мозъчната строма благоприятства еволюцията на метастатичните туморни клетки и модулира експресионния им профил и туморния фенотип (487). Подобен фенотипен преход се наблюдава при различни видове тумори не само при РБД (488).



Фигура 5.2. Основни етапи на колонизация на мозъчния паренхим от ракови белодробни клетки (489)

След като достигнат до микросъдовете в мозъка туморните клетки обикновено остават в продължение на няколко дни, през които променят формата си, за да се адаптират. Взаимодействието между съдовите ендотелни клетки (ЕК) и метастатичните туморни клетки по време на интравасацията и след екстравазацията все още не е изяснена напълно. ЕК са основни компоненти на триадата, включваща кръвно-мозъчната бариера, медиаторите на съдовото моделиране и мозъчната метастатична ниша (487). Същевременно ЕК са първият клетъчен тип, с който метастатичните туморни клетки се сблъскват по време на екстравазацията и последващата колонизация в мозъка (490). Следователно може да се предположи, че комуникацията между ЕК и туморните клетки играе ключова роля по време на метастазирането и колонизацията на туморните клетки. Повлияна е пропускливостта на кръвно-мозъчната бариера, активира се JAK/STAT3 сигналният път в ЕК и те служат като протекторни клетки, поддържащи растежа на туморните клетки, увеличавайки резистентността към химиотерапия и позволявайки метастазирането на туморните клетки (476,491).

Таргетите на микроРНКите от АК Клъстер 2 показват силно обогатяване BDNF сигналния път (Фигура 4.42, Раздел 4). Мозъчният невротрофичен фактор (BDNF) е член на суперсемејството на невротрофините (нервните растежни фактори) и е важен за фенотипното поведение на невроните, в норма и патология (492). Неотдавна няколко проучвания показаха, че BDNF и неговият рецептор, тропомиозин свързаната киназа В (TrkB), участват в растежа и метастазирането при рак на белия дроб (493,494). Експресията и активността на TrkB са докладвани и при други карциноми, но механизмът на действие още не е напълно изяснен (494–498). Chen и съавтори доказват *in vivo*, че стимулирането на BDNF увеличава активността на STAT3, което от своя страна насърчава синтеза на BDNF (493).

GO анализът на 6508 гена на четирите микроРНКи при ПК показва, че GO термините са силно свързани с ремоделиране на клетъчния цитоскелет (Фигура 4.43, Раздел 4). Цитоскелетните белтъци в клетката действат по координиран начин, за да позволят правилното функциониране на всички клетъчни и биохимични процеси чрез

регулиране на клетъчната структура, организация, трафик и подвижност. По време на злокачествена трансформация, цитоскелетната мрежа бива препрограмирана, за да подпомогне прогресията на ПК, чрез насърчаване на оцеляването, растежа и инвазията на туморните клетки (465).

Цитоскелетът е динамична сложна триизмерна мрежа от нишковидни белтъци, намираща се в цитоплазмата. Състои се от микрофиламенти (съставени основно от актин), междинни филаменти (кератини и виментини) и микротубули (α и β -тубулин) (499). Той има съществена роля за контролиране на формата на клетката като осигурява механична опора, дава възможност за клетъчна подвижност и улеснява вътреклетъчния транспорт и сигнализация (500). Клетъчният цитоскелет служи и като скеле за сигналните каскади, осигурявайки места за локализиране и закрепване на сигналните молекули (499). Освен това цитоскелетната структура има неразделна роля в преобразуването на външни механични сигнали от туморната среда към вътрешни реакции, повлияващи клетъчното поведение (501). Механотрансдукцията представлява комуникационен процес между туморната микросреда (ТМС) и туморите клетки, като ТМС предоставя променящи се механични стимули към механорецепторите на раковите клетки (463,464). Процесът на механотрансдукция преобразува механичните стимули в биохимични сигнали, променяйки генната експресия или ремоделирайки цитоскелета и мембранния трафик, за да промени клетъчните функции (502,503).

ПК на белия дроб се разпознава под светлинен микроскоп по наличието на кератинизация и междуклетъчни мостове (504). Кератините са специфични за епитела междинни филаменти, които имат важна роля в поддържането на структурната цялост на клетките и участват в клетъчната диференциация (465). Кератиноцитите следват уникална програма на терминална диференциация и апоптотична клетъчна смърт, което в крайна сметка води до образуването на кератиновия слой (505). Поради тези причини наличие на кератини в туморите се счита за маркер за добре диференциран плоскоклетъчен карцином на белите дробове (506).

Проучванията върху ПК на главата и шията показват връзка кератинизиращия подтип с по-ниска обща преживяемост, в сравнение с некератинизиращия подтип (507). При РБД малко проучвания разглеждат обща преживяемост по отношение на кератинизацията. Проучването Хiao и сътр., показват, че ИКК α се експресира в ниски нива при кератинизиращи и във високи нива при ПК без кератинизация. Тъй като е установено, че по-високото ниво на експресия на ИКК α гена корелира с добра обща преживяемост при пациенти с рак на белия дроб, кератинизиращият ПК се характеризира с по-лоша прогноза (508). Освен това, сред случаите на некератинизиращ ПК, диференцираният подтип корелира значително с по-висока експресия на ИКК α , спрямо недиференцирания подтип, т.е. кератинизирането на белодробния карцином корелира с намалена обща преживяемост, независимо от диференциацията (508).

Междинните филаменти също имат важна роля и в процеса на ЕМТ, където богатата на цитокератин мрежа се променя с богат на виментин цитоскелет (509). Експресията на виментин е увеличена при РБД, когато клетките се подлагат на мезенхимен преход (510). При миши модели с първичен белодробен тумор, виментин е необходим за формирането на метастази и намалената му експресия се наблюдава при карциноми с по-ниска степен на диференциация (510,511). Виментин също е

необходим за подвижността тумор-инфилтрираните фибробласти, подпомагащи ЕМТ (512).

Потенциално важен аспект в регулирането на туморните метастази е ендцитозата. Проведеният биоинформатичен анализ показва, че GO термините при ПК и АК Клъстер 2 се асоциират с вътреклетъчния транспорт и ендцитоза (Фигура 4.43, Раздел 4). Тя може да бъде пиноцитоза, фагоцитоза или рецептор-медирана ендцитоза, последната от които включва клатрин-, кавеола- и не-клатрин или кавеола-медирана (513). Ендцитозата преминава през няколко вътреклетъчни компартменти за сортиране и разпределяне, завършвайки с лизозомно разграждане, обратно рециклиране към клетъчната мембрана или секреция (514). Множество белтъци, участващи в ендцитозата се характеризират с аберантна експресия при НДРБД и регулират клетъчната миграция и инвазия. Например изоформата на леката верига на клатрина (CLCb) се регулира специфично в клетките на недребноклетъчен рак на белия дроб и повишената ѝ експресия е свързана с лоша прогноза (515). Раковите клетки, експресиращи *CLCb*, показват повишена скорост на клатринзависима ендцитоза чрез *Dynamin 2* (515). Това води до активиране на *Akt/GSK-3 β* сигналния път и до увеличаване на способността за миграция и метастазиране (516). *Dynamin 2* е от решаващо значение за ендцитозата на няколко белтъка, за които е известно, че участват в подвижността и инвазивността на раковите клетки (напр. β -1 интегрин и фокална адхезионна киназа) и свръхекспресията на *Dynamin 2* корелира с лоша прогноза (517–519).

След засегнатите клетъчни органели е апаратът на Голджи (АГ) (Фигури 4.42 и 4.43, Раздел 4). Той взема участие в опаковането, модифицирането и транспортирането на белтъци и липиди, които изпълняват биологичната си функция в клетката или биват секретирани (520). Освен това много проучвания показват съществена роля на апарата на Голджи в клетъчните процеси като вътреклетъчен сензор на стресови сигнали, индуктор на апоптозата и контролна точка на митозата (521). В допълнение към гореизброените му функции, последните проучвания показват, че АГ е медиатор на злокачествена трансформация, вродения имунитет, ангиогенезата, туморната миграция и инвазия (521–523). Освен това връзките между апарата на Голджи с ендоплазмения ретикулум, митохондриите, автофагозомите и ендцитозите, осигуряват предаването на клетъчни сигнали и последвалите ефекторни реакции (522,523). Jung и сътр., въз основа на алгоритми за машинно самообучение и биоинформатичен анализ създават оценка на риска, свързан с експресията на гените на АГ при АК. Установяват, че оценката се асоциира с мутации в *tp53* гена, активирането на имунната микросреда и ефекта от имунотерапията в хода на лечение на пациенти с АК. Jiang сътр., докладват задоволителни резултати при лечението на тумори с помощта на наносистеми за доставка на лекарства, насочени към апарата на Голджи (524).

5.3.4. Геномна нестабилност.

След като организъмът е изложен на ендогенни или екзогенни агенти, увреждащи ДНК, в клетките се задействат пътищата за възстановяване на ДНК, които поддържат генетичната стабилност и цялост. Дерегулацията на пътищата за поправка на ДНК се свързва с възникването и прогресирането на рака. Тъй като основните противоракови терапии, йонизиращата радиация и химиотерапевтичните агенти

предизвикват клетъчна смърт, чрез пряко или непряко причиняване на увреждане на ДНК, дерегулацията на отговора за увреждане на ДНК може да допринесе за свръхчувствителността или устойчивостта на раковите клетки към противоракови терапии (378,525).

Молекулярните механизми, които са в основата на геномната нестабилност са свързани с процесите, контрол на клетъчния цикъл, поправката на ДНК, транскрипцията, репликацията, епигенетичния контрол, ремоделирането на хроматина и сегрегацията на хромозомите по време на митозата. Различните видове генетична нестабилност включват хромозомна нестабилност, микросателитна нестабилност и мутации на базови двойки (526). При анализът на сигналните пътища при ПК се наблюдава десрегулация на клетъчния цикъл чрез промени в гените на туморните супресори *CDKN2A*, *RB1* и *CDKN2B* (Фигура 4.43, Раздел 4). Дефектният контрол на клетъчния цикъл е характеристика на геномна нестабилност, причиняваща хромозомни промени и злокачествена трансформация (527). По време на нормалния клетъчен цикъл може да настъпи увреждане на ДНК молекулата, което се възстановява от специфични клетъчни пътища. Невъзможността да се поправи увреждането на ДНК води до мутации, които се натрупват и в крайна сметка предизвикват канцерогенеза (528). Доказано е че при повече от половината пациенти с НДРБД се откриват мутации в *tp53* гена, докато *ATM* експресията се губи в над 40% от белодробните аденокарциноми (529). Анализът на сигналните пътища при ацинарните АК показва обогатяване на *ATM* зависимата поправка на ДНК (Фигура 4.42, Раздел 4).

5.3.5. Потискане на клетъчната смърт.

ГО анализът на 6508 таргетни гени на четирите микроРНКи при ПК показва, че ГО термините асоциират с апоптозата и автофагията (Фигура 4.43, Раздел 4). Две важни физиологични дейности, които контролират оцеляването и смъртта на клетките и имат много важна роля по време на прогресията на белодробния рак (530).

Автофагия е процес на „самоизяждане“, поглъщане на цитоплазмени белтъци, комплекси или органели в автофагозома (цитоплазмена структура с двойна мембрана). Автофагозомата се включва в късната ендозома, или се транспортира и слива с лизозомата. Така се образува автолизозома, където макромолекулите след това се освобождават и разграждат от киселинни хидролази. Генерираните продукти на разграждане, включително нуклеотиди, аминокиселини, мастни киселини и захари, могат след това да бъдат транспортирани обратно и рециклирани в общия клетъчен метаболизъм (531,532). Автофагията служи в две противоположни посоки: в ранните етапи на карциногенезата, тя функционира като онкосупресор, като потиска възпалението и поддържа стабилността на генома. В късните етапи функционира като защитен механизъм като повишава способността на раковите клетки да се адаптират към условията на клетъчен стрес или повишени метаболитни изисквания като хипоксия и лишаване от хранителни вещества (533). Четирите специфичните за ПК микроРНКи потискат синтеза на белтъци, свързани с автофагията, чрез свързване към целеви мРНК, като по този начин модулират процеса и повлияват чувствителността на ПК към таргетни лекарства, лъчетерапия и химиотерапия (Фигура 4.43, Раздел 4)

Апоптозата е интензивно изследвана в последните десетилетия и е най-добре проучена форма на програмирана клетъчна смърт. Каспазите са повсеместно експресирани цистеинови протеази, които играят централна роля в този процес. Те са

предимно свързани с апоптоза, но някои от тях участват в съзряването на цитокини и вродения имунитет (534). Апоптозата се регулира от вътреклетъчни и/или извънклетъчни сигнали и се характеризира с морфологични промени на клетката, насочена към смърт, които включват ядрена фрагментация и кондензация, пермеабилзация на външната мембрана на митохондриите, свиване на клетките и формиране на апоптотично тяло (535).

Взаимодействието между автофагията и апоптозата оказва влияние върху клетъчната хомеостаза, елиминирането на умиращите клетки, както и върху клиничните терапия. Автофагията може да играе положителна и отрицателна роля за насочването на клетките към апоптоза при НДРБД. Като цяло инхибирането на автофагията ограничава способността на клетките да преодоляват стреса и да поддържат хомеостазата (536). Въпреки това, има случаи, когато автофагията не насърчава клетъчната смърт (530). И двата процеса се активират, когато клетките са подложени на стрес. От една страна, автофагията винаги се инхибира от различни онкопротеини, като АКТ, Р13К, Bcl-1 или мутации в p53, което предотвратява прекомерното разграждане на белтъци в туморните клетки. От друга страна, постоянното активиране на автофагията причинява автофагична програмирана клетъчна смърт или апоптоза (537,538). Взаимодействието на автофагията и апоптозата при НДРБД е сложно и различни методи, като например имунопреципитация и масспектрометрия, биха спомогнали за по-доброто разбиране на взаимодействието между тези два процеса. Става ясно, че miR-21-3p, miR-744-3p, miR-140-3p и miR-195-5p участват във финната регулацията на апоптозата и автофагията при ПК (Фигура 4.43, Раздел 4). Необходими са последващи функционални изследвания, за да се оцени тяхната роля в разработването на нови терапевтични стратегии и преодоляване на лекарствената резистентност.

5.3.6. Индуциране/достъп до кръвоносни съдове.

Подобно на нормалните тъкани туморите се нуждаят от хранителни вещества и кислород, както и от възможност за изхвърляне на метаболитните отпадъци и въглеродния диоксид. Свързаната с тумора неоваскуларна система, създадена в процеса на ангиогенеза, отговаря на тези нужди. По време на туморната прогресия ангиогенният превключвател се активира и остава включен, като кара нормалната васкулатура непрекъснато да образува нови съдове, които помагат за поддържането на разширяващия се неопластичен растеж (539). Ангиогенезата е отличителен белег на рака, свързан с появата, пролиферацията и метастазирането на туморите. Анормалната структура и функция на туморната ангиогенеза улесняват развитието на неблагоприятна туморна микросреда, характеризираща се с повишено интерстициално налягане, хипоксия и ацидоза (222,540). При АК и при ПК процесът на ангиогенеза е засегнат от изследваните микроРНК (Фигури 4.42 и 4.43, Раздел 4).

Хипоксията индуцира експресията на гени, участващи в образуването на кръвоносни съдове и клетъчната пролиферация и по този начин променя ТМС (541). VEGFs, семейство секретирани гликопротеини, играят съществена роля в ангиогенезата на тумора, които включват VEGF-A, VEGF-B, VEGF-C, VEGF-D, VEGF-E, VEGF-F, плацентарен растежен фактор (PlGF) (542). Съществуват три рецептора за VEGF - VEGFR-1, -2 и -3. Ефектът на VEGF за стимулиране на ангиогенезата се

медира главно от VEGFR-2 като няколко ключови тирозинови остатъка (Try801, Try951, Try1175 и Try1214) медираят няколко ключови сигнални пътища: PLC γ -PKC, TSA γ -Src-PI3K-Akt, SHB-FAK-paxillin, SHB-PI3K-Akt и NCK-p38-МАРКАРК2/3 пътища (542). По този начин се насърчават ангиогенезата и съдовата пропускливост, регулира се диференциацията, миграцията, пролиферацията и оцеляването на микросъдовите ендотелни клетки (543).

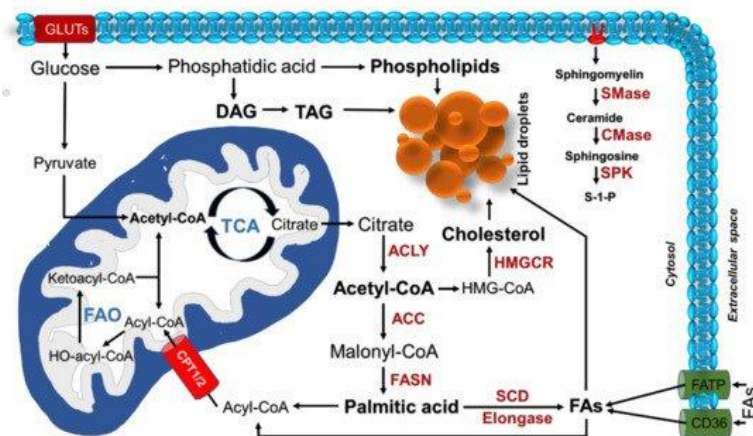
Генът *VEGFA* кодира лиганди, които участват в образуването на нови кръвоносни съдове по време на ембрионалното и постнаталното развитие на ендотелните клетки. Сигнализирането се регулира на множество нива е чрез три рецепторни тирозин кинази (VEGFR-1-3). Експресия на VEGF се регулира както от хипоксия, така и от сигнали на онкогени (544). Освен това лигандите на VEGF могат да бъдат отделени в екстрацелуларния матрикс в латентни форми, които подлежат на активиране от разграждащите екстрацелуларния матрикс протеази (MMP-9) (545). Други проангиогенни сигнали с повишена експресия, като например членове на семейството на фибробластните растежни фактори (FGF), участват в поддържането на туморната ангиогенеза, когато тяхната експресия е повишена (546).

Хипоксията стимулира специфичната експресия на VEGF и VEGFR-2 в туморните клетки и ендотелните клетки около тумора, което прави експресията на им значително по-висока, отколкото в нормалните тъкани (547). Това показва, че инхибирането на туморната ангиогенеза чрез блокиране на сигналния път VEGF/VEGFR-2 може да бъде ефективна стратегия за лечение, включваща: инхибиране на нивото на експресия на гените VEGF и VEGFR-2 чрез антисенс олигонуклеотиди (ASO), РНК интерференция (RNAi) и рибозим (Rz); блокиране на свързването на VEGF и VEGFR-2 на ниво белтък, чрез неутрализиращи антитела (nAbs), разтворим VEGFR-2 (sVEGFR-2); и потискане на растежа на съдовите ендотелни клетки, чрез насочена терапия с VEGF, комбинирана с малки молекули токсични вещества или моноклонално антитяло (548–550).

5.3.7. Препрограмиране на клетъчния метаболизъм.

Една от основните промени, които настъпват в раковите клетки е метаболитното препрограмиране (222). Заедно с ефекта на Варбург и повишеното разграждане на глутамин, липидният метаболизъм се очертава като съществен за развитието и прогресията на туморите. Напоследък се обръща все по-голямо внимание към нарушенията в липидната обмяна, наблюдаващи се при рак на белия дроб (551).

Проведеният онтологичен анализ показва, че липидният метаболизъм е повлиян при АК (Фигура 4.42, Раздел 4). Липидите са хетерогенна група биомолекули, включваща фосфолипиди, мастни киселини, триглицериди и сфинголипиди и имат значителна роля в пролиферацията, инвазията и метастазирането на туморните клетки (552). Раковите клетки показват повишена нужда от липиди, за да отговорят на непосредствените нужди от мембранен синтез, енергийно обезпечаване и да поддържат синтеза на сигнални и биоактивни молекули, подпомагащи клетъчните процеси, растежа на раковите клетки и метастазирането (552).



Фигура 5.3. Метаболизъм на липидите (553).

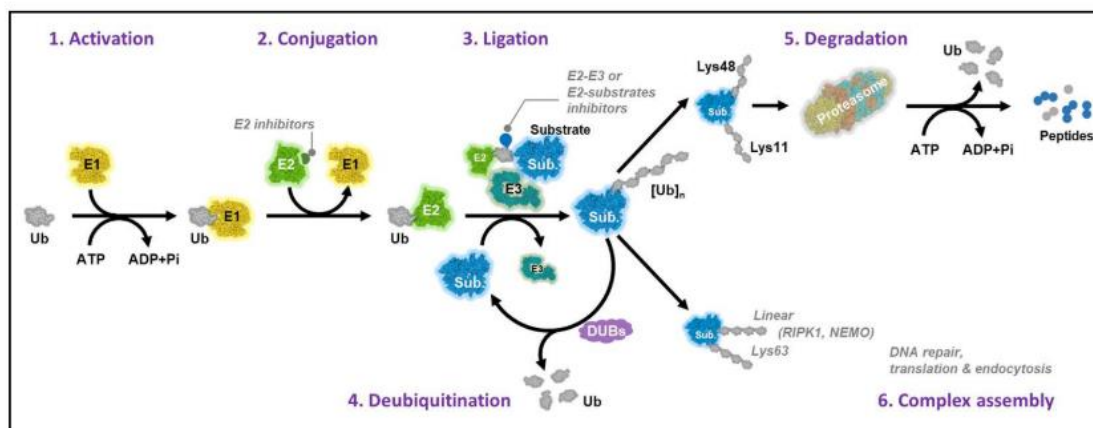
Клетъчните мембрани на нормалните клетки се характеризират с липидна асиметрия между вътрешния и външния слой (554). Дерегулацията на тази асиметрия често се среща при НДРБД и води до отрицателен повърхностен заряд на раковите клетки, корелиращ с по-кисело рН в междуклетъчното пространство (555). Също така различната структура и физични свойства на липидите в мембраната предизвикват промени в нейната флуидност. В зависимост от физикохимичното си състояние мембранните домени могат да бъдат в течно-подредена или течно-неподредена фаза. Едни от основните липиди, изграждащи течно-подредена фаза, са холестерол и сфингомиелин. Те от своя страна могат да образуват липидни рафтове (микродомени) (556). По този начин определят структурата и функцията на трансмембранните белтъци, разположени по клетъчната мембрана. Проучването на Rap и сътр., разкрива механизъм, включващ реекспресията на $ERR\alpha$ и поддържащ резистентността към gefитиниб и озимертиниб (557). Продължителното лечение с тези медикаменти води до натрупване на холестерол в липидните рафтове като по този начин се осигурява взаимодействието на EGFR и Src и повторно активиране на сигналния път EGFR/Src/Erk (557). Увеличената ядрена транслокация на SP1 отключва транскрипцията на $ERR\alpha$ чрез свързване с неговия промотор. Повторната експресия на $ERR\alpha$ предизвиква обезвреждане на активните кислородни видове, което позволява оцеляването на клетките, въпреки постоянното излагане на EGFR-TKIs (557).

Липидният метаболизъм също така има важна роля за туморната микросреда и имунотерапията (558). Туморните клетки произвеждат големи количества мастни киселини *de novo*, а обогатената с мастни киселини туморната микросреда засяга функцията на ефекторните Т-клетки и макрофагите, водейки до имunosупресивно действие и избягване на тумора от действието на имунната система (559,560). Lin и сътр., изследвайки стомашен аденокарцином, откриват че инфилтрираните Т-лимфоцити не използват глюкоза, а разчитат на окислението на мастни киселини за покриване на енергетичните си нужди (561). Също така те наблюдават, че раковите клетки и Т-лимфоцитите се конкурират за наличните мастни киселини и недостига им предизвиква апоптоза на Т-лимфоцитите (562). Блокирането на PD-L1 може да регулира метаболизма на Т-клетките, насърчавайки усвояването на липидите и допълнително да подобри антитуморната имунна способност (561). Cheng и сътр., предоставят доказателства, че натрупването на мутации в гените на белтъците, участващи в

липидния метаболизъм е свързано с удължена обща преживяемост при пациенти, лекувани с имунотерапия (563).

Убиквитинирането и фосфорилирането при злокачествени заболявания представляват двете най-важни посттранслационни белтъчни модификации (564). Все повече проучвания сочат, че убиквитинът играе критична роля в канцерогенезата и прогресията на РБД (565). Онтологичният анализ показва обогатяване на тези биологични процеси при АК и ПК (Фигури 4.42 и 4.43, Раздел 4).

Убиквитин-медираната система за разграждане е отговорна за контролирането на различни процеси, подпомагащи туморната прогресия, ДНК репарацията, спиране на клетъчния цикъл, рециклиране на рецептори на клетъчната мембрана, вътреклетъчен трафик, клетъчна пролиферация, апоптоза, ангиогенеза, миграция и инвазия, метастази и лекарствена резистентност. Два вида ензими, а именно E1/E2/E3 лигази и деубиквитирази регулират убиквитинирането на белтъците в клетките, контролирайки функцията и стабилността им (566).



Фигура 5.4. Механизъм на убиквитинация. (1) Активиране: убиквитинът (Ub) се активира от убиквитин-активация ензим (E1) по АТФ-зависим начин. (2) Конюгиране: убиквитин-конюгиращият ензим (E2) прехвърля активирания Ub към убиквитин-лигазния ензим (E3). (3) Лигиране: лигазата E3 се свързва с белтъци и отговаря за прехвърлянето на убиквитина към целевия субстрат. (4) Деубиквитиниране: убиквитиновите молекули могат да бъдат отстранени от субстратните протеини чрез действието на деубиквитиразните ензими (DUB). В зависимост от резултата от убиквитинирането на белтъка съдбата на субстратите ще бъде различна. (5) Каноничното свързване на K48 води до разпознаване от 20S протеазома, което предизвиква разграждане на субстрата, като се получават малки пептидни фрагменти и свободен убиквитин. (6) Могат да се образуват неканонични типове. Тук е показана връзка K63, която не се разпознава от протеазома, което води до променена функция на субстратния протеин, като например възстановяване на ДНК, трансляция и ендоцитоза (567).

Убиквитинът е малък 76-аминокиселинен белтък, който може да се свърже със специфични целеви белтъци в различни форми. E1/E2/E3 лигазите са отговорни за добавянето на убиквитин в целевите белтъци. Седемте лизинови остатъка в убиквитина осигуряват различни видове модификации, включително моноубиквитиниране, полиубиквитиниране и разклонено убиквитиниране (568). Моноубиквитинирането на белтъците влияе върху активността на ДНК репарацията, трансляцията, възпалението, регулацията на генната експресия, вътреклетъчния транспорт и ендоцитозата (569). Полиубиквитинирането и разклоненото убиквитиниране (в комплекса с APC/C) на белтъци, често води до разграждане от 26S-протеазомозависимо разграждане (568) .

Семейството на убиквитин-специфични пептидази (USP) е клас ензими с белтъчна деубиквитинираща активност, включваща по-голямата част от белтъчното деубиквитиниращо разнообразие в клетките (570). Много проучвания разглеждат тумор-супресорната или онкогена роля на различните USP при РБД. Все по-голям брой проучвания показват, че USP са критични за прогресията на рака и някои USP са използвани като мишени за разработване на инхибитори и служат като прогностични фактори (566).

Съществената роля на UPS в контрола на туморогенезата предизвиква голям интерес към разработването на протеазомни инхибитори като противоракови лекарства. През последните години беше постигнат огромен напредък в разбирането на потенциала на медните комплекси като противоракови лекарства чрез инхибиране на множество компоненти и регулатори на UPS, като 20S протеазома, 19S DUBs и NPLOC4 (571). Прави впечатление, че при анализа на сигналните пътища при ПК 40 от белтъците участват в хомеостазата на меда. Тя е незаменим микроелемент и участва в широк спектър от метаболитни процеси, включително аеробно дишане, пигментация, С-крайното амидиране на белтъци, биосинтеза на катехоламини, транспорт на желязо, действието на супероксид дисмутацата и биосинтеза на извънклетъчния матрикс. При редица видове тумори се наблюдават аберантно повишени нива на мед, което подпомага прогресията на тумора чрез увеличаване на клетъчната пролиферация и стимулиране на ангиогенезата и метастазирането (572,573).

5.3.8. Избягване на имунния отговор.

По време на всички етапи на канцерогенезата, непрекъснатата паракринна комуникация между раковите и стромалните клетки създава богата и динамична микросреда. VEGF има ключова роля в потискане на противотуморния имуен отговор. Инхибирането на VEGF/VEGFR може да препрограмира ТМС от имуносупресивна в имуностимулираща, като модулира функцията на имуносупресивните клетки и Т-клетките. Ето защо анти-VEGF/VEGFR терапията има не само антиангиогенен, но и подпомагащ имунния отговор ефект (574). Освен това лечението с анти-VEGFR2 води до секреция на IFN- γ от CD8⁺ Т-клетките, който индуцират експресията на PD-L1 върху туморните клетки. Тези механизми осигуряват теоретична основа за комбинирано лечение на напреднал НДРБД с имунотерапия и антиангиогенни агенти (575).

Трансформираният растежен фактор бета (TGF- β) е многофункционален цитокин, който има роля в пролиферацията, диференциацията и морфогенезата (576). Семейството TGF- β лиганди включва над 30 белтъка, а съответните им рецептори са серин - треонин киназни рецептори. Свързването на лиганда води до фосфорилиране на рецептора тип I от тип II и предаване на сигнала надолу по веригата чрез транскрипционните фактори SMAD, като се съобщава и за неканонична TGF- β сигнализация, която е независима от активирането на SMAD (577,578). Нарушения в TGF- β сигналния път водят до неопластичната трансформация, пролиферацията на туморните клетки, ЕМТ и имуносупресия (579). Един от механизмите, чрез които TGF- β води до имуносупресивна ТМЕ, е образуването на свързани с рака фибробласти, водещи до повишено производство на извънклетъчен матрикс в ТМС (580). Dalal и

сътр. установяват че, повишената концентрация на TGF- β в TME се поддържа от туморните клетки, фибробластите, инфилтрираните имунни клетки и съдовите клетки, чрез автокринни или паракринни сигнали (581). По този начин TGF- β сигналният път се очертава като главен регулатор на TME, който организира взаимодействието между туморните клетки и фибробластите, води до промени в извънклетъчния матрикс, които потиска имунните клетки и повлиява отговорите към имунотерапията.

5.3.9. Немутационни промени водещи до трансформация на клетките.

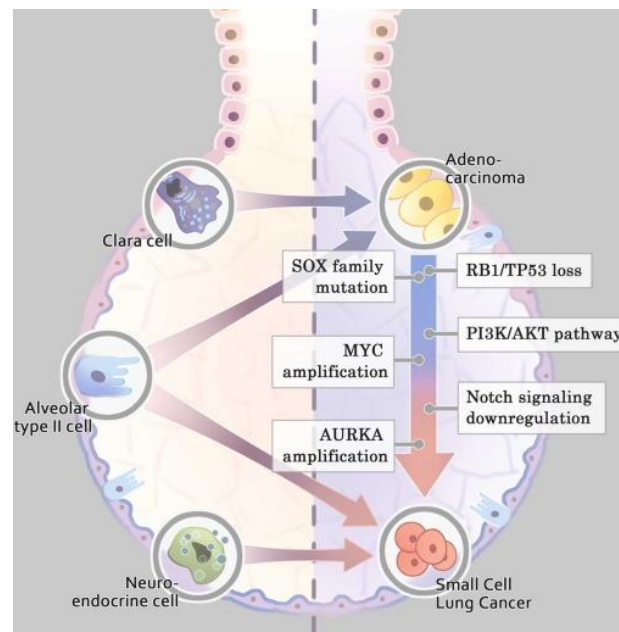
В последните години се наблюдава усилено изучаване на мутационните и немутационните причини за възникване на резистентност. Немутационните процеси до голяма степен се основават на пластичността на туморните клетки, която ги прави нечувствителни към приложения химио- или имунотерапевтик и по този начин улеснява оцеляването и растежа им. Концепцията за пластичност на туморните клетки подчертава значението на повторното активиране на програмите за развитие, тясно свързани с епително-мезенхимния преход, придобиване на свойства на ракови стволови клетки и потенциал за трансдиференциация (223). По отношение на промяната на фенотипа, клетъчната пластичност (фундаменталната способност на клетките да променят свойствата си по обратим начин, активно или пасивно) има важна роля при възстановяването и регенерацията на тъканите след травма, както и при възстановяването на нарушената хомеостаза (582).

Трансформацията от не-дребноклетъчен в дребноклетъчен белодробен карцином е докладвана за първи път през 2006 г. при 45-годишна жена непущач с установена делеция от 18 bp в екзон 19 на EGFR гена и лекувана с ерлотиниб (583). При повторна биопсия е установен хистологична трансформация в с дребноклетъчен белодробен карцином. Оттогава насам няколко проучвания са идентифицирали подобна пластичност на туморните клетки като тя може да възникне след таргетна терапия, имунотерапия или химиотерапия и е една от причините за придобитата резистентност (584). В повечето случаи е наблюдавана трансформация от АК към дребноклетъчен карцином, но само няколко проучвания са описали трансформация на ПК (585). Трансформацията към ДРБД може да се нарече и невроендокринна трансформация. Този феномен е наблюдаван и при рак на простатата, където невроендокринната диференциация е била индуцирана след химиотерапия или ендокринотерапия (586). Една от хипотезите за възникването ѝ е общия клетъчен произход.

Тиреоидният транскрипционен фактор-1 (TTF-1) е невроендокринен маркер при ДРБД и се експресира в 69%–80% АК (587). TTF-1 е 43-kDa хомеодомен-съдържащ протеин, експресиран в предния мозък, щитовидната жлеза и белия дроб. В белия дроб TTF-1 играе решаваща роля в регулацията на белодробната морфогенеза, диференциацията на епителните клетки и експресията на гени, от които зависи перинаталната дихателна адаптация (588). TTF-1 се открива в респираторните епителни клетки по време на цялата белодробна морфогенеза и в зряла възраст (589). В проучването на Satoshi Nozari и сътр., учените установяват различни места на свързване на TTF-1 с ДНК (587). При ДРБД TTF-1 подпомага растежа на раковите клетки и допринася за невронната диференциация чрез частична координация с ASCL1

или с транскрипционните фактори от Forkhead семейството, активирайки експресията на тъканно-специфични гени по време на развитието и клетъчното препрограмиране (587). В миши модел на АК TTF-1 се свързва и взаимодейства с FOXA, осигурявайки пряка връзка между транскрипционните програми за диференциация на белия дроб и иницирането на тумора (587).

Освен това на границата между компонентите на ДРБД и АК са открити клетки с преходна морфология, което показва, че двата карцинома вероятно произхождат от общи стволови клетки, експресиращи TTF-1 (590). Въпреки че невроендокринните клетки са основната група клетки, от които произлиза ДРБД, проучвания върху животни показват, че ДРБД може да възникне и от прогениторните клетки или алвеоларни клетки тип II. Аденокарциномът обикновено се развива от алвеоларни клетки тип II в периферната област, което предполага, че ДРБД и АК могат да имат един и същ клетъчен произход (591) като различни генетични промени иницират процеса - загубата на тумор-супресорните гени *RB1* и *TP53*, заедно с мутацията на в *SOX* гените, амплификацията на *MYC* и *AURKA*, както и нарушения в *PI3K/AKT* и Notch сигналните пътища (585).



Фигура 5.5. Схематична диаграма описваща механизмите на трансформация на АК в ДРБД (585).

Като заключение онтологичният анализ на 15808 гена-мишени на микроРНКите от АК Клъстер 2 показва, че ГО термините са силно свързани със транскрипцията, транслацията, организацията на клетъчната мембрана, апарата на Голджи, включително регулация на клетъчната диференциация и морфогенеза. След засегнатите иРНКи преобладават кадхерини, протеин кинази, РНК-свързващите се белтъци, убиквитин лигази и транскрипционни фактори. Онтологичният анализ на 11560 иРНКи-мишени на микроРНКите от АК Клъстер 1 показва, че получените ГО термини са силно свързани със специфични клетъчни процеси, улесняващи метастазирането в мозъка, клатрин-медирана ендоцитоза, липиден метаболизъм и др. (Фигура 4.42, Раздел 4).

Можем да заключим че при АК обогатените GO термини са свързани с метастазирането и прогресията, фокалната адхезия, ЕМТ и транзицията към дребноклетъчен белодробен карцином и са засегнати са повече цитокин – медириани сигнални пътища. Тези фенотипни промени показват че тези микроРНКи могат да са потенциални регулатори на подвижността, миграцията и инвазия на туморните клетки. Резултатите ни са в общо съгласие с предишни световни доклади, показващи че честотата на различните видове тумори се е променила през последните години, така че в днешно време аденокарциномът на белия дроб се среща по-често от плоскоклетъчния карцином и че е най-честият хистологичен тип на метастазиращия първичен рак на белия дроб.

Можем да заключим че при ПК обогатените GO термини са свързани с клетъчния цикъл, адхезията и организацията на цитоскелета. Получените данни показват, че тези микроРНКи могат да са потенциални биомаркери при ПК, тъй като регулират основни процеси свързани с инициацията и поддържането на туморния фенотип.

GO анализът на 6508 гена на четирите микроРНКи при ПК показва, че GO термините са силно свързани с регулация на генната експресия, клетъчния отговор към инсулин, вътреклетъчния транспорт, апоптоза, автофагия и убиквитилацията. След засегнатите иРНКи преобладават РНК-свързващите се белтъци, кадхерините, протеин кинази и транскрипционни фактори.

5.4. ОБСЪЖДАНЕ НА РЕЗУЛТАТИТЕ ОТ RT-qPCR за мРНКите и асоциация с клинично-патологичните характеристики.

Особено обещаващо в лечението на НДБРД е използването на моноклонални антитела срещу имунни контролни точки. Въпреки това много от пациенти получават рецидиви и развиват резистентност. Механизмите, лежащи в основата на резистентността към имунотерапията, не са напълно разбрани. Смята се, че различни транскрипционни, посттранскрипционни и транслационни механизми допринасят за свръхекспресията на PD-L1 при ракови заболявания и тези механизми се регулират от различни онкогенни сигнални пътища.

В настоящия дисертационен труд при проведения онтологичен анализ и анализ на сигналните пътища установихме няколко засегнати цитокин-медириани пътя. Измежду гените, обогатяващи се предимно в 20-те най-силно повлияни от микроРНКите пътя при АК и ПК, избрахме 4 (*EGFR*, *STAT3*, *PTEN* и *mTOR*) значително по-представени, чиято роля да изследваме при АК и ПК и асоциацията им с основните клинично – патологични характеристики и корелацията им с експресията на PD-L1.

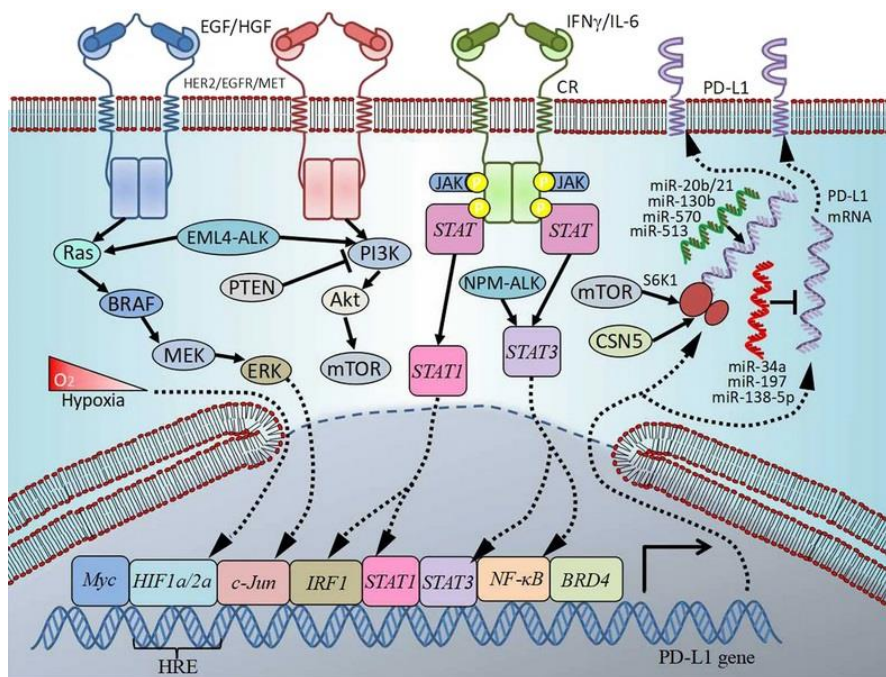
Един от най-значимите компоненти, поддържащи геномна стабилност е епигенетичната модификация на ДНК молекулата и хистоновите белтъци. Аберациите в метилиране на хистони H3 и H4 са свързани с иницирането и поддържането на карциногенезата при РБД (592). Такива аберации могат да бъдат предизвикани от гени на околната среда, активирани поради излагане на опасности вещества или канцерогени (107). Настоящото проучване изследва експресионните нива на един такъв ген - *MINA53*, в двата подтипа на НДБРД.

Корелацията на изследваните мРНК молекули и 14 микроРНКи е значима, но слаба до умерена и за валидацията ѝ е необходимо проследяването им в по-голяма кохорта пациенти. При АК нивата на *MINA53* и *mTOR* са повишени, а експресията на *PD-L1*, *STAT3*, *PTEN*, *EGFR* е намалена (Фигура 4.55А, Раздел 4). При ПК открихме, че *MINA53*, *PD-L1*, *PTEN*, *EGFR* и *mTOR* са с повишена експресия и само *STAT3* е намален в сравнение със съседните нетуморни тъкани (Фигура 4.55Б, Раздел 4).

5.4.1. Обсъждане на резултати получени при анализа на експресията на PD-L1 при АК и ПК.

Нивата на експресия на *PD-L1* бяха значително по-високи в ПК, отколкото при АК (Фигура 4.55В, Раздел 4) и се асоциираха с ниско диференцираните тумори, без кератинизация и без далечни метастази (Фигури 4.70, 4.69 и 4.74 , Раздел 4). При АК по-високата експресия се наблюдава при пациентите с нодални (Фигура 4.60, Раздел 4) и далечни метастази (Фигура 4.61, Раздел 4), ниско диференцираните тумори (Фигура 4.64, Раздел 4), при които преобладава солидна или микропапиларна морфологична компонента Фигура. Скорошни метаанализи показват неблагоприятната прогноза при *PD-L1* положителните тумори (593).

Предполага се, че много сигнални пътища вземат участие в контрола на експресия на *PD-L1* и той е различен, при двата подтипа белодробен карцином (594). За да се разработи по-ефективна терапевтична стратегия за имунотерапия, е необходимо да се изучат специфичните сложни регулаторни механизми на *PD-L1* при АК и ПК. Налице са някои проучвания, които показват, че отговорите и токсичността на имунотерапията, могат да бъдат повлияни от герминативни и/или соматични мутации, което би било от решаващо значение за определяне на прогнозата на пациентите (595). Най-често срещаните начини за епигенетична регулация на експресията на *PD-L1* са промоторното хиперметиране на ДНК, модифицирането на хистони, регулирането чрез некодиращи РНК молекули и N6-метиладенозин. Също така промяна в броя копия, амплификации, мутации в 3'-UTR на *PD-L1* гена (596). Аберантно активиране на *MET* при РБД индуцира експресията на негативните регулатори на имунния отговор (*CD274*, *SOCS1* и *PDCD1LG2*) и на факторите, отговорни за ангиогенезата и васкуларизацията (*VEGFA* и *NRP1*) (597). При тази група пациенти, въпреки високата експресия на *PD-L1*, туморите показват намален отговор към имунотерапията (598). Освен това експресията и/или активността на *PD-L1* се регулира чрез няколко посттранслационни модификации, включително фосфорилиране, гликозилиране, ацетилиране, убиквитиниране и палмитоилиране, които служат като сигнали за лизозомно или протеазомно медирано разграждане на *PD-L1* (599).



Фигура 5.6. Транскрипционен и посттранскрипционен контрол на експресията на PD-L1 при онкологичните заболявания. Онкогенната активация на RAS/RAF/MEK/MAPK-ERK и PI3K/PTEN/Akt/mTOR сигналните пътища може да доведе или до директно активиране на таргетните гени или активирането на различни транскрипционни фактори като STAT3, STAT1, c-Jun, HIFs, и NF-κB. Те преминават в ядрото и се свързват с промотора на PD-L1 гена и индуцират експресията. Също така експресията на PD-L1 може да бъде регулирана от различни микроРНКи. (596)

5.4.2. Обсъждане на резултати получени при анализа на експресията на STAT3 при АК и ПК.

STAT3 е член на голямото белтъчно семейството STAT (600). В отговор на цитокини и растежни фактори, STAT3 се фосфорилира от рецептор-асоциирани Janus кинази (JAK). Образува хомо- или хетеродимери и навлиза в клетъчното ядро. Активирането на STAT3 може да се получи и чрез фосфорилиране MAPK киназата и чрез c-src нерецепторна тирозин киназа (601). STAT3 действа като транскрипционен фактор на редица регулатори на клетъчния цикъл и участва в комуникацията тумор-имунна клетка, чрез намаляване на ефективния антитуморен имунитет, посредством различни механизми, включително регулиране на експресията на PD-L1 гена (601–603). Освен това медира освобождаването на различни цитокини и хемокини, като по този начин повлиява туморната микросреда и тумор-инфилтриращите лимфоцити (604). Смята се, че повишената експресия на STAT3 и неговите таргетни молекули допринасят за развитието на туморите, резистентност към радио-, химио- и EGFR таргетна терапия (602). В допълнение контролира функциите на митохондриите, клетъчния метаболизъм и стволово клетъчния потенциал (605).

STAT3 самостоятелно или заедно с NF-κB регулира експресията на много цитокини и хемокини (606). IFNγ, кодиран от *IFNG* гена, е първият цитокин, които се секретират от макрофагите и другите тумор инфилтрирани лимфоцити. Той регулира транскрипцията на PD-L1 чрез активиране на JAK/STAT3 сигналния път (607). Така чрез тази положителна връзка, секретирани от раковата клетка цитокини като TGF-β, IL-6, IL-10 и VEGF възпрепятстват имунния отговор, подпомагат имунното избягване

на тумора и стимулират туморната пролиферация и инвазия. От друга страна активираният STAT3 в туморната клетка е и директен транскрипционен фактор на PD-L1. Така STAT3 ограничава антитуморния ефект на тумор инфилтрираните лимфоцити и NK клетките (608).

В изследваните от нас тумори STAT3 има значително по-висока експресия при ниско диференцираните ПК тумори (Фигура 4.70, Раздел 4). При АК се наблюдава силна положителна корелация между експресионните нива на STAT3 и PD-L1 (Фигура 4.56, Раздел 4).

5.4.3. Обсъждане на резултати получени при анализа на експресията на PTEN при АК и ПК.

Генът PTEN (phosphatase and tensin homolog) е локализиран на дългото рамо на 10 хоромозома (10q23.31) и кодира, липид/белтък фосфатаза. Този белтък има роля в множество клетъчни процеси, включително растеж, пролиферация, оцеляване и поддържане на геномната стабилност. PTEN действа като туморен-супресор чрез отрицателно регулиране на PI3K/AKT сигналния път (609).

При изследваните ПК експресията му е значително по-ниска при кератинообразуващите карциноми (Фигура 4.74, Раздел 4) и при пациентите, при които се откриват далечни метастази (Фигура 4.69, Раздел 4). Соматичните мутации на PTEN се срещат при множество злокачествени заболявания, включително глиома, меланом, рак на простатата, ендометриума, рак на гърдата и яйчниците, бъбреците и белия дроб. При 3,33% от пациентите с НДБК се наблюдават мутации в PTEN гена (28). Тези пациенти се характеризират с висок мутационен товар и хромозомни пренареждания (610).

Подобно на RAS/MEK/ERK сигналния път PI3K/AKT/mTOR има важна роля в туморогенезата, контрола на експресията на PD-L1 и активацията на фактори, благоприятстващи клетъчното оцеляване и пролиферация (611). Загубата на PTEN е свързана и с промени в цитокинова и хемокинова секреция с имунопотискащ ефект на туморната микросреда. Туморите с дефицит на PTEN са свързани с по-висока експресия на индолеамин 2,3-диоксигеназа 1 (IDO1) и PD-L1 като по този начин намаляват активността на цитотоксичните имунни клетки (612).

5.4.4. Обсъждане на резултати получени при анализа на експресията на mTOR при АК и ПК.

Генът *mTOR* е локализиран на късото рамо на първа хромозома (1p36.2) и кодира белтък, принадлежащ към фамилията на фосфатидилинозитол киназа-свързаните серин-треонинови кинази. *mTOR* играе важна роля в контролирането на клетъчния растеж, метаболизъм и пролиферация, автофагия, оцеляване при стрес, имунитет и регулира транслацията на различни белтъци в клетката, вземащи участие в тези метаболитни процеси (613). mTOR белтъкът е мишена на комплекса рапамицин-FKBP12 като блокирането му води до прекъсване на клетъчния цикъл, имуносупресия и индуцира повишаване на експресия на PD-L1 (613). При 3.61% от пациентите с НДБК се наблюдават мутации в *mTOR* гена (28).

Установено е че, по-високите нива на експресия на *mTOR* са предиктивен фактор за чувствителност към лечение с *mTOR* инхибитори при панкреасни НЕТ (614). В нашето изследване експресията е по-висока сред АК тумори на белия дроб (Фигура 4.55В, Раздел 4), което предполага, че лечението с *mTOR* инхибитори може да се прилага при тези тумори. Освен това експресията му е значимо по-висока при карциномите, при които са налице нодални (Фигура 4.59, Раздел 4) и далечни метастази (Фигура 4.60, Раздел 4). *mTOR* и *PTEN* са белтъци от една и съща сигнално-трансдукционна каскада. Liang et al съобщават, че хиперактивираният сигнален път на *mTOR* води до химиотерапевтична резистентност в *KRAS*-мутантни ракови клетки на белия дроб (615). Инхибирането на *mTOR* възстановява чувствителността на раковите клетки към цисплатина в 2D и 3D клетъчни култури и миши модели (615). Това определя инхибирането на *mTOR* като рационален терапевтичен подход в комбинация с цисплатина. Съобщава се, че фосфорилирането на серинов остатък на 727 място в молекулата на *STAT3* от *mTOR*, води до неговото максимално активиране (616).

В допълнение гените като *PTEN* и *mTOR*, които са тясно свързани с развитието на тумора, могат косвено да регулират транскрипцията на *PD-L1* и по този начин да насърчат избягването на туморните клетки от имунната система (Фигура 4.56, Раздел 4).

5.4.5. Обсъждане на резултати получени при анализа на експресията на EGFR при АК и ПК.

EGFR е трансмембранен рецепторен белтък с тирозин киназна активност и се характеризира с повишена експресия в голям брой белодробни карциноми (617). Свързва се с туморния растеж, неоангиогенеза, туморна инвазия и клетъчно безсмъртие (618). В изследваните от нас тумори е значимо по-ниска експресията му при АК сравнение с ПК (Фигура 4.55В, Раздел 4). Нашите резултати потвърдиха предишни открития на Kosaka и сътр., Vrabender и сътр., че по-висока експресия на *EGFR* се наблюдава при плоскоклетъчен карцином, отколкото при аденокарцином (619,620). При 62% от НДБК *EGFR* показва повишена експресия, която се асоциира с по-лоша прогноза (621). Резултатите ни показват, че експресията на *EGFR* е значително по-висока при нискодиференцираните тумори (Фигура 4.61, Раздел 4) и АК със солидна и микропапиларна компонента (Фигура 4.67, Раздел 4).

При ПК също така се намериха позитивни корелации с нивата на експресия на *mTOR* и *MINA53*, които биха могли да се обяснят с връзката им с *EGFR* пътя (Фигура 4.57, Раздел 4). Известно е, че *MINA53* регулира експресията на *EGFR* при НДКБД (111,622). Активираният *EGFR* включва сигнално-трансдукционната каскада, част от която е и *mTOR*, а в нашето проучване установихме и на експресивно ниво връзка между *EGFR* и *mTOR* при ПК(623). *EGFR* участва в регулирането на експресията на *PD-L1* и клетъчната пролиферация чрез сигналния път *IL-6/JAK/STAT3* при НДРБД (Фигура 4.56 и 4.57, Раздел 4) (624). **Настоящото проучване показва потенциала на комбинираната таргетна терапия с имунотерапия при лечението на АК.**

5.4.6. Обсъждане на резултати получени при анализа на експресията на MINA53 при АК и ПК.

Някои фактори на околната среда като тютюнев дим, силициев диоксид и арсен, индуцират експресията на *MINA53*, чрез JNK-STAT3 сигналния път (625). Наблюдавахме положителна корелация между нивата на експресия на *STAT3* и *PD-L1* с *MINA53* (Фигура 4.57, Раздел 4). Въпреки това е необходим по-нататъшен функционален анализ и валидиране на резултатите, в по-големи кохорти, за да се установи по-добре биологичната функция и прогностичния потенциал на *MINA53* и връзката му със *STAT3* и *PD-L1* в АК и ПК.

Корелационният анализ при двата подтипа НДРБД показва корелация между експресия на *EGFR* и *MINA53* (Фигури 4.59 и 4.60, Раздел 4). Корелацията между двата биомаркера е наблюдавана и в изследването на Ма и сътр. (626). В клетъчните линии свръхекспресията на *MINA53* повишава нивата на *EGFR* и стимулира туморната пролиферация чрез регулиране на фазовия преход G1 към S (626).

В изследваните от нас пациенти няма статистически значима разлика в експресията на *MINA53* между АК и ПК в ранен и напреднал стадий (Фигури 4.61 и 4.69, Раздел 4). Въпреки това, данните показват повишена експресия на *MINA53* в ранен стадий АК и ПК. Тези констатации могат да се обяснят с хетерогенността на тумора и различни морфологични подтипове на пробите АК и ПК. Komia и сътр. показват, че повишената експресия на *MINA53* при пациенти с рак на белия дроб в стадий I-II е свързана с благоприятна прогноза (111). Обратно, намалените нива на *MINA53* в напреднал стадий на НДРБД индуцира инвазивен фенотип (111). Тези резултати показват двойната функция на *MINA53* и неговата роля като ранно събитие в канцерогенезата на рак на белия дроб, свързана с цитокиновата мрежа EGFR, IL-6 и HGF.

Открихме значително намалена експресия на *MINA53* при пациенти с АК и ПК с лимфни възли и далечни метастази (Фигури 4.63, 4.64, 4.70 и 4.71, Раздел 4). Нашите резултати са в съгласие с предишните проучвания, илюстриращи ролята на *MINA53* като ранно събитие по време на белодробната карциногенеза (110,111,627). Geng и сътр., демонстрират, че в клетъчни линии на рак на белия дроб, експресията на *MINA53* може инхибира инвазията и метастазите, чрез регулиране на белтъците на EMT (627).

От друга страна според Zhang и сътр. *MINA53* също така деметилира H3K9me3 и отключва експресията на генни, участващи в клетъчната адхезия, G-белтъчната сигнализация и експресията на белтъците на цитоскелета, които поддържат клетъчната инвазивност и метастази (96). По този начин двата колектива показват различната роля на *MINA53*, според етапа на карциногенезата.

5.5. Обсъждане на резултатите от проведеното машинно самообучение .

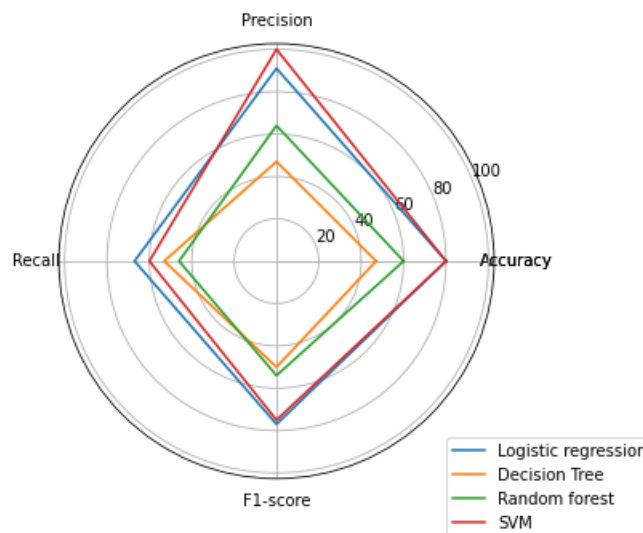
Откриването на асоциация на микроРНКите с АК или ПК може не само да помогне да се разбере ролята им в онкогенезата, но също така може да послужи като диагностичен маркер. Специфичните профили на експресия могат да допринесат за по-

точното субкласифициране на двата подтипа не-дребноклетъчни белодробни карциноми.

Като се имат предвид вариациите на микроРНКите в различните изследвания, касаещи НДРБД, откриването на универсална микроРНКа е все още проблематично. Хетерогенността на РБД усложнява класификацията, особено като се имат предвид и различните генетични и епигенетични промени, както и преданалитичните и аналитичните методи на детекция. Машинно самообучение, поддисциплина на изкуствения интелект, е полезен инструмент за анализ на данни от генна експресия и идентифицирането на модели, които не могат да бъдат точно уловени чрез традиционните математически подходи (628).

Моделите на машинно самообучение обикновено се разделят на три категории: контролирано, неконтролирано и обучение с подсилване. Методите за контролирано обучение използват информация за етикетите от извадките за обучение на модела, докато при неконтролираното обучение моделите се извличат директно от данни без етикети. В настоящото изследване анализирахме експресията на 14-те подбрани от микрочиповия анализ микроРНКи, при аденокарциноми и плоскоклетъчни белодробни тумори. За различните видове данни и проблеми за решаване са подходящи определени алгоритми за бинарна класификация. За да разберем, кой алгоритъм най-добре разграничава АК от ПК на база ко-експресията на 14-те микроРНКи, тренирахме и тествахме 4 алгоритъма за класификация. Едно от ограниченията на настоящето изследване е малкият размер на стратифицираната извадка за трениране и тестване като чрез прилагане на Grid Search и кръстосано валидиране частично се надяваме да намалим този ефект. При този метод всеки от моделите се обучава с всичките възможни комбинации от хиперпараметри и техните стойности докато не открие най-добрите на базата на конкретен оценител (F1-score).

Получените резултати при класификация АК и ПК са обобщени и представени графично на Фигура 5.7. На база ко-експресията на 14-те микроРНКи DT и RF не са подходящи и не могат да разграничат АК и ПК. От графиката се вижда, че логистичната регресия и методът на опорните вектори се представят по-добре от останалите два алгоритъма. Двата модела се характеризират с по-висока прецизност за сметка на по-ниска чувствителност. Като цяло и двата алгоритъма класифицират АК с висока положителна прогностична стойност, но класифицират погрешно ПК като АК. Това показва, че ко-експресията на 14-те микроРНКи не е подходяща да се използва като диагностичен панел за разграничаване на двата подтипа НДРБД. Необходимо е увеличаване на извадката и допълнителни маркери, които да се включат в анализа.



Фигура 5.7. Сравнение на метриците за оценка на всеки от приложените модели за машинно самообучение .

Чрез методите за неконтролирано машинно самообучение, включващи йерархичния клъстерен анализ (Фигури 4.84 и 4.85, Раздел 4) и метода на k-най-близките съседи (Фигури 4.86 и 4.87, Раздел 4), открихме специфичен морфологичен модел на ко-експресия на 14-те микроРНКи.

Клъстерирането, базирано на йерархичното групиране, може да анализира връзката между Евклидовото разстояние на всяка проба, така че проби с относително близки разстояния да бъдат групирани заедно. Предимството на йерархичния метод, както е отбелязано и от Eisen, е, че естественият начин за възприемане и анализиране на голямо количество от данни е първо да се открият големите характеристики и разлики, а след това да се насочи вниманието върху по-малките детайли (629).

Клъстерирането, базирано на центроидите, организира данните в не-йерархични клъстери, за разлика от йерархичното клъстериране. Методът на k-най-близките съседи, базирани на центроид, са ефективни, но чувствителни към началните условия и аутлаерите. **Като цяло наблюдаваме запазване на разделението на туморите според морфологията, което показва че изследваните 14 микроРНКи може да не са подходящи маркери за отдиференциране на АК от ПК, но са специфични за морфологичните им подтипове.**

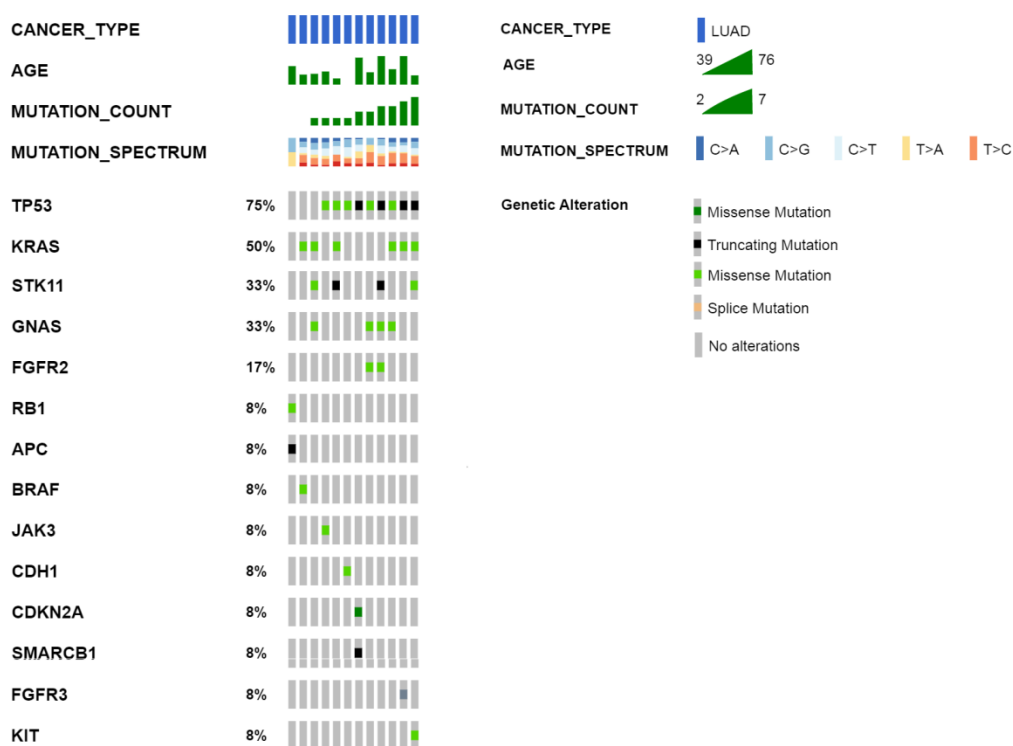
5.6. Обсъждане на резултатите от проведеното секвениране от следващо поколение.

Не-дребноклетъчният рак на белия дроб вече може да бъде класифициран в голям брой подтипове в зависимост както от хистологичния външен вид (плоскоклетъчен срещу аденокарцином), така и от наличието/отсъствието на определени водещи мутации (232).

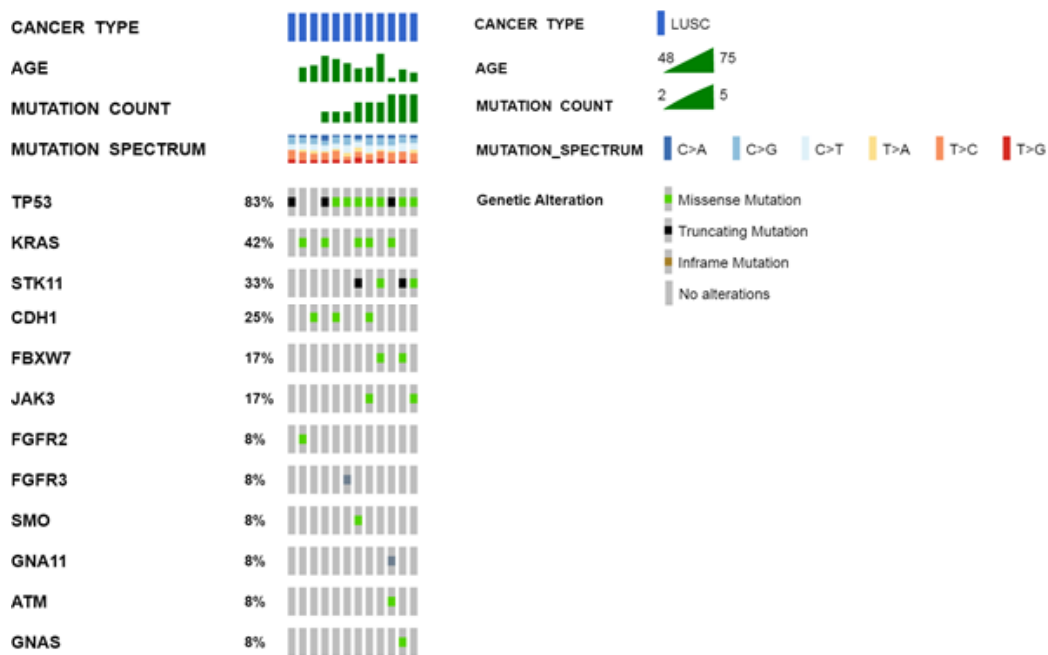
Част от поставените задачи в настоящия дисертационен труд е да се изследва соматичния спектър на мутации при РБД, събрани от 24 пациента с АК и ПК, посредством таргетно NGS секвениране, (TruSeq Amplicon-Cancer Panel) на платформа

MiSeq, Illumina. Бяха избрани АК и ПК от всеки морфологичен подтип. Непосредствено до представените патогенни варианти се очертаха и три варианта в *TP53*, *JAK3* и *EGFR* гените, които не са определени като патогенни в базите данни, но се свързват с канцерогенезата и лекарствената терапия (Таблицы 4.15 и 4.16, Раздел 4).

Най-често мутиралите гени в изследваната извадка, както и техния потенциален ефект върху канцерогенезата на туморите са подробно дискутирани. Останалите изследвани гени, при които са установени мутации, се откриват само при няколко пациента и допълват мутационния спектър. На Фигури 5.8 и 5.9 са представени генетични варианти открити при АК и ПК.



Фигура 5.8. Генетични варианти открити при АК (630).



Фигура 5.9. Генетични варианти открити при ПК (630).

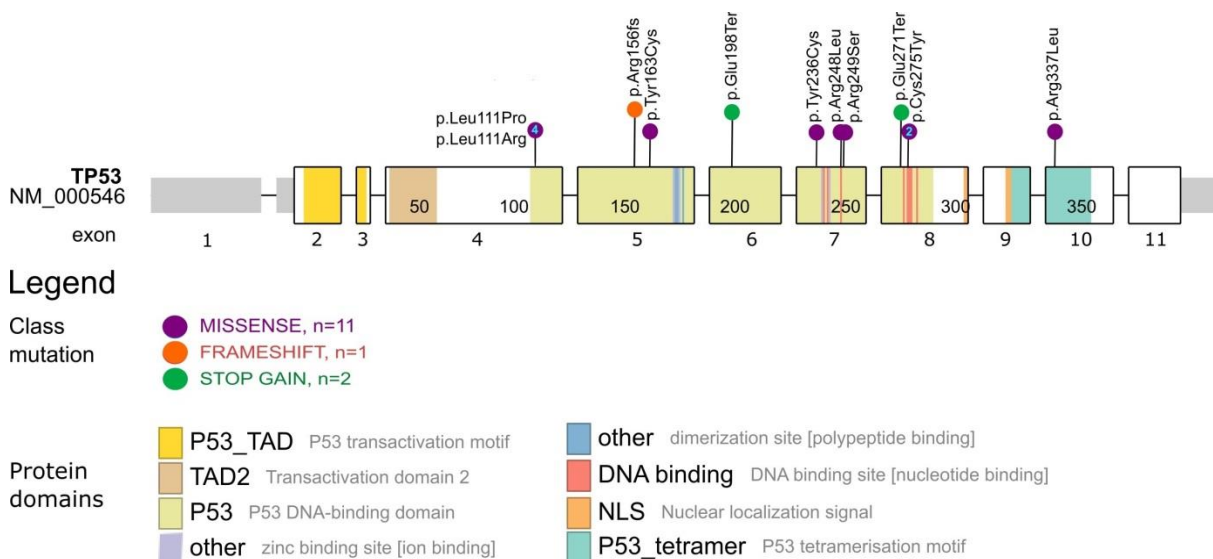
5.6.1. Мутации в *TP53* гена.

TP53 (17p13.1) е тумор супресорен ген, който има роля в поддържането на геномната стабилност, клетъчния цикъл, ДНК репарацията, апоптозата, ангиогенезата и стареенето. P53 белтъкът е хомо-тетрамер, съдържа 393 аминокиселини, четири белтъчни домена и характеризира се с високо ниво на консервативност между гръбначните животни (631). N-терминалният регион (АК остатъци 1-62) е разделен на два субдомейна (TAD1 и TAD2) и съдържа трансактивиращ домейн; пролин-богат регион (АК остатъци 63-94), със значение относно апоптотична активност; ДНК-свързващ домейн (DBD; АК остатъци 94-292), който е познат още като коров домейн, съдържащ електроположителни аргининови аминокиселини и един цинков атом, който взаимодейства с тандемни повтори на р53 отговарящия елемент (RRRCWWGYYY: R, G/A; W, A/T; Y, C/T) в ДНК веригата, разделени с до 13bp в промоторните региони на своите целеви гени. ДНК-свързващия домейн се свързва към нуклеарните сигнални домейни (АК остатъци 316-325), които са отговорни за тетрамеризацията, която е основна за р53 активността. С-терминалният домейн (АК остатъци 363-393) е включен в даун-регулацията на централния ДНК-свързващ домейн (632).

При изследваните пациенти с АК и ПК преобладават мутациите в *TP53* гена, което е в съгласие с твърдението, че *TP53* гена е един от най-често мутиралите гени при солидни тумори (633). Мутациите в *TP53* гена се появяват рано в процеса на канцерогенезата и са свързани с лоша обща преживяемост, резистентност към терапия и чести случаи на рецидиви. При нормални условия концентрацията на р53 в клетката се поддържа ниска. При хипоксичен стрес или ДНК повреда, ядрената концентрация на р53 се увеличава. В отговор на увреждането на ДНК, р53 се регулира по MDM2-зависим начин и се активира от CHK1 и CHK2, което води до спиране на клетъчния

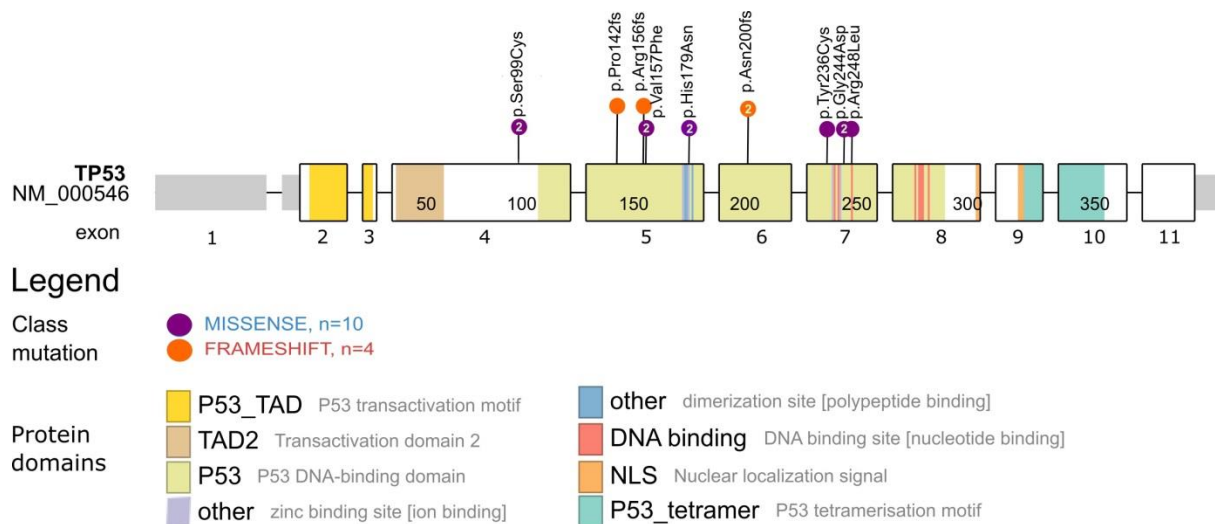
цикъл и апоптоза. При настъпване на мутация в гена и загуба на функцията, белтъка p53 не може да се свърже с промоторите на таргетните си гени и по този начин малигнените клетки не се насочват към апоптоза и се стимулира тяхната пролиферация (634). Интересно е да се отбележи, че в някои пациенти се откриват повече от една мутация в гена. Това може да се дължи на факта, че ДНК е изолирана от свеж туморен материал и отразява вътретуморната хетерегенност и присъствието на няколко клона туморни клетки.

Спектърът на мутациите в TP53 гена при ПК включва 11 missense, 2 безсмислени (nonsense) мутация и 1 вариант, при който се наблюдава изместване рамката на четене. Миссенс мутациите, които променят една аминокиселина в p53 белтъка влияят по различен начин на пространствена конформация. Преобладават мутациите засягащи ДНК свързващия домен (екзони 4-8) и една мутация се открива в 10 екзон на гена. Също така често се среща еднуклеотидната замяна C>A/G>T. Всички пациенти отбелязват в анамнезата си тютюнопушене в различна степен (тежки, леки или бивши пушачи), което обяснява преобладаващия мутационния профил, характерен за тютюнопушенето. Фигура 5.10. изобразява кодиращите екзони на TP53 гена, белтъчните домени и откритите мутации.



Фигура 5.10. Кодиращите екзони на TP53 гена, белтъчните домени и откритите мутации при ПК (635).

Спектърът на мутациите в TP53 гена при АК включва 10 missense и 4 варианта, при които се наблюдава изместване рамката на четене. Миссенс мутациите са разнообразни и не се открива преобладаване на определен мутационен профил. Повечето от пациентите с АК са непущачи. В резултат на делециите се измества рамката на четене и се транслира скъсен белтък, който не може да изпълнява основната си биологична роля. Фигура 5.11 изобразява кодиращите екзони, белтъчните домени и откритите мутации на TP53 гена.



Фигура 5.11. Кодиращите екзони на *TP53* гена, белтъчните домени и откритите мутации при АК (635).

Въз основа на вида на замените, мутациите в *TP53* показват забележително предпочитание към миссенс мутациите пред безсмислени мутации или мутации, променящи рамката на четене, които обикновено са доминиращи в други туморни супресорни гени като *RBI*, *APC* и *PTEN* (636). През последните години има нарастващ брой проучвания, обсъждащи значението на мутациите в *TP53* гена за прогнозата при пациенти с НДРБД. Предполага се, че мутациите въздействат по различен начин върху клиничния изход на пациентите.

Проучване на Labbé и сътр., показва, че пациенти с миссенс мутации в *TP53* гена, в сравнение с тези, при които се откриват малки делеции/инсерции, имат значително намалена преживяемост без прогресия при лечение с първа линия терапия с EGFR TKI (471). Скорошен метаанализ предполага, че наличието на мутации в *TP53* гена води до намалена клинична ефикасност на EGFR-TKIs и ALK-TKIs и е отрицателен прогностичен фактор за пациенти с напреднал НДРБД, при които се детектират мутации в EGFR гена или пренареждания на ALK (независимо от патологичния подтип и линията на лечение (637). Обратно, наличието на *TP53* мутация не повлиява преживяемост без прогресия или общата преживяемост на пациентите, лекувани с 2-ро поколение EGFR-TKI (638).

Според проучване на Poeta и сътр., мутациите в *TP53* могат да бъдат класифицирани като деструктивни и недеструктивни мутации въз основа на степента на нарушение на белтъчната структура (639). Деструктивните мутации водят до пълна или почти пълна загуба на функция на p53 белтъка и засягат ДНК свързващия домен. Докато недеструктивните мутации се отнасят до консервативни или неконсервативни замини (с изключение на стоп кодони) извън ДНК свързващия домен, които могат да запазят някои от функционалните свойства на p53 белтъка (639). Няколко поучавания показват, че недеструктивните мутации в *TP53* гена имат значима връзка със съкратена обща преживяемост, в сравнение с дивия тип (640–642).

Много изследвания се фокусират върху местоположението на мутациите. Най-често мутациите се откриват в екзон 4-8 на *TP53* гена, което представлява около 44,8% от случаите (643). Проучването на Jiao и сътр. показва, че пациенти с мутации в екзон 4, екзон 6, и множество мутации в *TP53* гена имат по-лоша прогноза, отколкото

пациенти, при които се открива комутация на EGFR екзон 19/21 и мутация в екзон 5, екзон 7, екзон 8 и екзон 9 на TP53 гена (643). Освен това е доказано, че T790M позитивните пациенти с мутация на TP53 R237C не се повлияват при лечение с osimertinib, а полиморфизмът на TP53 rs55863639 се свързва с по-лоша прогноза при пациенти с TP53/EGFR ко-мутация (644,645). Досега не са публикувани проучвания при лечение на пациенти с EGFR-TKIs и химиотерапия при напреднал NSCLC с ко-мутация на EGFR/TP53. В ClinicalTrials.gov в момента се провежда едно регистрирано клинично проучване във фаза III, сравняващо монотерапия с osimertinib с комбинирана терапия с osimertinib, карбоплатина и пеметрексед за нелекувани пациенти, диагностицирани с напреднал АК с мутации в *EGFR* и *TP53* гените.

Ретроспективно проучване на Liu Y изследва експресията на PD-L1 и асоциацията ѝ с единадесет генни мутации при 247 резецирани първични и 26 напреднали пациенти с НДРБД (646). Колективът открива, използвайки мултивариантна логистична регресия, че експресията на PD-L1 се асоциира с мутация в TP53 гена (646). Освен това, Hao Sun и сътр. демонстрират, че пациенти с TP53 миссенс мутации имат повишено ниво на PD-L1 и обогатени IFN- γ сигнатури в сравнение с подгрупата от TP53-див тип (647). Проучването на Jin Y и сътр., включващо голяма кохорта пациенти в Китай, предполага, че мутациите в TP53 са значително асоциирани с висока експресия на PD-L1 (648). Още повече високата експресия на PD-L1 е значително свързана с ко-мутация на EGFR и TP53, докато мутация само в EGFR гена сама по себе си не е свързана с висока експресия на PD-L1 (648). Тези резултати биха могли да обяснят защо пациентите в напреднал НДРБД, при които има само мутация на EGFR показват лош отговор към имунотерапията и пациентите с ко-мутация на EGFR/TP53 се повлияват по-добре при лечение с анти-PD-1/PD-L1.

5.6.2. Мутации в *KRAS* гена.

Белъкът KRAS, кодиран от протоонкогена *KRAS*, е малка гуанин трифосфатаза (GTPase), която служи като бинарен превключвател в сигналната трансдукция за повечето рецепторни тирозинкинази, включително EGFR, MET или ALK, и играе ключова роля в регулирането на различни клетъчни функции (649,650). В изследваните от нас пациенти мутации в *KRAS* се детектират както при АК (50%), така и при ПК (42%) и при тези пациенти често се откриват ко-мутации в *TP53* или *STK11* гените.

Онкогените мутации на гена *KRAS* възникват най-вече в екзон 2 в кодон 12, порядко в кодон 13 (3–5%) и още по-рядко в екзон 3 кодон 61 (по-малко от 1%) (650). Тези промени са миссенс мутации, които нарушават способността на KRAS да хидролизира GTP, което води до конститутивно активиране на неговите ефекторни пътища и прогресия на рака (651). Поради високата му честота в АК са проведени няколко предклинични и клинични изследвания, търсещи ефективни терапевтични подходи и насочени към мутацията в *KRAS* гена (652). Все пак към днешна дата в рутинната клинична практика не се използват ефективни инхибитори на RAS и подходите за лечение на АК с *KRAS* мутация отразяват тези за лечение на НДРБД, при който липсва известна водеща мутация. Интересното е, че тютюнопушенето също оставя молекулен отпечатък върху *KRAS*, тъй като мутацията G12D е по-често срещана

при непушачи, докато мутациите G12C и G12V се срещат по-често сред бивши или настоящи пушачи (653). В допълнение при последната група се наблюдава по-висока честота на съпътстващи мутации в *TP53* или *STK11* и все повече доказателства предполагат, че тези съпътстващи мутации са свързани с уникални туморни характеристики и биологично поведение (654).

Неотдавнашен ретроспективен анализ на пациенти с напреднал НДРБД установи, че мутацията в *KRAS* е предиктор за намалена обща преживяемост, при лечение с цитотоксична химиотерапия (655). Освен това съвместното детектиране на мутация в *TP53* гена предсказва по-лоша прогноза (656). Повечето публикувани данни, включително мета-анализ на 22 проучвания, предполагат, че *KRAS* мутационният статус е отрицателен предиктор за EGFR-ТКИ (657). Съответно пациентите с мутация на *KRAS*, лекувани с EGFR-ТКИ, имат тенденция за по-лоши нива на обективно повлияване, намалено време до прогресия и намалена обща преживяемост в сравнение с пациенти без мутация в *KRAS* (658). Въпреки това, все още съществуват противоречиви резултати и не всички проучвания са стигнали до същите заключения (659–662). Повишената експресия на PD-L1 е открита при *KRAS* мутирани тумори и също така е показано, че активирането на ERK води до повишена експресия на PD-L1 при АК (663).

5.6.3. Мутации в *STK11* гена.

Мутациите в *STK11* гена (*LKB1*) се откриват в приблизително 5–30% от случаите на НДРБД и наскоро бяха идентифицирани като важен регулатор на резистентността към анти-PD -1/PD-L1 терапии (58,664–666). В изследваните от нас пациенти мутациите в *STK11* гена се откриват с еднаква честота при АК и ПК.

STK11 е серин-треонин киназа, която е важен регулатор на клетъчния метаболизъм и биоенергетика (667). Загубата на функциите на *STK11* увеличава използването на серин и синтеза на S-аденозил метионин (*SAM*), субстрат за множество ензими, включително *DNMT1* и *EZH2*, които повлияват на експресията на гените, влияещи на имунното разпознаване и редица интерферони (668,669). Мутациите в *STK11* водят до ремоделиране на туморна микросреда, ниска експресия на PD-L1, високи нива на гранулоцитен колония стимулиращ фактор и цитокини от IL-8 семейството, висока плътност на неутрофилоподобни клетки и производство на хемокини като IL-6 (670,671). Често се откриват съвместно мутации в *STK11* и *KRAS* гените и наличието им е свързано с по-лоша прогноза и намалена обща преживяемост при лечение с химио- и именотерапия (666,672).

Настоящото изследване има своите ограничения. Едно от тях е сравнително малкият брой пациенти, които са анализирани, което не позволява да се направи заключение за разлика в мутационния профили на различните морфологични подтипове АК и ПК. Също така не може да се извърши статистически анализ и да се корелира експресията на изследваните микроРНКи и мРНКи с откритите мутации. В допълнение, броят на гените, включени в панела, също е ограничен.

Независимо от това, получените резултати показват нуждата преминаване към използване рутинно на масивно паралелно секвениране, не само изследване на EGFR гена, но и на други гени, имащи отношение към основните молекулни пътища, свързани с онкогенезата при рака на белия дроб и определящи терапевтични подход.

Необходимо е цялостно проучване на всички промени на геномно, транскриптомно и епигеномно ниво, за да се получи пълната картина и да се натрупат по-задълбочени познания за молекулните механизми на белодробната карциногенеза.

6. ИЗВОДИ:

- Установен беше специфичен микроРНК профил за АК и ПК и бяха валидирани с RT-PCR съответно 8 при АК (3 с повишена експресия и 5 с намалена експресия) и 10 при ПК - 4 с повишена експресия и 6 с намалена експресия:
- Единадесет микроРНК показаха статистически значима разлика в експресията между АК и ПК. Тези микроРНК са следните: miR-101-3p, miR-144-3p, miR-744-3p, miR-95-3p, miR-650, miR-21-3p, miR-7, miR-375, miR-139-3p, miR-195-5p, miR-4689.
- miR-375 показва статистически значима разлика в експресията спрямо съответстващите нормални тъкани и значима разлика при сравнение между АК и ПК.
- Наблюдава се асоциация между експресията на miR-101-3p, miR-144-3p, miR-6785 и miR-195-5p с Т-статуса при АК. С N-статуса асоциира експресията на miR-650, miR-7, miR-375, miR-139-3p, miR-195-3p, PD-L1, MINA53, mTOR и с М-статуса miR-101-3p, miR-744-3p, miR-650, miR-7, miR-375, miR-139-3p, miR-195-5p, miR-195-3p, PD-L1, MINA53 и mTOR
- Променената експресията на miR-195-3p, miR-650, miR-4689 и miR-6785 се асоциира с АК с преобладаваща ацинарна компонента; miR-101-3p, miR-140-3p, miR-375, miR-139-3p, miR-195-5p се асоциират с АК с преобладаваща солидна компонента; miR-21-3p се асоциира с АК с преобладаваща микропапиларна компонента; miR-7 и miR-95 се асоциират с АК с преобладаваща папиларна компонента.
- При ПК асоциация с Т-статуса показаха – miR-4689, miR-101-3p, miR-140-3p, miR-144-3p, miR-6785, miR-21-3p, miR-375, miR-139-3p, miR-195-3p; с N-статуса - miR-95 и MINA53 и с М-статуса miR-744, miR-95, miR-195-5p, PD-L1, PTEN, EGFR, MINA53 и mTOR.
- Експресията на miR-101-3p, miR-144-3p, miR-95 и miR-195-5p е значимо променена при с ПК, при които липсва кератинизация. Значимо по-висока експресия на miR-21-3p се наблюдава при ПК с кератинизация.
- Корелационно-базираните мрежи при АК показаха разделянето на микроРНКи на два клъстера с различна биологична и функционална роля. При АК обогатените GO термини са свързани с метастазирането и прогресията,

фокалната адхезия, ЕМТ и транзицията към дребноклетъчен белодробен карцином и са засегнати са повече цитокин – медирани сигнални пътища.

- Клъстер 1 е силно свързан със специфични клетъчни процеси, улесняващи метастазирането в мозъка, клатрин-медирана ендоцитоза, липиден метаболизъм и др. и засяга основно иРНК-мишени, транскрипционни фактори.
- Клъстер 2 е свързан с транскрипцията, транслацията, организацията на клетъчната мембрана, апарата на Голджи, включително регулация на клетъчната диференциация и морфогенеза. След засегнатите иРНК преобладават кадхерини, протеин кинази, РНК-свързващи се белтъци, убиквитин лигази и транскрипционни фактори.
- При ПК корелационно-базираните мрежи от микроРНК с най-много връзки са miR-21-3p, miR-744-3p, miR-195-5p и miR-144-3p, като онтологичния анализ показва, че таргетите им се асоциират с клетъчния цикъл, адхезията и организацията на цитоскелета, клетъчния отговор към инсулин, вътреклетъчния транспорт, апоптоза, автофагия и убиквитилацията. След засегнатите иРНК преобладават РНК-свързващите се белтъци, кадхерините, протеин кинази и транскрипционни фактори.
- На база ко-експресията на 14-те микроРНК бяха тренирани и тествани четири модела за машинно самообучение. DT и RF не са подходящи за разграничаване на АК и ПК.
- Двата модела LG и SVM се представят по-добре от останалите два алгоритъма. Двата модела се характеризират с по-висока прецизност за сметка на по-ниска чувствителност. Като цяло и двата алгоритъма класифицират АК с висока положителна прогностична стойност, но класифицират погрешно ПК като АК.
- Не е подходящо самостоятелното използване за анализ на експресията на 14-те микроРНК като диагностичен панел за разграничаване на двата подтипа НДРБД. Необходимо е увеличаване на извадката и допълнителни маркери, които да се включат в анализа.
- Йерархичният клъстерен анализ на микроРНК експресия потвърди туморната хетерогенност и показва добро разделяне, корелиращо с различните хистологични подтипове при АК и ПК.

- Въвеждането на NGS таргетно секвениране в рутинната практика позволява установяване на характерните за всеки подтип движещи мутации, както и разграничаването на различни мутационни профили при АК и ПК, които имат отношение към наличните таргетни терапии.

7. ПРИНОСИ:

- За първи път беше извършено цялостно микроРНК профилиране на български пациенти с не-дребно клетъчен карцином на белия дроб.
- Установена и валидирана в независима извадка беше значимо променената експресия на 8 микроРНК при АК и 10 микроРНК при ПК.
- За първи път се доказва променена експресия на miR-744, miR-650, miR-4689 и miR-6785 при НДРБД
- Беше изградена корелационно-базирана мрежа при АК, която показва специфичната роля на miR-195-3p и miR-95-3p при различните морфологични подпипове АК.
- Беше изградена корелационно-базирана мрежа при ПК, която показва специфичната роля на miR-21-3p, miR-744-3p, miR-195-5p и miR-144-3p при различните морфологични подпипове ПК.
- Чрез онтологичен анализ бяха идентифицирани засегнатите сигнални пътища при двата подтипа НДРБД.
- Бяха тренирани и тествани модели за машинно самообучение за разграничаване на АК и ПК на база ко-експресията на изследваните микроРНК като логистичната регресия и методът на опорните вектори показаха задоволителни точност (accuracy), 80% и прецизност на модела (precision) – 91% за логистичната регресия и 99% за метода на опорните вектори;
- Валидираните данни за специфична експресия на микроРНКите и техните иРНК-мишени могат да послужат за разработването им като биомаркери и предиктивни фактори за отговор на имунната система и действието на имунотерапията при пациенти с НДРБД.

8. ПУБЛИКАЦИИ, УЧАСТИЯ В НАУЧНИ ПРОЯВИ И ПРОЕКТИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД.

8.1. Публикации.

- В. Петкова, А. Миткова, В. Митев, Р. Кънева, Молекулярно-генетични биомаркери при рак на белия дроб, *Studia Oncologica*, год. IX, брой 1, 2019
- Petkova V, Marinova D, Kyurkchian S, Stancheva G, Mekov E, Kachakova-Yordanova D, Slavova Y, Kostadinov D, Mitev V, Kaneva R. Expression analysis of MINA53: correlation with aberrantly expressed mRNAs and pathological features in non-small lung cancer *Biotechnology & Biotechnological Equipment* 35, no. 1 (2021): 1791-1798. doi: 10.1080/13102818.2021.2019117 **IF= 1.762 (2021), SJR 0.377 (2021);**
- Petkova V, Marinova D, Kyurkchian S, Stancheva G, Mekov E, Kachakova-Yordanova D, Slavova Y, Kostadinov D, Mitev V, Kaneva R. MiRNA expression profiling in adenocarcinoma and squamous cell lung carcinoma reveals both common and specific deregulated microRNAs. *Medicine (Baltimore)*. 2022 Aug 19;101(33):e30027. doi: 10.1097/MD.00000000000030027. PMID: 35984198; PMCID: PMC9388044. **IF = 1.817 (2021); Citations: 1**

8.2. Участия в научни прояви:

- V. Petkova, A. Mitkova, G. Stancheva, D. Kachakova, S. Giragosyan, D. Marinova, S. Yanina, V. Mitev, R. Kaneva. Expression analysis of miR-21, miR-205, EGFR, MINA53 and mTOR in Bulgarian patients with non-small cell lung cancer. Abstract 120, page 51. *EACR-AACR-SIC SPECIAL CONFERENCE 2017 PROCEEDINGS BOOK*. 24-27 JUNE 2017; FLORENCE, ITALY; EACR-AACR-SIC SPECIAL CONFERENCE 2017
- V. Petkova, A. Mitkova, G. Stancheva, D. Kachakova, S. Giragosyan, D. Marinova, S. Yanina, V. Mitev, R. Kaneva., Analysis of somatic mutations in lung adenocarcinomas and squamous cell carcinomas with targeted NGS. 2th Balkan Congress of Human Genetics, Plovdiv 8-10.09.2017
- V Petkova, A Mitkova, D Kachakova, G Stancheva, S Giragosyan, D Marinova, Y Slavova-Marinova, V Mitev, R Kaneva, PO-375 ANALYSIS OF MIRNA

EXPRESSION IN NON-SMALL CELL LUNG CANCER, ESMO Open 2018;3(Suppl 2):A1–A463

- V. Petkova, S. Giragosyan, D. Kachakova, A. Mitkova, D. Marinova, Y. Slavova-Marinova, V. Mitev, R. Kaneva, P17.25B Analysis of miRNA expression in non-small cell lung cancer, European Human Genetics Conference, Gothenburg, Sweden, JUNE 15–18, 2019
- V. Petkova, S. Giragosyan, D. Kachakova, A. Mitkova, D. Marinova, Y. Slavova-Marinova, V. Mitev, R. Kaneva, MiRNA expression profiling in non-small cell lung cancer, TMex online course in Translational Medicine 2020
- V. Petkova, S. G. Kyurkchyan, D. Kachakova – Yordanova, G. S. Stancheva, D. Marinova, E. Mekov, Y. Slavova-Marinova, V.Mitev, R. Kaneva, - Arm specific miRNAs expression analysis of hsa-miR-195 in non-small cell lung cancer, European Human Genetics Conference Virtual Conference, JUNE 12–15, 2021
- V. Petkova, S. Kyurkchyan, D. Kachakova - Yordanova, G. S. Stancheva, D. Marinova, E. Mekov, Y. Slavova-Marinova, V.Mitev, R. Kaneva, Comparative MiRNA expression analysis in squamous cell carcinoma and adenocarcinoma of the lung, EACR 2021 Virtual Congress: Innovative Cancer Science. 09-12 June 2021
- V. Petkova, S. Kyurkchyan, D. Kachakova - Yordanova, G. S. Stancheva, D. Marinova, E. Mekov, Y. Slavova-Marinova, V.Mitev, R. Kaneva, MiRNA expression in non-small cell lung cancer, European Human Genetics Conference 2022 from June 11-14, 2022
- V. Petkova, S. Kyurkchyan, D. Kachakova - Yordanova, G. S. Stancheva, D. Marinova, E. Mekov, Y. Slavova-Marinova, V.Mitev, R. Kaneva, P2-440: Analysis of MicroRNA expression data using machine learning in non-small cell lung cancer, 28TH CONGRESS OF THE EUROPEAN ASSOCIATION FOR CANCER RESEARCH, 20-23 June 2022, Seville, Spain

8.3. Участие в проекти:

- Фундаментални научни изследвания, 2016
Фонд научни изследвания
Договор №ДН03/14 от 19.12.2016
„Микро РНК профилиране при аденокарциноми и плоскоклетъчни карциноми на белия дроб при български пациенти“ (Участник)
- Млад Изследовател‘ 2017
Медицински Университет – София
Договор № Д-138/02.05.2017
“Експресионен анализ на микроРНК-и при недребноклетъчни белодробни карциноми“ (Водещ изследовател)
- Млад изследовател‘2018
Медицински Университет – София
Договор № Д-138/02.05.2017
“Експресионен анализ на ключови гени, свързани с избягване на имунния отговор при белодробен кацином“ (Водещ изследовател)
- Млад изследовател‘2019
Медицински Университет – София
Договор № Д-№ D-70/23.04.2019
„TP53 мутационен статус при не-дребноклетъчен белодробен карцином“ (Водещ изследовател)

9. БИБЛИОГРАФИЯ

1. WD T, E B, HK M-H, CC H. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart.
2. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med.* 2011 Dec;32(4):605–44.
3. Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, et al. Mutational signatures associated with tobacco smoking in human cancer. *Science* (80-). 2016;354(6312):618–22.
4. Zamay TN, Zamay GS, Kolovskaya OS, Zukov RA, Petrova MM, Gargaun A, et al. Current and Prospective Protein Biomarkers of Lung Cancer. *Cancers (Basel).* 2017;9(11).
5. Travis WD, Linnoila RI, Tsokos MG, Hitchcock CL, Cutler GB, Nieman L, et al. Neuroendocrine tumors of the lung with proposed criteria for large-cell neuroendocrine carcinoma. An ultrastructural, immunohistochemical, and flow cytometric study of 35 cases. *Am J Surg Pathol.* 1991;15(6):529–53.
6. Travis WD, Rush W, Flieder DB, Falk R, Fleming M V., Gal AA, et al. Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. *Am J Surg Pathol.* 1998 Aug;22(8):934–44.
7. Merrill RM, Henson DE, Barnes M. Conditional survival among patients with carcinoma of the lung. *Chest.* 1999;116(3):697–703.
8. Rekhtman N. Neuroendocrine Tumors of the Lung: An Update. *Arch Pathol Lab Med.* 2010;134(11):1628–38.
9. Travis WD. Advances in neuroendocrine lung tumors. *Ann Oncol.* 2010;21:vii65--vii71.
10. Ivanova LB, Vukov MI, Vassileva-Valerianova ZG. Thyroid Cancer Incidence in Bulgaria Before and After The Introduction of Universal Salt Iodization- An Analysis of National Cancer Registry Data. *Balkan Med J.* 2020;
11. Travis WD. Lung Cancer Pathology: Current Concepts. *Clin Chest Med.* 2020;41(1):67–85.
12. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances since the 2004 Classification. Vol. 10, *Journal of Thoracic Oncology.* Elsevier Inc; 2015. p. 1243–60.
13. Nitadori JI, Bograd AJ, Kadota K, Sima CS, Rizk NP, Morales EA, et al. Impact of micropapillary histologic subtype in selecting limited resection vs lobectomy for lung adenocarcinoma of 2cm or smaller. *J Natl Cancer Inst.* 2013 Aug 21;105(16):1212–20.
14. Damjanov I. ROBBINS AND COTRAN PATHOLOGIC BASIS OF DISEASE, 7TH EDITION. *Shock.* 2005;23(5):482–3.
15. Mountain CF. New Prognostic Factors in Lung Cancer. *Chest.* 1995;108(1):246–54.
16. Appasani K. Introduction. In: *MicroRNAs.* Cambridge University Press; 2007. p. 1–4.
17. Okamoto I, Yoshida, Kaneda. Molecularly targeted approaches herald a new era of non-small-cell lung cancer treatment. *Cancer Manag Res.* 2013;91.
18. Reungwetwattana T, Weroha SJ, Molina JR. Oncogenic Pathways, Molecularly Targeted Therapies, and Highlighted Clinical Trials in Non-small-Cell Lung Cancer (NSCLC). *Clin Lung Cancer.* 2012;13(4):252–66.
19. Intratumor heterogeneity comparison among different subtypes of non-small-cell lung cancer through multi-region tissue and matched ctDNA sequencing.
20. Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational Implications of Tumor Heterogeneity. *Clin Cancer Res.* 2015 Mar 15;21(6):1258–66.
21. Chrysanthakopoulos NA. Molecular and Biological Basis of Lung Cancer-Part I. *Glob J Cancer Case Reports.* 2020;01(02):1–22.
22. Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, et al. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nat Rev Cancer.* 2019;19(1):9.
23. Gyanchandani R, Priedigkeit N, Basudan A, Lee A V. Intratumor Heterogeneity. In: *Oncogenomics.* Elsevier; 2019. p. 17–30.
24. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al.

- Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505(7484):495–501.
25. Larsen JE, Minna JD. Molecular biology of lung cancer: clinical implications. *Clin Chest Med*. 2011 Dec;32(4):703–40.
 26. Viktorsson K, Lewensohn R, Zhivotovsky B. Systems biology approaches to develop innovative strategies for lung cancer therapy. *Cell Death Dis*. 2014;5(5):e1260--e1260.
 27. Li T, Kung H-J, Mack PC, Gandara DR. Genotyping and Genomic Profiling of Non\textendashSmall-Cell Lung Cancer: Implications for Current and Future Therapies. *J Clin Oncol*. 2013;31(8):1039–49.
 28. Sweeney SM, Cerami E, Baras A, Pugh TJ, Schultz N, Stricker T, et al. AACR project genie: Powering precision medicine through an international consortium. *Cancer Discov*. 2017;7(8):818–31.
 29. Mogi A, Kuwano H. TP53 mutations in nonsmall cell lung cancer. *J Biomed Biotechnol*. 2011;2011.
 30. Tanaka T, Watanabe M, Yamashita K. Potential therapeutic targets of TP53 gene in the context of its classically canonical functions and its latest non-canonical functions in human cancer. *Oncotarget*. 2018;9(22):16234–47.
 31. Karachaliou N, Sosa AE, Rosell R. Unraveling the genomic complexity of small cell lung cancer. *Transl lung cancer Res*. 2016 Aug 1;5(4):363–6.
 32. Shajani-Yi Z, de Abreu FB, Peterson JD, Tsongalis GJ. Frequency of Somatic TP53 Mutations in Combination with Known Pathogenic Mutations in Colon Adenocarcinoma, Non–Small Cell Lung Carcinoma, and Gliomas as Identified by Next-Generation Sequencing. *Neoplasia*. 2018;20(3):256–62.
 33. Zienolddiny S, Ryberg D, Arab MO, Skaug V, Haugen A. Loss of heterozygosity is related to p53 mutations and smoking in lung cancer. *Br J Cancer*. 2001;84(2):226.
 34. Aisner DL, Sholl LM, Berry LD, Rossi MR, Chen H, Fujimoto J, et al. The impact of smoking and tp53 mutations in lung adenocarcinoma patients with targetable mutations—the lung cancer mutation consortium (LCMC2). *Clin Cancer Res*. 2018;24(5):1038–47.
 35. Zhao J, Han Y, Li J, Chai R, Bai C. Prognostic value of KRAS/TP53/PIK3CA in non-small cell lung cancer. *Oncol Lett*. 2019 Feb 4;17(3):3233.
 36. Villalobos P, Wistuba II. Lung Cancer Biomarkers. *Hematol Oncol Clin North Am*. 2017;31(1):13–29.
 37. Sholl LM. Biomarkers in Lung Adenocarcinoma: A Decade of Progress. *Arch Pathol Lab Med*. 2014;139(4):469–80.
 38. Sharma S V, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer*. 2007;7(3):169–81.
 39. Nalejska E, M\kaczyńska E, Lewandowska MA. Prognostic and Predictive Biomarkers: Tools in Personalized Oncology. *Mol Diagnosis Ther*. 2014;18(3):273–84.
 40. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of Tumor Specimens at the Time of Acquired Resistance to EGFR-TKI Therapy in 155 Patients with Mutant Lung Cancers. *Clin Cancer Res*. 2013;19(8):2240–7.
 41. Arulananda S, Do H, Musafir A, Mitchell P, Dobrovic A, John T. Combination Osimertinib and Gefitinib in C797S and T790M EGFR-Mutated Non\textendashSmall Cell Lung Cancer. *J Thorac Oncol*. 2017;12(11):1728–32.
 42. Murtuza A, Bulbul A, Shen JP, Keshavarzian P, Woodward BD, Lopez-Diaz FJ, et al. Novel Third-Generation EGFR Tyrosine Kinase Inhibitors and Strategies to Overcome Therapeutic Resistance in Lung Cancer. *Cancer Res*. 2019;79(4):689–98.
 43. Schmid S, Li JJN, Leighl NB. Mechanisms of osimertinib resistance and emerging treatment options. *Lung Cancer*. 2020;147(May):123–9.
 44. Schalm SS, Dineen T, Lim SM, Park CW, Hsieh J, Woessner R, et al. 384P BLU-945, a highly potent and selective 4th generation EGFR TKI for the treatment of EGFR T790M/C797S resistant NSCLC. *Ann Oncol*. 2020 Nov 1;31:S1391.
 45. Blumenschein GR, Paulus R, Curran WJ, Robert F, Fossella F, Werner-Wasik M, et al. Phase II study of cetuximab in combination with chemoradiation in patients with stage IIIA/B non-small-cell lung cancer: RTOG 0324. *J Clin Oncol*. 2011 Jun 10;29(17):2312–8.

46. Han C, Cho J, Moon J, Kim H, Kim S, Shin D, et al. Clinical significance of insulin receptor substrate-I down-regulation in non-small cell lung cancer. *Oncol Rep.* 2006;
47. Wei EX, Anga AA, Martin SS, Jackson JG, Nordberg ML, Herrera GA, et al. EGFR expression as an ancillary tool for diagnosing lung cancer in cytology specimens. *Mod Pathol.* 2007 Sep;20(9):905–13.
48. Kempf E, Rousseau B, Besse B, Paz-Ares L. oncogene in lung cancer: focus on molecularly driven clinical trials. *Eur Respir Rev.* 2016;25(139):71–6.
49. Dogan S, Shen R, Ang DC, Johnson ML, D'Angelo SP, Paik PK, et al. Molecular Epidemiology of EGFR and KRAS Mutations in 3,026 Lung Adenocarcinomas: Higher Susceptibility of Women to Smoking-Related KRAS-Mu. *Clin Cancer Res.* 2012;18(22):6169–77.
50. Machado AL, Martins F, Seruca R, Velho, S rgiaCarvalho PD. Targeting the Tumor Microenvironment: An Unexplored Strategy for Mutant KRAS Tumors. *Cancers* 2019, Vol 11, Page 2010. 2019 Dec 13;11(12):2010.
51. Fujimoto J, Wistuba II. Current concepts on the molecular pathology of non-small cell lung carcinoma. *Semin Diagn Pathol.* 2014;31(4):306–13.
52. Garnett MJ, Marais R. Guilty as charged. *Cancer Cell.* 2004;6(4):313–9.
53. Baik CS, Myall NJ, Wakelee HA. Targeting BRAF-Mutant Non-Small Cell Lung Cancer: From Molecular Profiling to Rationally Designed Therapy. *Oncologist.* 2017;22(7):786–96.
54. Planchard D, Kim TM, Mazieres J, Quoix E, Riely G, Barlesi F, et al. Dabrafenib in patients with BRAFV600E-positive advanced non-small-cell lung cancer: a single-arm, multicentre, open-label, phase 2 trial. *Lancet Oncol.* 2016;17(5):642–50.
55. Origins, genetic landscape, and emerging therapies of small cell lung cancer.
56. Heist RS, Shim HS, Gingipally S, Mino-Kenudson M, Le L, Gainor JF, et al. MET Exon 14 Skipping in Non-Small Cell Lung Cancer. *Oncologist.* 2016;21(4):481–6.
57. Reungwetwattana T, Liang Y, Zhu V, Ou S-HI. The race to target MET exon 14 skipping alterations in non-small cell lung cancer: The Why, the How, the Who, the Unknown, and the Inevitable. *Lung Cancer.* 2017;103:27–37.
58. Facchinetti F, Bluthgen MV, Tergemina-Clain G, Faivre L, Pignon JP, Planchard D, et al. LKB1/STK11 mutations in non-small cell lung cancer patients: Descriptive analysis and prognostic value. *Lung Cancer.* 2017 Oct 1;112:62–8.
59. Berger AH, Brooks AN, Wu X, Shrestha Y, Chouinard C, Piccioni F, et al. High-throughput Phenotyping of Lung Cancer Somatic Mutations. *Cancer Cell.* 2017;32(6):884.
60. Bean J, Brennan C, Shih J-Y, Riely G, Viale A, Wang L, et al. MET amplification occurs with or without greaterT790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci.* 2007;104(52):20932–7.
61. Du X, Shao Y, Qin H, Tai Y, Gao H. ALK-rearrangement in non-small-cell lung cancer (NSCLC). *Thorac Cancer.* 2018;9(4):423–30.
62. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4ALK fusion gene in non-small-cell lung cancer. *Nature.* 2007;448(7153):561–6.
63. Palmirotta R, Quresmini D, Lovero D, Mannavola F, Dammacco F, Silvestris F. Gene Fusion in NSCLC. In: *Oncogenomics.* Elsevier; 2019. p. 443–64.
64. Giri S, Patel JK, Mahadevan D. Novel mutations in a patient with ALK-rearranged lung cancer. *N Engl J Med.* 2014;371(17):1655–6.
65. Tuononen K, Sarhadi VK, Wirtanen A, R nty M, Salmenkivi K, Knuutila A, et al. Targeted Resequencing Reveals greaterALKFusions in Non-Small Cell Lung Carcinomas Detected by FISH, Immunohistochemistry, and Real-Time RT-PCR: A Comparison of Four Methods. *BioMed Res Int.* 2013;2013:1–9.
66. Kim MH, Shim HS, Kang DR, Jung JY, Lee CY, Kim DJ, et al. Clinical and prognostic implications of ALK and ROS1 rearrangements in never-smokers with surgically resected lung adenocarcinoma. *Lung Cancer.* 2014 Mar;83(3):389–95.
67. Davies KD, Doebele RC. Molecular pathways: ROS1 fusion proteins in cancer. *Clin Cancer Res.* 2013 Aug 1;19(15):4040–5.
68. Lin JJ, Shaw AT. Recent Advances in Targeting ROS1 in Lung Cancer. *J Thorac Oncol.* 2017

- Nov 1;12(11):1611–25.
69. Gainor JF, Tseng D, Yoda S, Dagogo-Jack I, Friboulet L, Lin JJ, et al. Patterns of Metastatic Spread and Mechanisms of Resistance to Crizotinib in greaterROS1-Positive Non\textendashSmall-Cell Lung Cancer. *JCO Precis Oncol.* 2017;(1):1–13.
 70. Kerr KM, Bubendorf L, Edelman MJ, Marchetti A, Mok T, Novello S, et al. Second ESMO consensus conference on lung cancer: pathology and molecular biomarkers for non-small-cell lung cancer. *Ann Oncol.* 2014;25(9):1681–90.
 71. Ju YS, Lee W-C, Shin J-Y, Lee S, Bleazard T, Won J-K, et al. A transforming greaterKIF5B and greaterRET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* 2011;22(3):436–45.
 72. Mendoza L. Clinical development of RET inhibitors in RET-rearranged non-small cell lung cancer: Update. *Oncol Rev.* 2018;12(2).
 73. Ju YS, Lee WC, Shin JY, Lee S, Bleazard T, Won JK, et al. A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* 2012 Mar;22(3):436–45.
 74. Roskoski R, Sadeghi-Nejad A. Role of RET protein-tyrosine kinase inhibitors in the treatment RET-driven thyroid and lung cancers. *Pharmacol Res.* 2018;128:1–17.
 75. Wang R, Hu H, Pan Y, Li Y, Ye T, Li C, et al. greaterRET Fusions Define a Unique Molecular and Clinicopathologic Subtype of Non\textendashSmall-Cell Lung Cancer. *J Clin Oncol.* 2012;30(35):4352–9.
 76. Alberti L, Carniti C, Miranda C, Roccato E, Pierotti MA. RET and NTRK1 proto-oncogenes in human diseases. *J Cell Physiol.* 2003;195(2):168–86.
 77. Vaishnavi A, Capelletti M, Le AT, Kako S, Butaney M, Ercan D, et al. Oncogenic and drug-sensitive NTRK1 rearrangements in lung cancer. *Nat Med.* 2013 Nov;19(11):1469–72.
 78. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science (80-).* 2011;331(6024):1565–70.
 79. Vareki SM, Garrigós C, Duran I. Biomarkers of response to PD-1/PD-L1 inhibition. *Crit Rev Oncol Hematol.* 2017;116:116–24.
 80. Dal Bello MG, Alama A, Coco S, Vanni I, Grossi F. Understanding the checkpoint blockade in lung cancer immunotherapy. *Drug Discov Today.* 2017;22(8):1266–73.
 81. Philips GK, Atkins M. Therapeutic uses of anti-PD-1 and anti-PD-L1 antibodies. *Int Immunol.* 2014;27(1):39–46.
 82. Guan J, Lim KS, Mekhail T, Chang C-C. Programmed Death Ligand-1 (PD-L1) Expression in the Programmed Death Receptor-1 (PD-1)/PD-L1 Blockade: A Key Player Against Various Cancers. *Arch Pathol Lab Med.* 2017;141(6):851–61.
 83. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science (80-).* 2018;359(6382):1350–5.
 84. Schultheis AM, Scheel AH, Ozretić L, George J, Thomas RK, Hagemann T, et al. PD-L1 expression in small cell neuroendocrine carcinomas. *Eur J Cancer.* 2015;51(3):421–6.
 85. Teixidó C, Vilariño N, Reyes R, Reguart N. PD-L1 expression testing in non-small cell lung cancer. *Ther Adv Med Oncol.* 2018;10:1758835918763493.
 86. Kouzarides T. Chromatin modifications and their function. *Cell.* 2007 Feb 23;128(4):693–705.
 87. Mellor J. The dynamics of chromatin remodeling at promoters. *Mol Cell.* 2005 Jul 22;19(2):147–57.
 88. Gibney ER, Nolan CM. Epigenetics and gene expression. *Hered* 2010 1051. 2010 May 12;105(1):4–13.
 89. Tian LJ, Wu YP, Wang D, Zhou ZH, Xue S Bin, Zhang DY, et al. Upregulation of long noncoding RNA (lncRNA) X-inactive specific transcript (XIST) is associated with cisplatin resistance in non-small cell lung cancer (NSCLC) by downregulating microRNA-144-3p. *Med Sci Monit.* 2019 Oct 29;25:8095–104.
 90. Zhang Q, Thakur C, Shi J, Sun J, Fu Y, Stemmer P, et al. New discoveries of mdig in the epigenetic regulation of cancers. *Semin Cancer Biol.* 2019;57(July):27–35.
 91. Field RW, Withers BL. Occupational and Environmental Causes of Lung Cancer. *Clin Chest Med.* 2012 Dec;33(4):681–703.
 92. Shi J, Thakur C, Zhao Y, Li Y, Nie L, Zhang Q, et al. Pathological and prognostic indications

- of the mdig gene in human lung cancer. *Cell Physiol Biochem*. 2021 Jan 1;55:13–28.
93. Zhang Y, Lu Y, Yuan B-Z, Castranova V, Shi X, Stauffer JL, et al. The Human mineral dust-induced gene, mdig, is a cell growth regulating gene associated with lung cancer. *Oncogene* 2005 2431. 2005 May 9;24(31):4873–82.
 94. Teye K, Tsuneoka M, Arima N, Koda Y, Nakamura Y, Ueta Y, et al. Increased Expression of a Myc Target Gene Mina53 in Human Colon Cancer. *Am J Pathol*. 2004;164(1):205–16.
 95. Zhou H, Geng F, Chen Y, Du J, Zhang X, Liu B, et al. The mineral dust-induced gene, mdig, regulates angiogenesis and lymphangiogenesis in lung adenocarcinoma by modulating the expression of VEGF-A/C/D via EGFR and HIF-1 α signaling. *Oncol Rep*. 2021;45(5):110001.
 96. Zhang Q, Thakur C, Fu Y, Bi Z, Wadgaonkar P, Xu L, et al. Mdig promotes oncogenic gene expression through antagonizing repressive histone methylation markers. *Theranostics*. 2020;10(2):602–14.
 97. Thakur C, Lu Y, Sun J, Yu M, Chen B, Chen F. Increased expression of mdig predicts poorer survival of the breast cancer patients. *Gene*. 2014 Feb 10;535(2):218.
 98. Huang MY, Xuan F, Liu W, Cui HJ. MINA controls proliferation and tumorigenesis of glioblastoma by epigenetically regulating cyclins and CDKs via H3K9me3 demethylation. *Oncogene*. 2017 Jan 19;36(3):387–96.
 99. Tsuneoka M, Fujita H, Arima N, Teye K, Okamura T, Inutsuka H, et al. Mina53 as a Potential Prognostic Factor for Esophageal Squamous Cell Carcinoma. *Clin Cancer Res*. 2004 Nov 1;10(21):7347–56.
 100. Ishizaki H, Yano H, Tsuneoka M, Ogasawara S, Akiba J, Nishida N, et al. Overexpression of the myc target gene Mina53 in advanced renal cell carcinoma. *Pathol Int*. 2007;57(10):672–80.
 101. S O, M K, O N, J A, M T, H Y. Accelerated expression of a Myc target gene Mina53 in aggressive hepatocellular carcinoma. *Hepatol Res*. 2010 Mar;40(4):330–6.
 102. XP T, WG D, Q Z, ZR Y, XF L, MH A. Potential effects of Mina53 on tumor growth in human pancreatic cancer. *Cell Biochem Biophys*. 2014;69(3):619–25.
 103. Wu K, Li L, Thakur C, Lu Y, Zhang X, Yi Z, et al. Proteomic Characterization of the World Trade Center dust-activated mdig and c-myc signaling circuit linked to multiple myeloma. *Sci Reports* 2016 61. 2016 Nov 11;6(1):1–12.
 104. Q Z, CM H, YS Y, CH H, Q Z, NZ L. Expression of Mina53 and its significance in gastric carcinoma. *Int J Biol Markers*. 2008;23(2):83–8.
 105. C T, F C. Current understanding of mdig/MINA in human cancers. *Genes Cancer*. 2015;6(7–8):1–15.
 106. Lian SL, Mihi B, Koyanagi M, Nakayama T, Bix M. A SNP uncoupling Mina expression from the TGF β signaling pathway. *Immunity, Inflamm Dis*. 2018 Mar 1;6(1):58.
 107. Thakur C, Wolfarth M, Sun J, Zhang Y, Lu Y, Battelli L, et al. Oncoprotein mdig contributes to silica-induced pulmonary fibrosis by altering balance between Th17 and Treg T cells. *Oncotarget*. 2015;6(6):3722–36.
 108. Hemmers S, Mowen KA. TH2 bias: Mina tips the balance. *Nat Immunol* 2009 108. 2009;10(8):806–8.
 109. Thakur C, Chen B, Li L, Zhang Q, Yang Z-Q, Chen F. Loss of mdig expression enhances DNA and histone methylation and metastasis of aggressive breast cancer. *Signal Transduct Target Ther*. 2018 Dec 1;3(1).
 110. Yu M, Sun J, Thakur C, Chen B, Lu Y, Zhao H, et al. Paradoxical Roles of Mineral Dust Induced Gene on Cell Proliferation and Migration/Invasion. *PLoS One*. 2014 Feb 4;9(2):87998.
 111. Komiya K, Sueoka-Aragane N, Sato A, Hisatomi T, Sakuragi T, Mitsuoka M, et al. Expression of Mina53, a novel c-Myc target gene, is a favorable prognostic marker in early stage lung cancer. *Lung Cancer*. 2010;69(2):232–8.
 112. Aziz N, Hong YH, Jo MK, Kim JK, Kim K-H, Ashktorab H, et al. Molecular Signatures of JMJD10/MINA53 in Gastric Cancer. *Cancers* 2020, Vol 12, Page 1141. 2020 May 2;12(5):1141.
 113. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2013;42(D1):D68–D73.
 114. Chen L, Sun H, Wang C, Yang Y, Zhang M, Wong G. miRNA arm switching identifies novel

- tumour biomarkers. *EBioMedicine*. 2018 Dec 1;38:37–46.
115. Kunej T, Godnic I, Horvat S, Zorc M, Calin GA. Cross Talk Between MicroRNA and Coding Cancer Genes. *Cancer J*. 2012;18(3):223–31.
 116. Hu Z, Chen J, Tian T, Zhou X, Gu H, Xu L, et al. Genetic variants of miRNA sequences and non-small cell lung cancer survival. *J Clin Invest*. 2008;118(7):2600–8.
 117. Zhang Y, Sui J, Shen X, Li C, Yao W, Hong W, et al. Differential expression profiles of microRNAs as potential biomarkers for the early diagnosis of lung cancer. *Oncol Rep*. 2017;37(6):3543–53.
 118. Liu C-G, Calin GA, Volinia S, Croce CM. MicroRNA expression profiling using microarrays. *Nat Protoc*. 2008;3(4):563–78.
 119. Lu J, Zhan Y, Feng J, Luo J, Fan S. MicroRNAs associated with therapy of non-small cell lung cancer. *Int J Biol Sci*. 2018;14(4):390–7.
 120. Koo T, Cho BJ, Kim DH, Park JM, Choi EJ, Kim HH, et al. MicroRNA-200c increases radiosensitivity of human cancer cells with activated EGFR-associated signaling. *Oncotarget*. 2017;8(39):65457–68.
 121. Zhou G, Zhang F, Guo Y, Huang J, Xie Y, Yue S, et al. miR-200c enhances sensitivity of drug-resistant non-small cell lung cancer to gefitinib by suppression of PI3K/Akt signaling pathway and inhibites cell migration via targeting ZEB1. *Biomed & Pharmacother = Biomed & Pharmacother*. 2017;85:113–9.
 122. Gao W, Yu Y, Cao H, Shen H, Li X, Pan S, et al. Deregulated expression of miR-21, miR-143 and miR-181a in non small cell lung cancer is related to clinicopathologic characteristics or patient prognosis. *Biomed & Pharmacother*. 2010;64(6):399–408.
 123. Zhao J, Fu W, Liao H, Dai L, Jiang Z, Pan Y, et al. The regulatory and predictive functions of miR-17 and miR-92 families on cisplatin resistance of non-small cell lung cancer. *BMC Cancer*. 2015;15(1).
 124. Hojbjerg JA, Ebert EBF, Clement MS, Winther-Larsen A, Meldgaard P, Sorensen B. Circulating miR-30b and miR-30c predict erlotinib response in EGFR-mutated non-small cell lung cancer patients. *Lung Cancer*. 2019;135:92–6.
 125. Remon J, Alvarez-Berdugo D, Majem M, Moran T, Reguart N, Lianes P. miRNA-197 and miRNA-184 are associated with brain metastasis in EGFR-mutant lung cancers. *Clin Transl Oncol*. 2016;18(2):153–9.
 126. Jiang Z, Yin J, Fu W, Mo Y, Pan Y, Dai L, et al. miRNA 17 Family Regulates Cisplatin-Resistant and Metastasis by Targeting TGFbetaR2 in NSCLC. *PLoS One*. 2014 Apr 10;9(4):e94639.
 127. Su W, Mo Y, Wu F, Guo K, Li J, Luo Y, et al. miR-135b reverses chemoresistance of non-small cell lung cancer cells by downregulation of FZD1. *Biomed Pharmacother*. 2016 Dec 1;84:123–9.
 128. Sun CC, Li SJ, Zhang F, Zhang YD, Zuo ZY, Xi YY, et al. The Novel miR-9600 Suppresses Tumor Progression and Promotes Paclitaxel Sensitivity in Non-small-cell Lung Cancer Through Altering STAT3 Expression. *Mol Ther Nucleic Acids*. 2016;5(11):e387.
 129. Hatley ME, Patrick DM, Garcia MR, Richardson JA, Bassel-Duby R, van Rooij E, et al. Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21. *Cancer Cell*. 2010;18(3):282–93.
 130. Zhang J guang, Wang J jun, Zhao F, Liu Q, Jiang K, Yang G hai. MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). *Clin Chim Acta*. 2010 Jun;411(11–12):846–52.
 131. Xiao-chun W, Wei W, Zhu-Bo Z, Jing Z, Xiao-Gang T, Jian-Chao L. Overexpression of miRNA-21 promotes radiation-resistance of non-small cell lung cancer. *Radiat Oncol*. 2013;8(1):146.
 132. The role of hypoxia-induced miR-210 in cancer progression. - PubMed - NCBI.
 133. Yoshino H, Yonemori M, Miyamoto K, Tatarano S, Kofuji S, Nohata N, et al. microRNA-210-3p depletion by CRISPR/Cas9 promoted tumorigenesis through revival of TWIST1 in renal cell carcinoma. *Oncotarget*. 2017;8(13):20881–94.
 134. Chen Q, Hu H, Jiao D, Yan J, Xu W, Tang X, et al. miR-126-3p and miR-451a correlate with clinicopathological features of lung adenocarcinoma: The underlying molecular mechanisms.

- Oncol Rep. 2016;36(2):909–17.
135. Jiang L, Zhang H, Tang J. MiR-30a: A Novel Biomarker and Potential Therapeutic Target for Cancer. *Journal of Oncology*. 2018.
 136. Tang R, Liang L, Luo D, Feng Z, Huang Q, He R, et al. Downregulation of MiR-30a is Associated with Poor Prognosis in Lung Cancer. *Med Sci Monit*. 2015;21:2514–20.
 137. Kumarswamy R, Mudduluru G, Ceppi P, Muppala S, Kozlowski M, Niklinski J, et al. MicroRNA-30a inhibits epithelial-to-mesenchymal transition by targeting Snail and is downregulated in non-small cell lung cancer. *Int J Cancer*. 2011;130(9):2044–53.
 138. Xu X, Jin S, Ma Y, Fan Z, Yan Z, Li W, et al. miR-30a-5p enhances paclitaxel sensitivity in non-small cell lung cancer through targeting BCL-2 expression. *J Mol Med (Berl)*. 2017;95(8):861–71.
 139. Yongchun Z, Linwei T, Xicai W, Lianhua Y, Guangqiang Z, Ming Y, et al. MicroRNA-195 inhibits non-small cell lung cancer cell proliferation, migration and invasion by targeting MYB. *Cancer Lett*. 2014 May 28;347(1):65–74.
 140. Chang KW, Kao SY, Wu YH, Tsai MM, Tu HF, Liu CJ, et al. Passenger strand miRNA miR-31* regulates the phenotypes of oral cancer cells by targeting RhoA. *Oral Oncol*. 2013 Jan;49(1):27–33.
 141. Chen L, Sun H, Wang C, Yang Y, Zhang M, Wong G. miRNA arm switching identifies novel tumour biomarkers. *EBioMedicine*. 2018;38:37–46.
 142. Flamini V, Dudley E, Jiang WG, Cui Y. Distinct mechanisms by which two forms of miR-140 suppress the malignant properties of lung cancer cells. *Oncotarget*. 2018;9(92):36474–91.
 143. Scagliotti G V, Novello S. The role of the insulin-like growth factor signaling pathway in non-small cell lung cancer and other solid tumors. *Cancer Treat Rev*. 2012;38(4):292–302.
 144. Hu G, Wei Y, Kang Y. The Multifaceted Role of MTDH/AEG-1 in Cancer Progression. *Clin Cancer Res*. 2009;15(18):5615–20.
 145. Sachdeva M, Zhu S, Wu F, Wu H, Walia V, Kumar S, et al. p53 represses c-Myc through induction of the tumor suppressor greatermiR-145. *Proc Natl Acad Sci*. 2009;106(9):3207–12.
 146. Suh SO, Chen Y, Zaman MS, Hirata H, Yamamura S, Shahryari V, et al. MicroRNA-145 is regulated by DNA methylation and p53 gene mutation in prostate cancer. *Carcinogenesis*. 2011;32(5):772–8.
 147. Du L, Schageman JJ, Subauste MC, Saber B, Hammond SM, Prudkin L, et al. miR-93, miR-98 and miR-197 regulate expression of tumor suppressor gene FUS1. *Mol Cancer Res*. 2009 Aug;7(8):1234.
 148. Fujita Y, Yagishita S, Hagiwara K, Yoshioka Y, Kosaka N, Takeshita F, et al. The Clinical Relevance of the miR-197/CKS1B/STAT3-mediated PD-L1 Network in Chemoresistant Non-small-cell Lung Cancer. *Mol Ther*. 2015;23(4):717–27.
 149. Remon J, Alvarez-Berdugo D, Majem M, Moran T, Reguart N, Lianes P. miRNA-197 and miRNA-184 are associated with brain metastasis in EGFR-mutant lung cancers. *Clin Transl Oncol*. 2015;18(2):153–9.
 150. Dacic S, Kelly L, Shuai Y, Nikiforova MN. MiRNA expression profiling of lung adenocarcinomas: Correlation with mutational status. *Mod Pathol*. 2010 Dec 3;23(12):1577–82.
 151. Del Vecovo V, Grasso M, Barbareschi M, Denti MA. MicroRNAs as lung cancer biomarkers. *World J Clin Oncol*. 2014 Oct 10;5(4):604.
 152. Mitchell TM. *Machine Learning*.
 153. Ohler U, Liao G, Niemann H, Rubin GM. Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol*. 2002/12/20. 2002;3(12):RESEARCH0087–RESEARCH0087.
 154. Degroeve S, De Baets B, Van de Peer Y, Rouze P. Feature subset selection for splice site prediction. *Bioinformatics*. 2002;18(Suppl 2):S75–83.
 155. Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J Mol Biol*. 1990;212(4):563–78.
 156. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39(3):311–8.

157. Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, et al. A genomic code for nucleosome positioning. *Nature*. 2006/07/19. 2006 Aug 17;442(7104):772–8.
158. Picardi E, Pesole G. *Computational Methods for Ab Initio and Comparative Gene Finding*. Methods in Molecular Biology. Humana Press; 2009. p. 269–84.
159. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet*. 2015;16(6):321–32.
160. Gandhi L, Rodríguez-Abreu D, Gadgeel S, Esteban E, Felip E, De Angelis F, et al. Pembrolizumab plus Chemotherapy in Metastatic Non–Small-Cell Lung Cancer. *N Engl J Med*. 2018;378(22):2078–92.
161. Paz-Ares L, Luft A, Vicente D, Tafreshi A, Gümüş M, Mazières J, et al. Pembrolizumab plus Chemotherapy for Squamous Non–Small-Cell Lung Cancer. *N Engl J Med*. 2018;379(21):2040–51.
162. Thomas A, Liu S V, Subramaniam DS, Giaccone G. Refining the treatment of NSCLC according to histological and molecular subtypes. *Nat Rev Clin Oncol*. 2015;12(9):511–26.
163. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Peadarallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet*. 2016/05/09. 2016 Jun;48(6):607–16.
164. Network CGAR. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012/09/09. 2012 Sep 27;489(7417):519–25.
165. Network CGAR. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014/07/09. 2014 Jul 31;511(7511):543–50.
166. Shiao T-H, Chang Y-L, Yu C-J, Chang Y-C, Hsu Y-C, Chang S-H, et al. Epidermal Growth Factor Receptor Mutations in Small Cell Lung Cancer: A Brief Report. *J Thorac Oncol*. 2011;6(1):195–8.
167. Lam VK, Tran HT, Banks KC, Lanman RB, Rinsurongkawong W, Peled N, et al. Targeted Tissue and Cell-Free Tumor DNA Sequencing of Advanced Lung Squamous-Cell Carcinoma Reveals Clinically Significant Prevalence of Actionable Alterations. *Clin Lung Cancer*. 2019;20(1):30-36.e3.
168. Tang M, Abbas HA, Negrao M V, Ramineni M, Hu X, Hubert SM, et al. The histologic phenotype of lung cancers is associated with transcriptomic features rather than genomic characteristics. *Nat Commun*. 2021 Dec 6;12(1):7081.
169. Beer MA, Tavazoie S. Predicting Gene Expression from Sequence. *Cell*. 2004;117(2):185–98.
170. Karlič R, Chung H-R, Lasserre J, Vlahovicek K, Vingron M. Histone modification levels are predictive for gene expression. *Proc Natl Acad Sci U S A*. 2010/02/01. 2010 Feb 16;107(7):2926–31.
171. Ouyang Z, Zhou Q, Wong WH. ChIP-Seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. *Proc Natl Acad Sci U S A*. 2009/12/07. 2009 Dec 22;106(51):21521–6.
172. Friedman N. Inferring Cellular Networks Using Probabilistic Graphical Models. *Science (80-)*. 2004;303(5659):799–805.
173. Rosa GJM. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* by HASTIE, T., TIBSHIRANI, R., and FRIEDMAN, J. *Biometrics*. 2010;66(4):1315.
174. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2018/04/03. 2018 Apr;15(4):233–4.
175. Xu C, Jackson SA. Machine learning and complex biological data. *Genome Biol*. 2019 Apr 16;20(1):76.
176. D’andrade PN, Fulmer-Smentek S. Agilent MicroRNA microarray profiling system. *Methods Mol Biol*. 2012;822:85–102.
177. Nguyen D V., Arpat AB, Wang N, Carroll RJ. DNA microarray experiments: biological and technological aspects. *Biometrics*. 2002;58(4):701–17.
178. Quackenbush J. Microarray data normalization and transformation. *Nat Genet* 2002 324. 2002;32(4):496–501.
179. Ransohoff DF. How to improve reliability and efficiency of research about molecular markers: roles of phases, guidelines, and study design. *J Clin Epidemiol*. 2007 Dec 1;60(12):1205–19.
180. Huber W, von Heydebreck A, Sueltmann H, Poustka A, Vingron M. Parameter estimation for

- the calibration and variance stabilization of microarray data. *Stat Appl Genet Mol Biol*. 2003 Dec 22;2(1).
181. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*. 2003 Feb 15;31(4):e15.
 182. Schadt EE, Li C, Ellis B, Wong WH. Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. *J Cell Biochem Suppl*. 2001;Suppl 37(SUPPL. 37):120–5.
 183. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*. 2002;30(4).
 184. Ratert N, Meyer HA, Jung M, Mollenkopf HJ, Wagner I, Miller K, et al. Reference miRNAs for miRNAome Analysis of Urothelial Carcinomas. *PLoS One*. 2012 Jun 20;7(6):e39309.
 185. Shwetha S, Gouthamchandra K, Chandra M, Ravishankar B, Khaja MN, Das S. Circulating miRNA profile in HCV infected serum: novel insight into pathogenesis. *Sci Reports* 2013 31. 2013 Apr 3;3(1):1–7.
 186. Ringnér M. What is principal component analysis? *Nat Biotechnol* 2008 263. 2008 Mar;26(3):303–4.
 187. Introduction to Probability and Statistics - William Mendenhall, Robert J. Beaver, Barbara M. Beaver - Google Книги [Internet]. [cited 2022 Nov 12]. Available from: <https://books.google.com.pk/books?id=fQsKAAAAQBAJ&printsec=frontcover#v=onepage&q&f=false>
 188. Batushansky A, Toubiana D, Fait A. Correlation-Based Network Generation, Visualization, and Analysis as a Powerful Tool in Biological Studies: A Case Study in Cancer Cell Metabolism. *Biomed Res Int*. 2016;2016.
 189. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020 Mar 1;17(3):261–72.
 190. Caswell TA, Droettboom M, Lee A, Andrade ES de, Hunter J, Hoffmann T, et al. matplotlib/matplotlib: REL: v3.4.2. 2021 May 8;
 191. Waskom M, Botvinnik O, O’Kane D, Hobson P, Lukauskas S, Gemperline DC, et al. mwaskom/seaborn: v0.8.1 (September 2017). 2017 Sep 3;
 192. Hagberg hagberg AA, -Los lanlgov, Schult DA, Swart swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. 2008.
 193. Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, et al. The multiMiR R package and database: Integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res*. 2014 Jun 27;42(17):133.
 194. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019 Jan 8;47(D1):D419–26.
 195. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. *Nucleic Acids Res*. 2021 Jan 8;49(D1):D613–21.
 196. Fang Z. GSEAPy: Gene Set Enrichment Analysis in Python. 2020 Apr 11;
 197. Interpretable Machine Learning [Internet]. [cited 2022 Nov 12]. Available from: <https://christophm.github.io/interpretable-ml-book/>
 198. Pedregosa FABIANPEDREGOSA F, Michel V, Grisel OLIVIERGRISEL O, Blondel M, Prettenhofer P, Weiss R, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011;12(85):2825–30.
 199. Samariya D, Aryal S, Ting KM, Ma J. A new effective and efficient measure for outlying aspect mining. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2020 Apr 28;12343 LNCS:463–74.
 200. Jo J-M. Effectiveness of Normalization Pre-Processing of Big Data to the Machine Learning Performance. *J Korea Inst Electron Commun Sci*. 2019;14(3):547–52.
 201. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*. 2013 Mar 22;14(1):1–16.
 202. Bonaccorso G. Machine Learning Algorithms Reference guide for popular algorithms for data

- science and machine learning. 2017;
203. Liu Y, Liao S, Jiang S, Ding L, Lin H, Wang W. Fast Cross-Validation for Kernel-Based Algorithms. *IEEE Trans Pattern Anal Mach Intell.* 2020 May 1;42(5):1083–96.
 204. Fushiki T. Estimation of prediction error by using K-fold cross-validation. *Stat Comput* 2009 212. 2009 Oct 10;21(2):137–46.
 205. Hans C. Bayesian lasso regression. *Biometrika.* 2009 Dec 1;96(4):835–45.
 206. Cortes C, Mohri M, Rostamizadeh A. L2 Regularization for Learning Kernels. *Proc 25th Conf Uncertain Artif Intell UAI 2009.* 2012 May 9;109–16.
 207. De Mol C, De Vito E, Rosasco L. Elastic-Net Regularization in Learning Theory. *J Complex.* 2008 Jul 22;25(2):201–30.
 208. Kelley CT. Iterative Methods for Linear and Nonlinear Equations. *Iterative Methods for Linear and Nonlinear Equations.* 1995 Jan;
 209. Zhou J, Gandomi AH, Chen F, Holzinger A. Evaluating the Quality of Machine Learning Explanations: A Survey on Methods and Metrics. *Electron* 2021, Vol 10, Page 593. 2021 Mar 4;10(5):593.
 210. Susmaga R. Confusion Matrix Visualization. *Intell Inf Process Web Min.* 2004;107–16.
 211. Kannan R, Vasanthi V. Machine learning algorithms with ROC curve for predicting and diagnosing the heart disease. *SpringerBriefs Appl Sci Technol.* 2019;63–72.
 212. Huang J, Ling CX. Using AUC and accuracy in evaluating learning algorithms. *IEEE Trans Knowl Data Eng.* 2005 Mar;17(3):299–310.
 213. Lance GN, Williams WT. A General Theory of Classificatory Sorting Strategies1. *Hierarchical Systems. Comput J.* 1967 Feb 1;9(4):373–80.
 214. Quackenbush J. Microarray data normalization and transformation. *Nat Genet.* 2002;32(S4):496–501.
 215. Halvorsen AR, Ragle Aure M, Õjlert ÅK, Brustugun OT, Solberg S, Nebdal D, et al. Identification of microRNAs involved in pathways which characterize the expression subtypes of NSCLC. *Mol Oncol.* 2019 Dec 1;13(12):2604–15.
 216. Bjaanæs MM, Halvorsen AR, Solberg S, Jørgensen L, Dragani TA, Galvan A, et al. Unique microRNA-profiles in EGFR-mutated lung adenocarcinomas. *Int J Cancer.* 2014 Oct 15;135(8):1812–21.
 217. Heberle H, Meirelles VG, da Silva FR, Telles GP, Minghim R. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics.* 2015 May 22;16(1).
 218. Zhang J, Duy Le T, Liu L, He J, Li J. Identifying miRNA synergistic regulatory networks in heterogeneous human data via network motifs. *Mol Biosyst.* 2016;12(2):454–63.
 219. Xu J, Li C-X, Li Y-S, Lv J-Y, Ma Y, Shao T-T, et al. MiRNA–miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. *Nucleic Acids Res.* 2011 Feb 1;39(3):825–36.
 220. Deng S, Qi J, Stephen M, Qiu L, Yang H. Network-based identification of reliable bio-markers for cancers. *J Theor Biol.* 2015;383:20–7.
 221. Bracken CP, Scott HS, Goodall GJ. A network-biology perspective of microRNA function and dysfunction in cancer. *Nat Rev Genet* 2016 1712. 2016 Oct 31;17(12):719–32.
 222. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell.* 2011 Mar 4;144(5):646–74.
 223. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022 Jan 1;12(1):31–46.
 224. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019 Jan 8;47(D1):D607–13.
 225. Delzell DAP, Magnuson S, Peter T, Smith M, Smith BJ. Machine Learning and Feature Selection Methods for Disease Classification With Application to Lung Cancer Screening Image Data. *Front Oncol.* 2019;9(December):1–8.
 226. Sherafatian M, Arjmand F. Decision tree-based classifiers for lung cancer diagnosis and subtyping using TCGA miRNA expression data. *Oncol Lett.* 2019/06/10. 2019 Aug;18(2):2125–31.
 227. Mingers J. An Empirical Comparison of Pruning Methods for Decision Tree Induction. *Mach Learn* 1989 42. 1989;4(2):227–43.

228. Jayatilake SMDAC, Ganegoda GU. Involvement of Machine Learning Tools in Healthcare Decision Making. *J Healthc Eng.* 2021;2021.
229. Qi Y. Random Forest for Bioinformatics. *Ensemble Machine Learning.* Springer US; 2012. p. 307–23.
230. Breiman L. Random forests. *Mach Learn.* 2001 Oct;45(1):5–32.
231. Cortes C, Vapnik V, Saitta L. Support-vector networks. *Mach Learn* 1995 203. 1995 Sep;20(3):273–97.
232. Cai Z, Poulos RC, Liu J, Zhong Q. Machine learning for multi-omics data integration in cancer. *iScience.* 2022;25(2):103798.
233. Statnikov A, Aliferis CF, Tsamardinos I, Hardin D, Levy S. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. *Bioinformatics.* 2005 Mar 1;21(5):631–43.
234. Statnikov A, Wang L, Aliferis CF. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinformatics.* 2008 Jul 22;9(1):1–10.
235. Huang S, Nianguang CAI, Penzuti Pacheco P, Narandes S, Wang Y, Wayne XU. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. *Cancer Genomics Proteomics.* 2018 Jan 1;15(1):41–51.
236. Zheng D, Chen Y, Gao C, Wei Y, Cao G, Lu N, et al. Polymorphisms of p53 and MDM2 genes are associated with severe toxicities in patients with non-small cell lung cancer. *Cancer Biol Ther.* 2014 Nov 1;15(11):1542–51.
237. Liu W, Liu L, Wang R, Gong G, Ding X, Yang B, et al. Bevacizumab Combined With Oxaliplatin/Capecitabine in Patient With Refractory and Recurrent Mucinous Adenocarcinoma of the Appendix: A Case Report. *Front Oncol.* 2019;9(FEB).
238. Acuna-Hidalgo R, Sengul H, Steehouwer M, van de Vorst M, Vermeulen SH, Kiemeny LALM, et al. Ultra-sensitive Sequencing Identifies High Prevalence of Clonal Hematopoiesis-Associated Mutations throughout Adult Life. *Am J Hum Genet.* 2017 Jul 6;101(1):50–64.
239. Kashima J, Kitadai R, Okuma Y. Molecular and morphological profiling of lung cancer: A foundation for “next-generation” pathologists and oncologists. Vol. 11, *Cancers.* MDPI AG; 2019.
240. Huang JY, Cui SY, Chen YT, Song HZ, Huang GC, Feng B, et al. MicroRNA-650 Was a Prognostic Factor in Human Lung Adenocarcinoma and Confers the Docetaxel Chemoresistance of Lung Adenocarcinoma Cells via Regulating Bcl-2/Bax Expression. *PLoS One.* 2013;8(8):1–16.
241. Esquela-Kerscher A, Slack FJ. Oncomirs — microRNAs with a role in cancer. *Nat Rev Cancer* 2006 64. 2006 Apr;6(4):259–69.
242. Peng Y, Croce CM. The role of microRNAs in human cancer. Vol. 1, *Signal Transduction and Targeted Therapy.* Springer Nature; 2016.
243. Dhawan A, Scott JG, Harris AL, Buffa FM. Pan-cancer characterisation of microRNA across cancer hallmarks reveals microRNA-mediated downregulation of tumour suppressors. *Nat Commun.* 2018 Dec 1;9(1):1–13.
244. Iorio M V, Croce CM. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol Med.* 2012/02/20. 2012 Mar;4(3):143–59.
245. Iorio M V, Croce CM. Causes and consequences of microRNA dysregulation. *Cancer J.* 2012;18(3):215–22.
246. Di Leva G, Garofalo M, Croce CM. MicroRNAs in cancer. *Annu Rev Pathol.* 2013/09/25. 2014;9:287–314.
247. Kim Y-K, Kim VN. Processing of intronic microRNAs. *EMBO J.* 2007/01/25. 2007 Feb 7;26(3):775–83.
248. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 2004/09/13. 2004 Oct;14(10A):1902–10.
249. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell.* 2006;9(3):189–98.
250. Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, et al. A

- Polycistronic MicroRNA Cluster, *miR-17-92*, Is Overexpressed in Human Lung Cancers and Enhances Cell Proliferation. *Cancer Res.* 2005;65(21):9628–32.
251. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer.* 2006;6(11):857–66.
 252. Diederichs S, Haber DA. Sequence Variations of MicroRNAs in Human Cancer: Alterations in Predicted Secondary Structure Do Not Affect Processing. *Cancer Res.* 2006;66(12):6097–104.
 253. Czubak K, Lewandowska MA, Klonowska K, Roszkowski K, Kowalewski J, Figlerowicz M, et al. High copy number variation of cancer-related microRNA genes and frequent amplification of DICER1 and DROSHA in lung cancer. *Oncotarget.* 2015 Sep 15;6(27):23399–416.
 254. Liu T, Zhang H, Fang J, Yang Z, Chen R, Wang Y, et al. AGO2 phosphorylation by c-Src kinase promotes tumorigenesis. *Neoplasia.* 2020/01/22. 2020 Mar;22(3):129–41.
 255. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature.* 2004;429(6990):457–63.
 256. Lodygin D, Tarasov V, Epanchintsev A, Berking C, Knyazeva T, Körner H, et al. Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. *Cell Cycle.* 2008;7(16):2591–600.
 257. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setién F, et al. Genetic Unmasking of an Epigenetically Silenced microRNA in Human Cancer Cells. *Cancer Res.* 2007;67(4):1424–9.
 258. Qin A, Reddy HG, Weinberg FD, Kalemkerian GP. Cyclin-dependent kinase inhibitors for the treatment of lung cancer. *Expert Opin Pharmacother.* 2020;21(8):941–52.
 259. Enkhbaatar Z, Terashima M, Oktyabri D, Tange S, Ishimura A, Yano S, et al. KDM5B histone demethylase controls epithelial-mesenchymal transition of cancer cells by regulating the expression of the microRNA-200 family. *Cell Cycle.* 2013/06/06. 2013 Jul 1;12(13):2100–12.
 260. Ceppi P, Mudduluru G, Kumarswamy R, Rapa I, Scagliotti G V, Papotti M, et al. Loss of miR-200c Expression Induces an Aggressive, Invasive, and Chemoresistant Phenotype in Non-Small Cell Lung Cancer. *Mol Cancer Res.* 2010;8(9):1207–16.
 261. Liang G, Meng W, Huang X, Zhu W, Yin C, Wang C, et al. MiR-196b-5p-mediated downregulation of TSPAN12 and GATA6 promotes tumor progression in non-small cell lung cancer. *Proc Natl Acad Sci U S A.* 2020 Feb 25;117(8):4347–57.
 262. Zhong S, Golpon H, Zardo P, Borlak J. miRNAs in lung cancer. A systematic review identifies predictive and prognostic miRNA candidates for precision medicine in lung cancer. *Transl Res.* 2021;230:164–96.
 263. Kok MGM, de Ronde MWJ, Moerland PD, Ruijter JM, Creemers EE, Pinto-Sietsma SJ. Small sample sizes in high-throughput miRNA screens: A common pitfall for the identification of miRNA biomarkers. *Biomol Detect Quantif.* 2018 May 1;15:1–5.
 264. Zhao T, Khadka VS, Deng Y. Identification of lncRNA biomarkers for lung cancer through integrative cross-platform data analyses. *Aging (Albany NY).* 2020 Jul 31;12(14):14506–27.
 265. Zhang Y-K, Zhu W-Y, He J-Y, Chen D-D, Huang Y-Y, Le H-B, et al. miRNAs expression profiling to distinguish lung squamous-cell carcinoma from adenocarcinoma subtypes. *J Cancer Res Clin Oncol.* 2012;138(10):1641–50.
 266. HAMAMOTO J, SOEJIMA K, YODA S, NAOKI K, NAKAYAMA S, SATOMI R, et al. Identification of microRNAs differentially expressed between lung squamous cell carcinoma and lung adenocarcinoma. *Mol Med Rep.* 2013;8(2):456–62.
 267. Hu Z, Wang X, Yang Y, Zhao Y, Shen Z, Huang Y. MicroRNA expression profiling of lung adenocarcinoma in Xuanwei, China: A preliminary study. *Medicine (Baltimore).* 2019 May;98(21):e15717–e15717.
 268. Hu Y, Wang L, Gu J, Qu K, Wang Y. Identification of microRNA differentially expressed in three subtypes of non-small cell lung cancer and in silico functional analysis. *Oncotarget.* 2017;8(43):74554–66.
 269. Sherafatian M, Arjmand F. Decision tree-based classifiers for lung cancer diagnosis and subtyping using TCGA miRNA expression data. *Oncol Lett.* 2019 Aug 1;18(2):2125–31.
 270. Gallach S, Jantus-Lewintre E, Calabuig-Fariñas S, Montaner D, Alonso S, Sirera R, et al. MicroRNA profiling associated with non-small cell lung cancer: Next generation sequencing

- detection, experimental validation, and prognostic value. *Oncotarget*. 2017;8(34):56143–57.
271. Wu K-L, Tsai Y-M, Lien C-T, Kuo P-L, Hung, Jen-Yu. The Roles of MicroRNA in Lung Cancer. *Int J Mol Sci*. 2019;20(7):1611.
 272. Zhu X, Kudo M, Huang X, Sui H, Tian H, Croce CM, et al. Frontiers of MicroRNA Signature in Non-small Cell Lung Cancer. *Front cell Dev Biol*. 2021 Apr 7;9:643942.
 273. Yu H, Guan Z, Cuk K, Zhang Y, Brenner H. Circulating MicroRNA Biomarkers for Lung Cancer Detection in East Asian Populations. *Cancers (Basel)*. 2019 Mar 23;11(3):415.
 274. Niu Y, Su M, Wu Y, Fu L, Kang K, Li Q, et al. Circulating Plasma miRNAs as Potential Biomarkers of Non-Small Cell Lung Cancer Obtained by High-Throughput Real-Time PCR Profiling. *Cancer Epidemiol Biomarkers & Prev*. 2019;28(2):327–36.
 275. Mestdagh P, Hartmann N, Baeriswyl L, Andreasen D, Bernard N, Chen C, et al. Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nat Methods*. 2014;11(8):809–15.
 276. Wu KL, Tsai YM, Lien CT, Kuo PL, Hung JY. The roles of microRNA in lung cancer. Vol. 20, *International Journal of Molecular Sciences*. MDPI AG; 2019.
 277. Ahn YH, Ko YH. Diagnostic and therapeutic implications of micrnas in non-small cell lung cancer. Vol. 21, *International Journal of Molecular Sciences*. MDPI AG; 2020. p. 1–17.
 278. Naeli P, Yousefi F, Ghasemi Y, Savardashtaki A, Mirzaei H. The Role of MicroRNAs in Lung Cancer: Implications for Diagnosis and Therapy. *Curr Mol Med*. 2020;20(2):90–101.
 279. Hamamoto J, Soejima K, Yoda S, Naoki K, Nakayama S, Satomi R, et al. Identification of microRNAs differentially expressed between lung squamous cell carcinoma and lung adenocarcinoma. *Mol Med Rep*. 2013 Aug 1;8(2):456–62.
 280. Griffiths-Jones S, Hui JHL, Marco A, Ronshaugen M. MicroRNA evolution by arm switching. *EMBO Rep*. 2011/01/07. 2011 Feb;12(2):172–7.
 281. Shen J, Xia W, Khotskaya YB, Huo L, Nakanishi K, Lim S-O, et al. EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature*. 2013/05/01. 2013 May 16;497(7449):383–7.
 282. Ro S, Park C, Young D, Sanders KM, Yan W. Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res*. 2007/08/28. 2007;35(17):5944–53.
 283. Jagadeeswaran G, Zheng Y, Sumathipala N, Jiang H, Arrese EL, Soulages JL, et al. Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. *BMC Genomics*. 2010 Jan 20;11:52.
 284. Kuchenbauer F, Mah SM, Heuser M, McPherson A, Rüschemann J, Rouhi A, et al. Comprehensive analysis of mammalian miRNA* species and their role in myeloid cells. *Blood*. 2011;118(12):3350–8.
 285. Yang J-S, Phillips MD, Betel D, Mu P, Ventura A, Siepel AC, et al. Widespread regulatory activity of vertebrate microRNA* species. *RNA*. 2010/12/22. 2011 Feb;17(2):312–26.
 286. Medley JC, Panzade G, Zinovyeva AY. microRNA strand selection: Unwinding the rules. *Wiley Interdiscip Rev RNA*. 2021;12(3):1–22.
 287. Johnstone IM, Lu AY. On Consistency and Sparsity for Principal Components Analysis in High Dimensions. *J Am Stat Assoc*. 2009 Jun 1;104(486):682–93.
 288. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116(2):281–97.
 289. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res*. 1996;6(10):986–94.
 290. Hunt EA, Broyles D, Head T, Deo SK. MicroRNA Detection: Current Technology and Research Strategies. *Annu Rev Anal Chem*. 2015;8(1):217–37.
 291. Wang B, Xi Y. Challenges for MicroRNA Microarray Data Analysis. *Microarrays (Basel, Switzerland)*. 2013 Jun;2(2):34–50.
 292. Jaksik R, Iwanaszko M, Rzeszowska-Wolny J, Kimmel M. Microarray experiments and factors which affect their reliability. *Biol Direct*. 2015 Sep 3;10:46.
 293. Benes V, Castoldi M. Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods*. 2010;50(4):244–9.
 294. Meyer SU, Pfaffl MW, Ulbrich SE. Normalization strategies for microRNA profiling

- experiments: a ‘normal’ way to a hidden layer of complexity? *Biotechnol Lett.* 2010;32(12):1777–88.
295. Becker C, Riedmaier I, Pfaffl M. Biomarker Discovery via RT-qPCR and Bioinformatical Validation. *PCR Technol.* 2013;259–70.
 296. Chang KH, Mestdagh P, Vandesompele J, Kerin MJ, Miller N. MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer.* 2010 Apr 29;10:173.
 297. Peltier HJ, Latham GJ. Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA.* 2008/03/28. 2008 May;14(5):844–52.
 298. Han S, Semba S, Abe T, Makino N, Furukawa T, Fukushima S, et al. Infrequent somatic mutations of the p73 gene in various human cancers. *Eur J Surg Oncol.* 1999;25(2):194–8.
 299. Testa JR, Siegfried JM, Liu Z, Hunt JD, Feder MM, Litwin S, et al. Cytogenetic analysis of 63 non-small cell lung carcinomas: Recurrent chromosome alterations amid frequent and widespread genomic upheaval. *Genes, Chromosom Cancer.* 1994;11(3):178–94.
 300. Nomoto S, Haruki N, Tatematsu Y, Konishi H, Mitsudomi T, Takahashi T, et al. Frequent allelic imbalance suggests involvement of a tumor suppressor gene at 1p36 in the pathogenesis of human lung cancers. *Genes, Chromosom Cancer.* 2000;28(3):342–6.
 301. Sunpaweravong P, Thu KL, Lam WL, Mai S. Assessment of the clinical relevance of 17q25.3 copy number and three-dimensional telomere organization in non-small lung cancer patients. *J Cancer Res Clin Oncol.* 2015;142(4):749–56.
 302. Yu C, Chen D, Liu H, Li W, Lu J, Feng J. Rosmarinic acid reduces the resistance of gastric carcinoma cells to 5-fluorouracil by downregulating FOXO4-targeting miR-6785-5p. *Biomed & Pharmacother.* 2019;109:2327–34.
 303. Guo X, Chen M, Cao L, Hu Y, Li X, Zhang Q, et al. Cancer-Associated Fibroblasts Promote Migration and Invasion of Non-Small Cell Lung Cancer Cells via miR-101-3p Mediated VEGFA Secretion and AKT/eNOS Pathway. *Front cell Dev Biol.* 2021 Dec 16;9:764151.
 304. Shen Q, Bae HJ, Eun JW, Kim HS, Park SJ, Shin WC, et al. MiR-101 functions as a tumor suppressor by directly targeting nemo-like kinase in liver cancer. *Cancer Lett.* 2014;344(2):204–11.
 305. Xiaoping L, Zhibin Y, Wenjuan L, Zeyou W, Gang X, Zhaohui L, et al. CPEB1, a histone-modified hypomethylated gene, is regulated by miR-101 and involved in cell senescence in glioma. *Cell Death Dis.* 2013 Jun 20;4(6):e675–e675.
 306. Yin J, Wang M, Jin C, Qi Q. miR-101 sensitizes A549 NSCLC cell line to CDDP by activating caspase 3-dependent apoptosis. *Oncol Lett.* 2013/12/03. 2014 Feb;7(2):461–5.
 307. Zhang X, He X, Liu Y, Zhang H, Chen H, Guo S, et al. MiR-101-3p inhibits the growth and metastasis of non-small cell lung cancer through blocking PI3K/AKT signal pathway by targeting MALAT-1. *Biomed & Pharmacother.* 2017;93:1065–73.
 308. Wu B, Lei D, Wang L, Yang X, Jia S, Yang Z, et al. MiRNA-101 inhibits oral squamous-cell carcinoma growth and metastasis by targeting zinc finger E-box binding homeobox 1. *Am J Cancer Res.* 2016;6(6):1396–407.
 309. Li J-T, Jia L-T, Liu N-N, Zhu X-S, Liu Q-Q, Wang X-L, et al. MiRNA-101 inhibits breast cancer growth and metastasis by targeting CX chemokine receptor 7. *Oncotarget.* 2015 Oct 13;6(31):30818–30.
 310. Lv P, Zhang P, Li X, Chen Y. Micro ribonucleic acid (RNA)-101 inhibits cell proliferation and invasion of lung cancer by regulating cyclooxygenase-2. *Thorac cancer.* 2015/06/23. 2015 Nov;6(6):778–84.
 311. Tang XR, Wen X, He QM, Li YQ, Ren XY, Yang XJ, et al. MicroRNA-101 inhibits invasion and angiogenesis through targeting ITGA3 and its systemic delivery inhibits lung metastasis in nasopharyngeal carcinoma. *Cell Death Dis.* 2017;8(1):e2566-9.
 312. Meng X, Sun Y, Liu S, Mu Y. miR-101-3p sensitizes lung adenocarcinoma cells to irradiation via targeting BIRC5. *Oncol Lett.* 2021/02/10. 2021 Apr;21(4):282.
 313. Tian T, Mingyi M, Qiu X, Qiu Y. MicroRNA-101 reverses temozolomide resistance by inhibition of GSK3 β in glioblastoma. *Oncotarget.* 2016 Nov 29;7(48):79584–95.
 314. Jiang R, Zhang C, Liu G, Gu R, Wu H. MicroRNA-101 inhibits proliferation, migration and

- invasion in osteosarcoma cells by targeting ROCK1. *Am J Cancer Res.* 2017;7(1):88.
315. Guo F, Guo L, Li Y, Zhou Q, Li Z. MALAT1 is an oncogenic long non-coding RNA associated with tumor invasion in non-small cell lung cancer regulated by DNA methylation. *Int J Clin Exp Pathol.* 2015;8(12):15903.
 316. Gutschner T, Hämmerle M, Eissmann M, Hsu J, Kim Y, Hung G, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* 2012/12/14. 2013 Feb 1;73(3):1180–9.
 317. Li P, Zhang X, Wang H, Wang L, Liu T, Du L, et al. MALAT1 Is Associated with Poor Response to Oxaliplatin-Based Chemotherapy in Colorectal Cancer Patients and Promotes Chemoresistance through EZH2. *Mol Cancer Ther.* 2017;16(4):739–51.
 318. Chen W, Zhao W, Zhang L, Wang L, Wang J, Wan Z, et al. MALAT1-miR-101-SOX9 feedback loop modulates the chemoresistance of lung cancer cell to DDP via Wnt signaling pathway. *Oncotarget.* 2017;8(55):94317–29.
 319. Santos JC, Carrasco-Garcia E, Garcia-Puga M, Aldaz P, Montes M, Fernandez-Reyes M, et al. SOX9 Elevation Acts with Canonical WNT Signaling to Drive Gastric Cancer Progression. *Cancer Res.* 2016;76(22):6735–46.
 320. Ma F, Ye H, He HH, Gerrin SJ, Chen S, Tanenbaum BA, et al. SOX9 drives WNT pathway activation in prostate cancer. *J Clin Invest.* 2016/04/04. 2016 May 2;126(5):1745–58.
 321. Liu H, Liu Z, Jiang B, Peng R, Ma Z, Lu J. SOX9 Overexpression Promotes Glioma Metastasis via Wnt/ β -Catenin Signaling. *Cell Biochem Biophys.* 2015;73(1):205–12.
 322. Wu K, Liang W-C, Feng L, Pang J, Waye MM-Y, Zhang J-F, et al. H19 mediates methotrexate resistance in colorectal cancer through activating Wnt/ β -catenin pathway. *Exp Cell Res.* 2017;350(2):312–7.
 323. Dong W, Yao C, Teng X, Chai J, Yang X, Li B. MiR-140-3p suppressed cell growth and invasion by downregulating the expression of ATP8A1 in non-small cell lung cancer. *Tumor Biol.* 2016;37(3):2973–85.
 324. Huang H, Wang Y, Li Q, Fei X, Ma H, Hu R. miR-140-3p functions as a tumor suppressor in squamous cell lung cancer by regulating BRD9. *Cancer Lett.* 2019;446:81–9.
 325. Li D, Xu T, Wang X, Ma X, Liu T, Wang Y, et al. The role of ATP8A1 in non-small cell lung cancer. *Int J Clin Exp Pathol.* 2017;10(7):7760–6.
 326. Song B, Wang Y, Xi Y, Kudo K, Bruheim S, Botchkina GI, et al. Mechanism of chemoresistance mediated by miR-140 in human osteosarcoma and colon cancer cells. *Oncogene.* 2009/09/07. 2009 Nov 19;28(46):4065–74.
 327. Wu S, Wang H, Pan Y, Yang X, Wu D. miR-140-3p enhances cisplatin sensitivity and attenuates stem cell-like properties through repressing Wnt/ β -catenin signaling in lung adenocarcinoma cells. *Exp Ther Med.* 2020;20(2):1664–74.
 328. Yang H, Fang F, Chang R, Yang L. MicroRNA-140-5p suppresses tumor growth and metastasis by targeting transforming growth factor β receptor 1 and fibroblast growth factor 9 in hepatocellular carcinoma. *Hepatology.* 2013;58(1):205–17.
 329. Kong XM, Zhang GH, Huo YK, Zhao XH, Cao DW, Guo SF, et al. MicroRNA-140-3p inhibits proliferation, migration and invasion of lung cancer cells by targeting ATP6AP2. *Int J Clin Exp Pathol.* 2015;8(10):12845–52.
 330. Jude JA, Dileepan M, Subramanian S, Solway J, Panettieri Jr RA, Walseth TF, et al. miR-140-3p regulation of TNF- α -induced CD38 expression in human airway smooth muscle cells. *Am J Physiol Lung Cell Mol Physiol.* 2012/07/06. 2012 Sep;303(5):L460–8.
 331. Zhong Y, Du Y, Yang X, Mo Y, Fan C, Xiong F, et al. Circular RNAs function as ceRNAs to regulate and control human cancer progression. *Mol Cancer.* 2018 Apr 7;17(1):79.
 332. Wang Y, Wo Y, Lu T, Sun X, Liu A, Dong Y, et al. Circ-AASDH functions as the progression of early stage lung adenocarcinoma by targeting miR-140-3p to activate E2F7 expression. *Transl Lung Cancer Res.* 2021 Jan 1;10(1):57–70.
 333. Uchida A, Seki N, Mizuno K, Misono S, Yamada Y, Kikkawa N, et al. Involvement of dual-strand of the miR-144 duplex and their targets in the pathogenesis of lung squamous cell carcinoma. *Cancer Sci.* 2019 Jan 1;110(1):420–32.
 334. Liu C, Yang Z, Deng Z, Zhou Y, Gong Q, Zhao R, et al. Downregulated miR-144-3p

- contributes to progression of lung adenocarcinoma through elevating the expression of EZH2. *Cancer Med.* 2018 Nov 1;7(11):5554–66.
335. Wen X, Han XR, Wang YJ, Fan SH, Zhuang J, Zhang ZF, et al. Effects of long noncoding RNA SPRY4-IT1-mediated EZH2 on the invasion and migration of lung adenocarcinoma. *J Cell Biochem.* 2018 Feb 1;119(2):1827–40.
 336. Jiang W, Xu Z, Yu L, Che J, Zhang J, Yang J. MicroRNA-144-3p suppressed TGF- β 1-induced lung cancer cell invasion and adhesion by regulating the Src–Akt–Erk pathway. *Cell Biol Int.* 2020 Jan 1;44(1):51–61.
 337. Wu C, Xu B, Zhou Y, Ji M, Zhang D, Jiang J, et al. Correlation between serum IL-1 β and miR-144-3p as well as their prognostic values in LUAD and LUSC patients. *Oncotarget.* 2016;7(52):85876–87.
 338. Yin Y, Liu H, Xu J, Shi D, Zhai L, Liu B, et al. MiR-144-3p regulates the resistance of lung cancer to cisplatin by targeting Nrf2. *Oncol Rep.* 2018 Dec 1;40(6):3479–88.
 339. Niture SK, Jaiswal AK. Nrf2 protein up-regulates antiapoptotic protein Bcl-2 and prevents cellular apoptosis. *J Biol Chem.* 2012/01/24. 2012 Mar 23;287(13):9873–86.
 340. Solis LM, Behrens C, Dong W, Suraokar M, Ozburn NC, Moran CA, et al. Nrf2 and Keap1 abnormalities in non-small cell lung carcinoma and association with clinicopathologic features. *Clin Cancer Res.* 2010/06/09. 2010 Jul 15;16(14):3743–53.
 341. Hayes JD, McMahon M, Chowdhry S, Dinkova-Kostova AT. Cancer Chemoprevention Mechanisms Mediated Through the Keap1–Nrf2 Pathway. *Antioxidants & Redox Signal.* 2010;13(11):1713–48.
 342. Guan H, Liu C, Fang F, Huang Y, Tao T, Ling Z, et al. MicroRNA-744 promotes prostate cancer progression through aberrantly activating Wnt/ β -catenin signaling. *Oncotarget.* 2017;8(9):14693–707.
 343. Zhang M, Li H, Zhang Y, Li H. Oncogenic miR-744 promotes prostate cancer growth through direct targeting of LKB1. *Oncol Lett.* 2019 Feb 1;17(2):2257–65.
 344. Lin F, Ding R, Zheng S, Xing D, Hong W, Zhou Z, et al. Decrease expression of microRNA-744 promotes cell proliferation by targeting c-Myc in human hepatocellular carcinoma. *Cancer Cell Int.* 2014 Jun 22;14(1):1–9.
 345. Deng Y, Li Y, Fang Q, Luo H, Zhu G. microRNA-744 is downregulated in glioblastoma and inhibits the aggressive behaviors by directly targeting NOB1. *Am J Cancer Res.* 2018;8(11):2238–53.
 346. Kleemann M, Schneider H, Unger K, Sander P, Schneider EM, Fischer-Posovszky P, et al. MiR-744-5p inducing cell death by directly targeting HNRNPC and NFIX in ovarian cancer cells. *Sci Rep.* 2018 Dec 1;8(1):1–15.
 347. Shen J, Li M. MicroRNA-744 inhibits cellular proliferation and invasion of colorectal cancer by directly targeting oncogene Notch1. *Oncol Res.* 2018;26(9):1401–9.
 348. Chen S, Shi F, Zhang W, Zhou Y, Huang J. Mir-744-5p inhibits non-small cell lung cancer proliferation and invasion by directly targeting pax2. *Technol Cancer Res Treat.* 2019 Sep 15;18:1–7.
 349. Wang J, Cai H, Dai Z, Wang G. Down-regulation of lncRNA XIST inhibits cell proliferation via regulating miR-744/RING1 axis in non-small cell lung cancer. *Clin Sci.* 2019 Jul 31;133(14):1567–79.
 350. Sui Y, Lin G, Zheng Y, Huang W. LncRNA MAFG-AS1 boosts the proliferation of lung adenocarcinoma cells via regulating miR-744-5p/MAFG axis. *Eur J Pharmacol.* 2019 Sep 15;859.
 351. Zhu X. MiR-744 facilitates non-small cell lung cancer progression by direct transcriptional regulation and indirect MAPK pathway activation mediated upregulation of C-FOS. :1–24.
 352. Du M, Zhuang Y, Tan P, Yu Z, Zhang X, Wang A. microRNA-95 knockdown inhibits epithelial–mesenchymal transition and cancer stem cell phenotype in gastric cancer cells through MAPK pathway by upregulating DUSP5. *J Cell Physiol.* 2019;235(2):944–56.
 353. Geng Y, Zhao S, Jia Y, Xia G, Li H, Fang Z, et al. miR-95 promotes osteosarcoma growth by targeting SCNN1A. *Oncol Rep.* 2020/02/24. 2020 May;43(5):1429–36.
 354. Xi M, Cheng L, Hua W, Zhou YL, Gao QL, Yang JX, et al. MicroRNA-95-3p promoted the development of prostatic cancer via regulating DKK3 and activating Wnt/ β -catenin pathway.

- Eur Rev Med Pharmacol Sci. 2019;23(3):1002–11.
355. Huang Z, Huang S, Wang Q, Liang L, Ni S, Wang L, et al. MicroRNA-95 Promotes Cell Proliferation and Targets Sorting Nexin 1 in Human Colorectal Carcinoma. *Cancer Res.* 2011;71(7):2582–9.
 356. FAN BO, JIAO B-H, FAN F-S, LU S-K, SONG J, GUO C-Y, et al. Downregulation of miR-95-3p inhibits proliferation, and invasion promoting apoptosis of glioma cells by targeting CELF2. *Int J Oncol.* 2015;47(3):1025–33.
 357. Seki N. Identification of novel molecular targets regulated by tumor suppressive miR-1/miR-133a in maxillary sinus squamous cell carcinoma. *Int J Oncol.* 2011;
 358. Wu C-D, Kuo Y-S, Wu H-C, Lin C-T. MicroRNA-1 induces apoptosis by targeting prothymosin alpha in nasopharyngeal carcinoma cells. *J Biomed Sci.* 2011 Nov 7;18(1):80.
 359. Hwang SJ, Lee HW, Kim HR, Song HJ, Lee DH, Lee H, et al. Overexpression of microRNA-95-3p suppresses brain metastasis of lung adenocarcinoma through downregulation of cyclin D1. *Oncotarget.* 2015;6(24):20434–48.
 360. Huang X, Taeb S, Jahangiri S, Emmenegger U, Tran E, Bruce J, et al. miRNA-95 Mediates Radioresistance in Tumors by Targeting the Sphingolipid Phosphatase SGPP1. *Cancer Res.* 2013;73(23):6972–86.
 361. Chen X, Chen S, Hang W, Huang H, Ma H. MiR-95 induces proliferation and chemo- or radioresistance through directly targeting sorting nexin1 (SNX1) in non-small cell lung cancer. *Biomed & Pharmacother.* 2014;68(5):589–95.
 362. Takuwa Y, Okamoto Y, Yoshioka K, Takuwa N. Sphingosine-1-phosphate signaling in physiology and diseases. *BioFactors.* 2012;38(5):329–37.
 363. Das S. Evolutionary origin and genomic organization of micro-RNA genes in immunoglobulin lambda variable region gene family. *Mol Biol Evol.* 2009/02/26. 2009 May;26(5):1179–89.
 364. Zhang XL, Zhu WY, Zhang JF, Huo SJ, Zhou LM, Gu Z, et al. MicroRNA-650 targets ING4 to promote gastric cancer tumorigenicity. *Biochem Biophys Res Commun.* 2010 Apr 30;395(2):275–80.
 365. Sun B, Pu B, Chu D, Chu X, Li W, Wei D. MicroRNA-650 expression in glioma is associated with prognosis of patients. *J Neurooncol.* 2013 Dec 24;115(3):375–80.
 366. Zeng ZL, Li FJ, Gao F, Sun DS, Yao L. Upregulation of miR-650 is correlated with down-regulation of ING4 and progression of hepatocellular carcinoma. *J Surg Oncol.* 2013 Feb 1;107(2):105–10.
 367. Tang X, Ding Y, Wang X, Wang X, Zhao L, Bi H. MiR-650 promotes non-small cell lung cancer cell proliferation and invasion by targeting ING4 through Wnt-1/ β -catenin pathway. *Oncol Lett.* 2019 Nov 1;18(5):4621–8.
 368. Korać P, Antica M, Matulić M. Mir-7 in cancer development. Vol. 9, *Biomedicines.* MDPI AG; 2021.
 369. Gajda E, Grzanka M, Godlewska M, Gawel D. The role of miRNA-7 in the biology of cancer and modulation of drug resistance. Vol. 14, *Pharmaceuticals.* MDPI AG; 2021. p. 1–24.
 370. Chou YT, Lin HH, Lien YC, Wang YH, Hong CF, Kao YR, et al. EGFR promotes lung tumorigenesis by activating miR-7 through a Ras/ERK/Myc pathway that targets the Ets2 transcriptional repressor ERF. *Cancer Res.* 2010 Nov 1;70(21):8822–31.
 371. Barshack I, Lithwick-Yanai G, Afek A, Rosenblatt K, Tabibian-Keissar H, Zepeniuk M, et al. MicroRNA expression differentiates between primary lung tumors and metastases to the lung. *Pathol Res Pract.* 2010;206(8):578–84.
 372. Petriella D, Galetta D, Rubini V, Savino E, Paradiso A, Simone G, et al. Molecular profiling of thin-prep FNA samples in assisting clinical management of non-small-cell lung cancer. *Mol Biotechnol.* 2013;54(3):913–9.
 373. Meza-Sosa KF, Pérez-García EI, Camacho-Concha N, López-Gutiérrez O, Pedraza-Alva G, Pérez-Martínez L. MiR-7 promotes epithelial cell transformation by targeting the tumor suppressor KLF4. *PLoS One.* 2014;9(9).
 374. Dragomir M, Mafra ACP, Dias SMG, Vasilescu C, Calin GA. Using microRNA Networks to Understand Cancer. *Int J Mol Sci.* 2018 Jul 1;19(7).
 375. Hong CF, Lin SY, Chou YT, Wu CW. MicroRNA-7 compromises p53 protein-dependent apoptosis by controlling the expression of the chromatin remodeling factor SMARCD1. *J Biol*

- Chem. 2016 Jan 22;291(4):1877–89.
376. Ray Chaudhuri A, Nussenzweig A. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nat Rev Mol Cell Biol.* 2017/07/05. 2017 Oct;18(10):610–21.
 377. Teloni F, Altmeyer M. Readers of poly(ADP-ribose): designed to be fit for purpose. *Nucleic Acids Res.* 2015/12/15. 2016 Feb 18;44(3):993–1006.
 378. Moynahan ME, Jasin M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol.* 2010 Mar;11(3):196–207.
 379. Lai J, Yang H, Zhu Y, Ruan M, Huang Y, Zhang Q. MiR-7-5p-mediated downregulation of PARP1 impacts DNA homologous recombination repair and resistance to doxorubicin in small cell lung cancer. *BMC Cancer.* 2019;19(1):1–9.
 380. Cheng MW, Shen ZT, Hu GY, Luo LG. Prognostic Significance of microRNA-7 and its Roles in the Regulation of Cisplatin Resistance in Lung Adenocarcinoma. *Cell Physiol Biochem.* 2017 Jul 1;42(2):660–72.
 381. Su T, Huang S, Zhang Y, Guo Y, Zhang S, Guan J, et al. miR-7/TGF- β 2 axis sustains acidic tumor microenvironment-induced lung cancer metastasis. *Acta Pharm Sin B.* 2022;12(2):821–37.
 382. Ouyang M, Li Y, Ye S, Ma J, Lu L, Lv W, et al. MicroRNA profiling implies new markers of chemoresistance of triple-negative breast cancer. *PLoS One.* 2014 May 2;9(5).
 383. Han Y, Xu GX, Lu H, Yu DH, Ren Y, Wang L, et al. Dysregulation of miRNA-21 and their potential as biomarkers for the diagnosis of cervical cancer. *Int J Clin Exp Pathol.* 2015 Jun 1;8(6):7131–9.
 384. Cybula M, Wieteska L, Józefowicz-Korczyńska M, Karbownik MS, Grzelczyk WL, Szemraj J. New miRNA expression abnormalities in laryngeal squamous cell carcinoma. *Cancer Biomarkers.* 2016 Apr 13;16(4):559–68.
 385. Wang Y, Li J, Tong L, Zhang J, Zhai A, Xu K, et al. The prognostic value of miR-21 and miR-155 in non-small-cell lung cancer: A meta-analysis. *Jpn J Clin Oncol.* 2013 Aug;43(8):813–20.
 386. Zheng YY, Fei Y, Wang Z, Chen Y, Qiu C, Li FR. Tissue microRNAs in non-small cell lung cancer detected with a new kind of liquid bead array detection system. *J Transl Med.* 2020 Mar 2;18(1):108.
 387. Jiang C, Guo Y, Yu H, Lu S, Meng L. Pleiotropic microRNA-21 in pulmonary remodeling: Novel insights for molecular mechanism and present advancements. *Allergy, Asthma Clin Immunol.* 2019 May 20;15(1):1–9.
 388. Báez-Vega PM, Vargas IME, Valiyeva F, Rosado JE, Roman A, Flores J, et al. Targeting miR-21-3p inhibits proliferation and invasion of ovarian cancer cells. *Oncotarget.* 2016;7(24):36321–37.
 389. Ferragut Cardoso AP, Banerjee M, Nail AN, Lykoudi A, States JC. miRNA dysregulation is an emerging modulator of genomic instability. *Semin Cancer Biol.* 2021;
 390. Yan J-W, Lin J-S, He X-X. The emerging role of miR-375 in cancer. *Int J Cancer.* 2013;135(5):1011–8.
 391. Jin Y, Liu Y, Zhang J, Huang W, Jiang H, Hou Y, et al. The expression of MIR-375 is associated with carcinogenesis in three subtypes of lung cancer. *PLoS One.* 2015 Dec 1;10(12):144187.
 392. Bhaumik S, Ahmad F, Das BR. Molecular binary classification of NSCLC: miR-375 is a potential biomarker to differentiate SQCC from ADCC in Indian NSCLC patients. *Appl Cancer Res.* 2017 Dec 25;37(1):1–10.
 393. Xu N, Qiao L, Yin L, Li H. Long noncoding RNA ROR1-AS1 enhances lung adenocarcinoma metastasis and induces epithelial-mesenchymal transition by sponging miR-375. *JBUON.* 2019;24(6):2273–9.
 394. Zehentmayr F, Hauser-Kronberger C, Zellinger B, Hlubek F, Schuster C, Bodenhofer U, et al. Hsa-miR-375 is a predictor of local control in early stage breast cancer. *Clin Epigenetics.* 2016 Mar 8;8:28.
 395. Porzycki P, Ciszakowicz E, Semik M, Tyrka M. Combination of three miRNA (miR-141, miR-21, and miR-375) as potential diagnostic tool for prostate cancer recognition. *Int Urol Nephrol.* 2018/07/16. 2018 Sep;50(9):1619–26.
 396. Wu Q, Zhao Y, Sun Y, Yan X, Wang P. miR-375 inhibits IFN- γ -induced programmed death 1

- ligand $\frac{1}{2}$ surface expression in head and neck squamous cell carcinoma cells by blocking JAK2/STAT1 signaling. *Oncol Rep.* 2018;
397. Lynn FC. Meta-regulation: microRNA regulation of glucose and lipid metabolism. *Trends Endocrinol Metab.* 2009 Nov;20(9):452–9.
 398. Li LM, Hu Z Bin, Zhou ZX, Chen X, Liu FY, Zhang JF, et al. Serum microRNA profiles serve as novel biomarkers for HBV infection and diagnosis of HBV-positive hepatocarcinoma. *Cancer Res.* 2010 Dec 1;70(23):9798–807.
 399. Lu TX, Lim EJ, Wen T, Plassard AJ, Hogan SP, Martin LJ, et al. MiR-375 is down-regulated in epithelial cells after IL-13 stimulation and regulates an IL-13 induced epithelial transcriptome. *Mucosal Immunol.* 2012 Jul;5(4):388.
 400. Erener S, Mojibian M, Fox JK, Denroche HC, Kieffer TJ. Circulating miR-375 as a biomarker of β -cell death and diabetes in mice. *Endocrinology.* 2013 Feb 1;154(2):603–8.
 401. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, MacDonald PE, et al. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature.* 2004 Nov 11;432(7014):226–30.
 402. Tang X, Tang G, Özcan S. Role of microRNAs in diabetes. *Biochim Biophys Acta.* 2008 Nov;1779(11):697–701.
 403. Zhang B, Li Y, Hou D, Shi Q, Yang S, Li Q. MicroRNA-375 Inhibits Growth and Enhances Radiosensitivity in Oral Squamous Cell Carcinoma by Targeting Insulin Like Growth Factor 1 Receptor. *Cell Physiol Biochem.* 2017 Oct 1;42(5):2105–17.
 404. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA.* 2005 Mar;11(3):241–7.
 405. Watanabe K, Amano Y, Ishikawa R, Sunohara M, Kage H, Ichinose J, et al. Histone methylation-mediated silencing of miR-139 enhances invasion of non-small-cell lung cancer. *Cancer Med.* 2015/08/08. 2015 Oct;4(10):1573–82.
 406. Zhang M, Lin B, Liu Y, Huang T, Chen M, Lian D, et al. LINC00324 affects non-small cell lung cancer cell proliferation and invasion through regulation of the miR-139-5p/IGF1R axis. *Mol Cell Biochem.* 2020 Oct 1;473(1–2):193–202.
 407. Huang LL, Huang LW, Wang L, Tong BD, Wei Q, Ding XS. Potential role of mir-139-5p in cancer diagnosis, prognosis and therapy (Review). *Oncol Lett.* 2017;14(2):1215–22.
 408. Li J, He X, Wu X, Liu X, Huang Y, Gong Y, et al. MiR-139-5p Inhibits Lung Adenocarcinoma Cell Proliferation, Migration, and Invasion by Targeting MAD2L1. *Comput Math Methods Med.* 2020;2020.
 409. Sun C, Sang M, Li S, Sun X, Yang C, Xi Y, et al. Hsa-miR-139-5p inhibits proliferation and causes apoptosis associated with down-regulation of c-Met. *Oncotarget.* 2015;6(37):39756–92.
 410. Luo HN, Wang ZH, Sheng Y, Zhang Q, Yan J, Hou J, et al. MiR-139 targets CXCR4 and inhibits the proliferation and metastasis of laryngeal squamous carcinoma cells. *Med Oncol.* 2014 Jan 1;31(1).
 411. Sun C, Sang M, Li S, Sun X, Yang C, Xi Y, et al. Hsa-miR-139-5p inhibits proliferation and causes apoptosis associated with down-regulation of c-Met. *Oncotarget.* 2015 Nov 24;6(37):39756–92.
 412. DU H, BAO Y, LIU C, ZHONG A, NIU Y, TANG X. miR-139-5p enhances cisplatin sensitivity in non-small cell lung cancer cells by inhibiting cell proliferation and promoting apoptosis via the targeting of Homeobox protein Hox-B2. *Mol Med Rep.* 2021 Feb 1;23(2):1–1.
 413. Bach DH, Luu TTT, Kim D, An YJ, Park S, Park HJ, et al. BMP4 Upregulation Is Associated with Acquired Drug Resistance and Fatty Acid Metabolism in EGFR-Mutant Non-Small-Cell Lung Cancer Cells. *Mol Ther - Nucleic Acids.* 2018 Sep 7;12:817–28.
 414. Finnerty JR, Wang W-X, Hébert SS, Wilfred BR, Mao G, Nelson PT. The miR-15/107 group of microRNA genes: evolutionary biology, cellular functions, and roles in human diseases. *J Mol Biol.* 2010/08/01. 2010 Sep 24;402(3):491–509.
 415. Li D, Zhao Y, Liu C, Chen X, Qi Y, Jiang Y, et al. Analysis of MiR-195 and MiR-497 Expression, Regulation and Role in Breast Cancer. *Clin Cancer Res.* 2011;17(7):1722–30.
 416. Volinia S, Calin GA, Liu C-G, Ambs S, Cimmino A, Petrocca F, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A.* 2006/02/03. 2006 Feb 14;103(7):2257–61.

417. Deng H, Guo Y, Song H, Xiao B, Sun W, Liu Z, et al. MicroRNA-195 and microRNA-378 mediate tumor growth suppression by epigenetical regulation in gastric cancer. *Gene*. 2013;518(2):351–9.
418. Itesako T, Seki N, Yoshino H, Chiyomaru T, Yamasaki T, Hidaka H, et al. The microRNA expression signature of bladder cancer by deep sequencing: the functional significance of the miR-195/497 cluster. *PLoS One*. 2014 Feb 10;9(2):e84311–e84311.
419. Yu W, Liang X, Li X, Zhang Y, Sun Z, Liu Y, et al. MicroRNA-195: A review of its role in cancers. *Onco Targets Ther*. 2018;11:7109–23.
420. Jin L, Li X, Li Y, Zhang Z, He T, Hu J, et al. Identification of miR-195-3p as an oncogene in RCC. *Mol Med Rep*. 2017;15(4):1916–24.
421. Jin M, Wang L, Zheng T, Yu J, Sheng R, Zhu H. MiR-195-3p inhibits cell proliferation in cervical cancer by targeting BCDIN3D. *J Reprod Immunol*. 2021;143:103211.
422. Angell H, Galon J. From the immune contexture to the Immunoscore: the role of prognostic and predictive immune markers in cancer. *Curr Opin Immunol*. 2013;25(2):261–7.
423. Oliver AJ, Lau PKH, Unsworth AS, Loi S, Darcy PK, Kershaw MH, et al. Tissue-Dependent Tumor Microenvironments and Their Impact on Immunotherapy Responses. *Front Immunol*. 2018 Jan 31;9:70.
424. Lao Y, Li T, Xie X, Chen K, Li M, Huang L. MiR-195-3p is a Novel Prognostic Biomarker Associated with Immune Infiltrates of Lung Adenocarcinoma. *Int J Gen Med*. 2022 Jan 6;15:191–203.
425. Wang J, Yu XF, Ouyang N, Zhao S, Yao H, Guan X, et al. MicroRNA and mRNA Interaction Network Regulates the Malignant Transformation of Human Bronchial Epithelial Cells Induced by Cigarette Smoke. *Front Oncol*. 2019;9(October):1–15.
426. Li L, Feng T, Zhang W, Gao S, Wang R, Lv W, et al. MicroRNA Biomarker hsa-miR-195-5p for Detecting the Risk of Lung Cancer. *Int J Genomics*. 2020;2020.
427. Du W, Liu T, Zhang Y, Zeng Y, Zhu J, Tang H, et al. MiR-195-5p is a potential factor responsible for CPNE1 differential expression between subtypes of non-small cell lung cancer. *J Cancer*. 2020;11(9):2610–20.
428. Dong S, Wu C, Song C, Qi B, Liu L, Xu Y. Identification of Primary and Metastatic Lung Cancer-Related lncRNAs and Potential Targeted Drugs Based on ceRNA Network. *Front Oncol*. 2021 Feb 3;10:3356.
429. Guo H, Li W, Zheng T, Liu Z. miR-195 Targets HDGF to inhibit proliferation and invasion of NSCLC cells. *Tumor Biol*. 2014 Oct 9;35(9):8861–6.
430. Wang X, Wang Y, Lan H, Li J. miR-195 inhibits the growth and metastasis of NSCLC cells by targeting IGF1R. *Tumor Biol*. 2014 Oct 9;35(9):8765–70.
431. Liu B, Qu J, Xu F, Guo Y, Wang Y, Yu H, et al. MiR-195 suppresses non-small cell lung cancer by targeting CHEK1. *Oncotarget*. 2015;6(11):9445–56.
432. Hill M, Tran N. MicroRNAs Regulating MicroRNAs in Cancer. *Trends in Cancer*. 2018;4(7):465–8.
433. Hill M, Tran N. miRNA interplay: Mechanisms and consequences in cancer. *DMM Dis Model Mech*. 2021;14(4):1–24.
434. Sales G, Coppe A, Bisognin A, Biasiolo M, Bortoluzzi S, Romualdi C. MAGIA, a web-based tool for miRNA and Genes Integrated Analysis. *Nucleic Acids Res*. 2010/05/19. 2010 Jul;38(Web Server issue):W352–9.
435. Hsu JB-K, Chiu C-M, Hsu S-D, Huang W-Y, Chien C-H, Lee T-Y, et al. miRTar: an integrated system for identifying miRNA-target interactions in human. *BMC Bioinformatics*. 2011 Jul 26;12:300.
436. Romano P, Céol A, Dräger A, Fiannaca A, Giugno R, La Rosa M, et al. The 2017 Network Tools and Applications in Biology (NETTAB) workshop: aims, topics and outcomes. *BMC Bioinformatics*. 2019 Apr 18;20(Suppl 4):125.
437. Kutmon M, Kelder T, Mandaviya P, Evelo CTA, Coort SL. CyTargetLinker: a cytoscape app to integrate regulatory interactions in network analysis. *PLoS One*. 2013 Dec 5;8(12):e82160–e82160.
438. Politano G, Benso A, Savino A, Di Carlo S. ReNE: a cytoscape plugin for regulatory network enhancement. *PLoS One*. 2014 Dec 26;9(12):e115585–e115585.

439. Yuan X, Liu C, Yang P, He S, Liao Q, Kang S, et al. Clustered microRNAs' coordination in regulating protein-protein interaction network. *BMC Syst Biol.* 2009 Jun 26;3:65.
440. Xu J, Li C-X, Li Y-S, Lv J-Y, Ma Y, Shao T-T, et al. MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. *Nucleic Acids Res.* 2010/10/06. 2011 Feb;39(3):825–36.
441. Zhao X, Song H, Zuo Z, Zhu Y, Dong X, Lu X. Identification of miRNA-miRNA synergistic relationships in colorectal cancer. *Int J Biol Macromol.* 2013;55:98–103.
442. Boross G, Orosz K, Farkas IJ. Human microRNAs co-silence in well-separated groups and have different predicted essentialities. *Bioinformatics.* 2009;25(8):1063–9.
443. Geng X, Kong X, Chen Q, Su S, Hu Y. Analysis of miRNA Functional Synergistic Network in Breast Cancer. *Proceedings of the 6th International Conference on Bioinformatics and Biomedical Science.* ACM; 2017.
444. Sahu M, Mallick B. Deciphering synergistic regulatory networks of microRNAs in hESCs and fibroblasts. *Int J Biol Macromol.* 2018;113:1279–86.
445. Zhang J, Duy Le T, Liu L, He J, Li J. Identifying miRNA synergistic regulatory networks in heterogeneous human data via network motifs. *Mol Biosyst.* 2016;12(2):454–63.
446. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr LA, Kinzler KW. Cancer genome landscapes. *Science.* 2013 Mar 29;339(6127):1546–58.
447. Kandasamy K, Mohan SS, Raju R, Keerthikumar S, Kumar GSS, Venugopal AK, et al. NetPath: a public resource of curated signal transduction pathways. *Genome Biol.* 2010 Jan 12;11(1):R3–R3.
448. Wang J, Zhou F, Li F, Wang B, Hu Y, Li X. Autocrined leptin promotes proliferation of non-small cell lung cancer (NSCLC) via PI3K/AKT and p53 pathways. *Ann Transl Med.* 2021 Apr;9(7):568.
449. Masjedi A, Hajizadeh F, Beigi Dargani F, Beyzai B, Aksoun M, Hojjat-Farsangi M, et al. Oncostatin M: A mysterious cytokine in cancers. *Int Immunopharmacol.* 2021;90(July):107158.
450. Jones SA, Jenkins BJ. Recent insights into targeting the IL-6 cytokine family in inflammatory diseases and cancer. *Nat Rev Immunol.* 2018;18(12):773–89.
451. Lauber S, Wong S, Cutz J-C, Tanaka M, Barra N, Lhoták Š, et al. Novel function of Oncostatin M as a potent tumour-promoting agent in lung. *Int J Cancer.* 2014;136(4):831–43.
452. Tseng CH. Human insulin therapy is associated with an increased risk of lung cancer: A population-based retrospective cohort study. *Front Endocrinol (Lausanne).* 2019;10(JULY):443.
453. Vella V, Milluzzo A, Scalisi NM, Vigneri P, Sciacca L. Insulin Receptor Isoforms in Cancer. *Int J Mol Sci.* 2018 Nov 16;19(11):3615.
454. Mancarella C, Scotlandi K. 40 YEARS OF IGF1: IGF system in sarcomas: a crucial pathway with many unknowns to exploit for therapy. *J Mol Endocrinol.* 2018;61(1):T45–60.
455. Vella V, Milluzzo A, Scalisi NM, Vigneri P, Sciacca L. Insulin Receptor Isoforms in Cancer. *Int J Mol Sci* 2018, Vol 19, Page 3615. 2018 Nov 16;19(11):3615.
456. Metz HE, McGarry Houghton A. Insulin Receptor Substrate Regulation of Phosphoinositide 3-Kinase. *Clin Cancer Res.* 2011;17(2):206–11.
457. Vuori K, Ruoslahti E. Association of insulin receptor substrate-1 with integrins. *Science (80-).* 1994;266(5190):1576–8.
458. Senthil D, Ghosh Choudhury G, Bhandari BK, Kasinath BS. The type 2 vascular endothelial growth factor receptor recruits insulin receptor substrate-1 in its signalling pathway. *Biochem J.* 2002 Nov 15;368(Pt 1):49–56.
459. Gibson SL, Ma Z, Shaw LM. Divergent Roles for IRS-1 and IRS-2 in Breast Cancer Metastasis. *Cell Cycle.* 2007;6(6):631–7.
460. Chan BT-Y, Lee A V. Insulin receptor substrates (IRSs) and breast tumorigenesis. *J Mammary Gland Biol Neoplasia.* 2008/11/22. 2008 Dec;13(4):415–22.
461. Gorgisen G, Hapil FZ, Yilmaz O, Cetin Z, Pehlivanoglu S, Ozbudak IH, et al. Identification of novel mutations of Insulin Receptor Substrate 1(IRS1) in tumor samples of non-small cell lung cancer (NSCLC): Implications for aberrant insulin signaling in development of cancer. *Genet Mol Biol.* 2019 Jan 1;42(1):15.

462. Jiang J, Ren H-Y, Geng G-J, Mi Y-J, Liu Y, Li N, et al. Oncogenic activity of insulin in the development of non-small cell lung carcinoma. *Oncol Lett.* 2017/11/03. 2018 Jan;15(1):447–52.
463. Paddillaya N, Mishra A, Kondaiah P, Pullarkat P, Menon GI, Gundiah N. Biophysics of Cell-Substrate Interactions Under Shear. *Front cell Dev Biol.* 2019 Nov 8;7:251.
464. Azadi S, Tafazzoli Shadpour M. The microenvironment and cytoskeletal remodeling in tumor cell invasion. 1st ed. Vol. 356, *International Review of Cell and Molecular Biology.* Elsevier Inc.; 2020. 257–289 p.
465. Aseervatham J. Cytoskeletal Remodeling in Cancer. *Biology (Basel).* 2020 Nov 1;9(11):1–40.
466. Burkhart DL, Sage J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer.* 2008 Sep;8(9):671–82.
467. Deshpande A, Sicinski P, Hinds PW. Cyclins and cdks in development and cancer: a perspective. *Oncogene.* 2005;24(17):2909–15.
468. Sherr CJ, McCormick F. The RB and p53 pathways in cancer. *Cancer Cell.* 2002;2(2):103–12.
469. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature.* 2018;553(7689):446–54.
470. Petitjean A, Achatz MIW, Borresen-Dale AL, Hainaut P, Olivier M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. *Oncogene.* 2007;26(15):2157–65.
471. Labbé C, Cabanero M, Korpanty GJ, Tomasini P, Doherty MK, Mascaux C, et al. Prognostic and predictive effects of TP53 co-mutation in patients with EGFR -mutated non-small cell lung cancer (NSCLC). *Lung Cancer.* 2017;111:23–9.
472. Wang X, Sun Q. TP53 mutations, expression and interaction networks in human cancers. *Oncotarget.* 2017 Jan 3;8(1):624–43.
473. Chen M, Liu X, Du J, Wang X-J, Xia L. Differentiated regulation of immune-response related genes between LUAD and LUSC subtypes of lung cancers. *Oncotarget.* 2017 Jan 3;8(1):133–44.
474. Baronsky T, Ruhlandt D, Brückner BR, Schäfer J, Karedla N, Isbaner S, et al. Cell–Substrate Dynamics of the Epithelial-to-Mesenchymal Transition. *Nano Lett.* 2017;17(5):3320–6.
475. Charalabopoulos K, Gogali A, Kostoula OK, Constantopoulos SH. Cadherin superfamily of adhesion molecules in primary lung cancer. *Exp Oncol.* 2004;26(4):256–60.
476. Ma S-C, Li Q, Peng J-Y, Zhouwen J-L, Diao J-F, Niu J-X, et al. Claudin-5 regulates blood-brain barrier permeability by modifying brain microvascular endothelial cell proliferation, migration, and adhesion to prevent lung cancer metastasis. *CNS Neurosci Ther.* 2017/09/29. 2017 Dec;23(12):947–60.
477. Nelson WJ, Nusse R. Convergence of Wnt, beta-catenin, and cadherin pathways. *Science.* 2004 Mar 5;303(5663):1483–7.
478. Thiery JP. Epithelial–mesenchymal transitions in tumour progression. *Nat Rev Cancer.* 2002;2(6):442–54.
479. Ciruna B, Rossant J. FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. *Dev Cell.* 2001;1(1):37–49.
480. Nishita M, Hashimoto MK, Ogata S, Laurent MN, Ueno N, Shibuya H, et al. Interaction between Wnt and TGF- β signalling pathways during formation of Spemann’s organizer. *Nature.* 2000;403(6771):781–5.
481. Stewart DJ. Wnt signaling pathway in non-small cell lung cancer. *J Natl Cancer Inst.* 2014 Jan 1;106(1).
482. Singh M, Venugopal C, Tokar T, McFarlane N, Subapanditha MK, Qazi M, et al. Therapeutic Targeting of the Premetastatic Stage in Human Lung-to-Brain Metastasis. *Cancer Res.* 2018;78(17):5124–34.
483. Neman J, Termini J, Wilczynski S, Vaidehi N, Choy C, Kowolik CM, et al. Human breast cancer metastases to the brain display GABAergic properties in the neural niche. *Proc Natl Acad Sci U S A.* 2014/01/06. 2014 Jan 21;111(3):984–9.
484. Van Swearingen AE, Siegel MB, Anders CK. Breast cancer brain metastases: evidence for neuronal-like adaptation in a “breast-to-brain” transition? *Breast Cancer Res.* 2014 May 6;16(3):304.

485. Brastianos P, Carter S, Santagata S, Cahill D, Taylor-Weiner A, Jones R, et al. 2905 Genomic characterization of brain metastases and paired primary tumors reveals branched evolution and potential therapeutic targets. *Eur J Cancer*. 2015;51:S586.
486. de Sousa VML, Carvalho L. Heterogeneity in Lung Cancer. *Pathobiology*. 2018;85(1–2):96–107.
487. Boire A, Brastianos PK, Garzia L, Valiente M. Brain metastasis. *Nat Rev Cancer*. 2019;20(1):4–11.
488. Wingrove E, Liu ZZ, Patel KD, Arnal-Estapé A, Cai WL, Melnick M-A, et al. Transcriptomic Hallmarks of Tumor Plasticity and Stromal Interactions in Brain Metastasis. *Cell Rep*. 2019 Apr 23;27(4):1277-1292.e7.
489. Rybarczyk-Kasiuchnicz A, Ramlau R, Stencel K. Treatment of Brain Metastases of Non-Small Cell Lung Carcinoma. *Int J Mol Sci*. 2021 Jan 8;22(2):593.
490. Zhu TS, Costello MA, Talsma CE, Flack CG, Crowley JG, Hamm LL, et al. Endothelial cells create a stem cell niche in glioblastoma by providing NOTCH ligands that nurture self-renewal of cancer stem-like cells. *Cancer Res*. 2011/07/25. 2011 Sep 15;71(18):6061–72.
491. Kim K-J, Kwon S-H, Yun J-H, Jeong H-S, Kim H-R, Lee EH, et al. STAT3 activation in endothelial cells is important for tumor metastasis via increased cell adhesion molecule expression. *Oncogene*. 2017;36(39):5445–59.
492. Bathina S, Das UN. Brain-derived neurotrophic factor and its clinical implications. *Arch Med Sci*. 2015 Dec 12;11(6):1164.
493. Chen B, Liang Y, He Z, An Y, Zhao W, Wu J. Autocrine activity of BDNF induced by the STAT3 signaling pathway causes prolonged TrkB activation and promotes human non-small-cell lung cancer proliferation. *Sci Rep*. 2016 Jul 26;6:30404.
494. Ricci A, Greco S, Mariotta S, Felici L, Bronzetti E, Cavazzana A, et al. Neurotrophins and Neurotrophin Receptors in Human Lung Cancer. *Am J Respir Cell Mol Biol*. 2001;25(4):439–46.
495. Aoyama M, Asai K, Shishikura T, Kawamoto T, Miyachi T, Yokoi T, et al. Human neuroblastomas with unfavorable biologies express high levels of brain-derived neurotrophic factor mRNA and a variety of its variants. *Cancer Lett*. 2001;164(1):51–60.
496. Miknyoczki SJ, Lang D, Huang L, Klein-Szanto AJP, Dionne CA, Ruggeri BA. Neurotrophins and Trk receptors in human pancreatic ductal adenocarcinoma: Expression patterns and effects on *In vitro* invasive behavior. *Int J Cancer*. 1999;81(3):417–27.
497. Montano X, Djamgoz MBA. Epidermal growth factor, neurotrophins and the metastatic cascade in prostate cancer. *FEBS Lett*. 2004;571(1–3):1–8.
498. Yang ZF, Ho DW, Lam CT, Luk JM, Lum CT, Yu WC, et al. Identification of Brain-Derived Neurotrophic Factor as a Novel Functional Protein in Hepatocellular Carcinoma. *Cancer Res*. 2005;65(1):219–25.
499. Zhang X, Pei Z, Ji C, Zhang X, Xu J, Wang J. Novel Insights into the Role of the Cytoskeleton in Cancer. *Cytoskelet - Struct Dyn Funct Dis*. 2017 May 17;
500. Strube F, Infanger M, Wehland M, Delvinioti X, Romswinkel A, Dietz C, et al. Alteration of Cytoskeleton Morphology and Gene Expression in Human Breast Cancer Cells under Simulated Microgravity. *Cell J*. 2020;22(1):106.
501. Martino F, Perestrelo AR, Vinarský V, Pagliari S, Forte G. Cellular mechanotransduction: From tension to function. *Front Physiol*. 2018 Jul 5;9(JUL):824.
502. Seong J, Wang N, Wang Y. Mechanotransduction at focal adhesions: from physiology to cancer development. *J Cell Mol Med*. 2013 May 1;17(5):597–604.
503. Broders-Bondon F, Ho-Bouidoires THN, Fernandez-Sanchez ME, Farge E. Mechanotransduction in tumor progression: The dark side of the force. *J Cell Biol*. 2018 May 1;217(5):1571–87.
504. Suarez E, Knollmann-Ritschel BEC. Squamous Cell Carcinoma of the Lung. *Acad Pathol*. 2017;4.
505. Eckert RL, Rorke EA. Molecular biology of keratinocyte differentiation. *Environ Health Perspect*. 1989;80:109.
506. Park HJ, Cha Y-J, Kim SH, Kim A, Kim EY, Chang YS. Keratinization of Lung Squamous Cell Carcinoma Is Associated with Poor Clinical Outcome. *Tuberc Respir Dis (Seoul)*. 2017

- Apr 1;80(2):179.
507. Cardesa A, Nadal A B S T R A C T K E Y W O R D S A. Carcinoma of the head and neck in the HPV era.
 508. Xiao D, Jia J, Shi Y, Fu C, Chen L, Jiang Y, et al. Opposed expression of IKK α : loss in keratinizing carcinomas and gain in non-keratinizing carcinomas. *Oncotarget*. 2015 Aug 6;6(28):25499–505.
 509. Yamasaki T, Seki N, Yamada Y, Yoshino H, Hidaka H, Chiyomaru T, et al. Tumor suppressive microRNA-138 contributes to cell migration and invasion through its targeting of vimentin in renal cell carcinoma. *Int J Oncol*. 2012 Sep;41(3):805–17.
 510. Dauphin M, Barbe C, Lemaire S, Nawrocki-Raby B, Lagonotte E, Delepine G, et al. Vimentin expression predicts the occurrence of metastases in non small cell lung carcinomas. *Lung Cancer*. 2013 Jul;81(1):117–22.
 511. Berr AL, Wiese K, Santos G dos, Davis JM, Koch CM, Anekalla KR, et al. Vimentin is Required for Tumor Progression and Metastasis in a Mouse Model of Non-Small Cell Lung Cancer. *bioRxiv*. 2020 Jun 5;2020.06.04.130963.
 512. Richardson AM, Havel LS, Koyen AE, Konen JM, Shupe J, Wiles WG, et al. Vimentin is required for lung adenocarcinoma metastasis via heterotypic tumor cell-cancer-associated fibroblast interactions during collective invasion. *Clin Cancer Res*. 2018 Jan 1;24(2):420.
 513. Khan I, Steeg PS. Endocytosis: a pivotal pathway for regulating metastasis. *Br J Cancer* 2020 1241. 2020 Dec 2;124(1):66–75.
 514. Mellman I, Yarden Y. Endocytosis and Cancer. *Cold Spring Harb Perspect Biol*. 2013 Dec 1;5(12):a016949.
 515. Chen P-H, Bendris N, Hsiao Y-J, Reis CR, Mettlen M, Chen H-Y, et al. Crosstalk between CLCb/Dyn1-Mediated Adaptive Clathrin-Mediated Endocytosis and Epidermal Growth Factor Receptor Signaling Increases Metastasis. *Dev Cell*. 2017 Feb 6;40(3):278-288.e5.
 516. Reis CR, Chen P, Srinivasan S, Aguet F, Mettlen M, Schmid SL. Crosstalk between Akt/GSK3 β signaling and dynamin-1 regulates clathrin-mediated endocytosis. *EMBO J*. 2015 Aug 13;34(16):2132–46.
 517. Wang Y, Cao H, Chen J, McNiven MA. A direct interaction between the large GTPase dynamin-2 and FAK regulates focal adhesion dynamics in response to active Src. *Mol Biol Cell*. 2011 May 5;22(9):1529.
 518. Briñas L, Vassilopoulos S, Bonne G, Guicheney P, Bitoun M. Role of dynamin 2 in the disassembly of focal adhesions. *J Mol Med (Berl)*. 2013 Jul;91(7):803–9.
 519. Xu B, Teng LH, Silva SD da, Bijian K, Al Bashir S, Jie S, et al. The significance of dynamin 2 expression for prostate cancer progression, prognostication, and therapeutic targeting. *Cancer Med*. 2013/12/18. 2014 Feb;3(1):14–24.
 520. Liu J, Huang Y, Li T, Jiang Z, Zeng L, Hu Z. The role of the Golgi apparatus in disease (Review). *Int J Mol Med*. 2021 Apr 1;47(4):1–1.
 521. Migita T, Inoue S. Implications of the Golgi apparatus in prostate cancer. *Int J Biochem & Cell Biol*. 2012;44(11):1872–6.
 522. Li H, Zhang P, Luo J, Hu D, Huang Y, Zhang ZR, et al. Chondroitin Sulfate-Linked Prodrug Nanoparticles Target the Golgi Apparatus for Cancer Metastasis Treatment. *ACS Nano*. 2019 Aug 27;13(8):9386–96.
 523. Tao Y, Yang Y, Zhou R, Gong T. Golgi Apparatus: An Emerging Platform for Innate Immunity. *Trends Cell Biol*. 2020 Jun 1;30(6):467–77.
 524. Jiang Y, Ouyang W, Zhang C, Yu Y, Yao H. Prognosis and Immunotherapy Response With a Novel Golgi Apparatus Signature-Based Formula in Lung Adenocarcinoma. *Front Cell Dev Biol*. 2022 Jan 20;9:3831.
 525. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, et al. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Rep*. 2018;23(1):239-254.e6.
 526. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability — an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* 2010 113. 2010 Mar;11(3):220–8.
 527. Broustas CG, Lieberman HB. DNA damage response genes and the development of cancer metastasis. *Radiat Res*. 2014/01/07. 2014 Feb;181(2):111–30.

528. Friedberg EC. How nucleotide excision repair protects against cancer. *Nat Rev Cancer*. 2001;1(1):22–33.
529. Villaruz LC, Jones H, Dacic S, Abberbock S, Kurland BF, Stabile LP, et al. ATM protein is deficient in over 40% of lung adenocarcinomas. *Oncotarget*. 2016 Sep 6;7(36):57714–25.
530. Liu G, Pei F, Yang F, Li L, Amin AD, Liu S, et al. Role of Autophagy and Apoptosis in Non-Small-Cell Lung Cancer. *Int J Mol Sci*. 2017 Feb 10;18(2):367.
531. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. *Cell*. 2008 Jan 11;132(1):27–42.
532. Galluzzi L, Pietrocola F, Bravo-San Pedro JM, Amaravadi RK, Baehrecke EH, Cecconi F, et al. Autophagy in malignant transformation and cancer progression. *EMBO J*. 2015/02/23. 2015 Apr 1;34(7):856–80.
533. Yun CW, Jeon J, Go G, Lee JH, Lee SH. The Dual Role of Autophagy in Cancer Development and a Therapeutic Strategy for Cancer by Targeting Autophagy. *Int J Mol Sci*. 2020 Dec 26;22(1):179.
534. Martinon F, Tschopp J. Inflammatory Caspases. *Cell*. 2004;117(5):561–74.
535. Mariño G, Niso-Santano M, Baehrecke EH, Kroemer G. Self-consumption: the interplay of autophagy and apoptosis. *Nat Rev Mol Cell Biol*. 2014/01/08. 2014 Feb;15(2):81–94.
536. Kroemer G, Mariño G, Levine B. Autophagy and the integrated stress response. *Mol Cell*. 2010 Oct 22;40(2):280–93.
537. White E. The role for autophagy in cancer. *J Clin Invest*. 2015/01/02. 2015 Jan;125(1):42–6.
538. Ye M-X, Li Y, Yin H, Zhang J. Curcumin: updated molecular mechanisms and intervention targets in human lung cancer. *Int J Mol Sci*. 2012/03/22. 2012;13(3):3959–78.
539. Hanahan D, Folkman J. Patterns and Emerging Mechanisms of the Angiogenic Switch during Tumorigenesis. *Cell*. 1996;86(3):353–64.
540. Fouad YA, Aanei C. Revisiting the hallmarks of cancer. *Am J Cancer Res*. 2017;7(5):1016.
541. Li Y, Zhao L, Li XF. Hypoxia and the Tumor Microenvironment. *Technol Cancer Res Treat*. 2021 Aug 5;20.
542. Holmes DR, Zachary I. The vascular endothelial growth factor (VEGF) family: Angiogenic factors in health and disease. *Genome Biol*. 2005 Feb 1;6(2):1–10.
543. Wang X, Bove AM, Simone G, Ma B. Molecular Bases of VEGFR-2-Mediated Physiological Function and Pathological Role. *Front Cell Dev Biol*. 2020 Nov 16;8:599281.
544. Mac Gabhann F, Popel AS. Systems biology of vascular endothelial growth factors. *Microcirculation*. 2008 Nov;15(8):715–38.
545. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell*. 2010 Apr 2;141(1):52–67.
546. Baeriswyl V, Christofori G. The angiogenic switch in carcinogenesis. *Semin Cancer Biol*. 2009;19(5):329–37.
547. Tamura R, Tanaka T, Akasaki Y, Murayama Y, Yoshida K, Sasaki H. The role of vascular endothelial growth factor in the hypoxic and immunosuppressive tumor microenvironment: perspectives for therapeutic implications. *Med Oncol*. 2020 Jan 1;37(1):1–14.
548. Marchand GS, Noiseux N, Tanguay J-F, Sirois MG. Blockade of in vivo VEGF-mediated angiogenesis by antisense gene therapy: role of Flk-1 and Flt-1 receptors. *Am J Physiol Circ Physiol*. 2002;282(1):H194–204.
549. Jain RK, Duda DG, Clark JW, Loeffler JS. Lessons from phase III clinical trials on anti-VEGF therapy for cancer. *Nat Clin Pract Oncol*. 2006;3(1):24–40.
550. Heymach J V., Desai J, Manola J, Davis DW, McConkey DJ, Harmon D, et al. Phase II study of the antiangiogenic agent SU5416 in patients with advanced soft tissue sarcomas. *Clin Cancer Res*. 2004 Sep 1;10(17):5732–40.
551. Mendes C, Serpa J. Metabolic Remodelling: An Accomplice for New Therapeutic Strategies to Fight Lung Cancer. *Antioxidants*. 2019 Dec 1;8(12).
552. Orita H, Coulter J, Lemmon C, Tully E, Vadlamudi A, Medghalchi SM, et al. Selective inhibition of fatty acid synthase for lung cancer treatment. *Clin Cancer Res*. 2007 Dec 1;13(23):7139–45.
553. Eltayeb K, La Monica S, Tiseo M, Alfieri R, Fumarola C. Reprogramming of Lipid Metabolism in Lung Cancer: An Overview with Focus on EGFR-Mutated Non-Small Cell

- Lung Cancer. *Cells*. 2022 Jan 25;11(3):413.
554. Rothman JE, Lenard J. Membrane asymmetry. *Science*. 1977;195(4280):743–53.
555. Preta G. New Insights Into Targeting Membrane Lipids for Cancer Therapy. *Front Cell Dev Biol*. 2020 Sep 2;8.
556. Bernardes N, Fialho AM. Perturbing the Dynamics and Organization of Cell Membrane Components: A New Paradigm for Cancer-Targeted Therapies. *Int J Mol Sci* 2018, Vol 19, Page 3871. 2018 Dec 4;19(12):3871.
557. Pan Z, Wang K, Wang X, Jia Z, Yang Y, Duan Y, et al. Cholesterol promotes EGFR-TKIs resistance in NSCLC by inducing EGFR/Src/Erk/SP1 signaling-mediated ERR α re-expression. *Mol Cancer*. 2022 Dec 1;21(1):1–17.
558. Baek AE, Yu YRA, He S, Wardell SE, Chang CY, Kwon S, et al. The cholesterol metabolite 27 hydroxycholesterol facilitates breast cancer metastasis through its actions on immune cells. *Nat Commun* 2017 81. 2017 Oct 11;8(1):1–11.
559. Ma X, Bi E, Lu Y, Su P, Huang C, Liu L, et al. Cholesterol Induces CD8+ T Cell Exhaustion in the Tumor Microenvironment. *Cell Metab*. 2019 Jul 2;30(1):143-156.e5.
560. Xiang Y, Miao H. Lipid Metabolism in Tumor-Associated Macrophages. *Adv Exp Med Biol*. 2021;1316:87–101.
561. Li J, Li Q, Su Z, Sun Q, Zhao Y, Feng T, et al. Lipid metabolism gene-wide profile and survival signature of lung adenocarcinoma. *Lipids Health Dis*. 2020 Dec 1;19(1):1–9.
562. Gaber T, Strehl C, Buttgerit F. Metabolic regulation of inflammation. *Nat Rev Rheumatol*. 2017 May 1;13(5):267–79.
563. Cheng T, Zhang J, Liu D, Lai G, Wen X. Prognosis of Non-small-cell Lung Cancer Patients With Lipid Metabolism Pathway Alternations to Immunotherapy. *Front Genet*. 2021 Jul 14;12:1156.
564. Wang Y, Wang F. Post-Translational Modifications of Deubiquitinating Enzymes: Expanding the Ubiquitin Code. *Front Pharmacol*. 2021 Jun 10;12:685011.
565. Gallo LH, Ko J, Donoghue DJ. The importance of regulatory ubiquitination in cancer and metastasis. *Cell Cycle*. 2017/02/06. 2017 Apr 3;16(7):634–48.
566. Young M-J, Hsu K-C, Lin TE, Chang W-C, Hung J-J. The role of ubiquitin-specific peptidases in cancer progression. *J Biomed Sci*. 2019 May 27;26(1):42.
567. Bui QT, Hong JH, Kwak M, Lee JY, Lee PCW. Ubiquitin-Conjugating Enzymes in Cancer. *Cells* 2021, Vol 10, Page 1383. 2021 Jun 4;10(6):1383.
568. Swatek KN, Komander D. Ubiquitin modifications. *Cell Res*. 2016/03/25. 2016 Apr;26(4):399–422.
569. Haglund K, Dikic I. Ubiquitylation and cell signaling. *EMBO J*. 2005/09/08. 2005 Oct 5;24(19):3353–9.
570. Pal A, Donato NJ. Ubiquitin-specific proteases as therapeutic targets for the treatment of breast cancer. *Breast Cancer Res*. 2014;16(5):461.
571. Chen X, Dou QP, Liu J, Tang D. Targeting Ubiquitin–Proteasome System With Copper Complexes for Cancer Therapy. *Front Mol Biosci*. 2021 Apr 13;8:116.
572. Denoyer D, Masaldan S, La Fontaine S, Cater MA. Targeting copper in cancer therapy: ‘Copper That Cancer.’ *Metallomics*. 2015;7(11):1459–76.
573. Li Y. Copper homeostasis: Emerging target for cancer treatment. *IUBMB Life*. 2020;72(9):1900–8.
574. Ren S, Xiong X, You H, Shen J, Zhou P. The Combination of Immune Checkpoint Blockade and Angiogenesis Inhibitors in the Treatment of Advanced Non-Small Cell Lung Cancer. *Front Immunol*. 2021 Jun 2;12:689132.
575. Lee WS, Yang H, Chon HJ, Kim C. Combination of anti-angiogenic therapy and immune checkpoint blockade normalizes vascular-immune crosstalk to potentiate cancer immunity. *Exp Mol Med*. 2020/09/11. 2020 Sep;52(9):1475–85.
576. Morikawa M, Derynck R, Miyazono K. TGF- β and the TGF- β Family: Context-Dependent Roles in Cell and Tissue Physiology. *Cold Spring Harb Perspect Biol*. 2016 May 2;8(5):a021873.
577. Derynck R, Akhurst RJ, Balmain A. TGF- β signaling in tumor suppression and cancer progression. *Nat Genet*. 2001;29(2):117–29.

578. Itoh S, Itoh F, Goumans M-J, ten Dijke P. Signaling of transforming growth factor- β family members through Smad proteins. *Eur J Biochem.* 2000;267(24):6954–67.
579. Massagué J. TGF β in Cancer. *Cell.* 2008 Jul 25;134(2):215–30.
580. Lebrun J-J. The Dual Role of TGF β in Human Cancer: From Tumor Suppression to Cancer Metastasis. *ISRN Mol Biol.* 2012 Dec 24;2012:381428.
581. EL-DANASOURI I, FRANCES A, WESTPHAL LM. Immunocytochemical Localization of Transforming Growth Factor- α and Epidermal Growth Factor Receptor in Human Fallopian Tubes and Cumulus Cells. *Am J Reprod Immunol.* 1993;30(2–3):82–7.
582. Fanelli GN, Naccarato AG, Scatena C. Recent Advances in Cancer Plasticity: Cellular Mechanisms, Surveillance Strategies, and Therapeutic Optimization. *Front Oncol.* 2020;10(April):1–10.
583. Zakowski MF, Ladanyi M, Kris MG. *EGFR* Mutations in Small-Cell Lung Cancers in Patients Who Have Never Smoked. *N Engl J Med.* 2006;355(2):213–5.
584. Clamon G, Zeitler W, An J, Hejleh TA. Transformational Changes Between Non-Small Cell and Small Cell Lung Cancer—Biological and Clinical Relevance—A Review. *Am J Clin Oncol.* 2020;43(9):670–5.
585. Yin X, Li Y, Wang H, Jia T, Wang E, Luo Y, et al. Small cell lung cancer transformation: From pathogenesis to treatment. *Semin Cancer Biol.* 2022 Nov 1;86:595–606.
586. Walker GE, Antoniono RJ, Ross HJ, Paisley TE, Oh Y. Neuroendocrine-like differentiation of non-small cell lung carcinoma cells: regulation by cAMP and the interaction of mac25/IGFBP-rP1 and 25.1. *Oncogene.* 2006;25(13):1943–54.
587. Hokari S, Tamura Y, Kaneda A, Katsura A, Morikawa M, Murai F, et al. Comparative analysis of TTF-1 binding DNA regions in small-cell lung cancer and non-small-cell lung cancer. *Mol Oncol.* 2020 Feb 1;14(2):277–93.
588. Zhang L, Whitsett JA, Stripp BR. Regulation of Clara cell secretory protein gene transcription by thyroid transcription factor-1. *Biochim Biophys Acta.* 1997 Mar 28;1350(3):359–67.
589. Yatabe Y, Mitsudomi T, Takahashi T. TTF-1 expression in pulmonary adenocarcinomas. *Am J Surg Pathol.* 2002;26(6):767–73.
590. Bellusci S, Grindley J, Emoto H, Itoh N, Hogan BL. Fibroblast growth factor 10 (FGF10) and branching morphogenesis in the embryonic mouse lung. *Development.* 1997;124(23):4867–78.
591. Oser MG, Niederst MJ, Sequist L V, Engelman JA. Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin. *Lancet Oncol.* 2015 Apr;16(4):e165–72.
592. Lu Y, Chang Q, Zhang Y, Beezhold K, Rojanasakul Y, Zhao H, et al. Lung cancer-associated JmjC domain protein mdig suppresses formation of tri-methyl lysine 9 of histone H3. *Cell Cycle.* 2009 Jul 1;8(13):2101–9.
593. Wu P, Wu D, Li L, Chai Y, Huang J. PD-L1 and Survival in Solid Tumors: A Meta-Analysis. *PLoS One.* 2015 Jun 26;10(6):e0131403–e0131403.
594. Pawelczyk K, Piotrowska A, Ciesielska U, Jablonska K, Gletzel-Plucinska N, Grzegorzolka J, et al. Role of PD-L1 expression in non-small cell lung cancer and their prognostic significance according to clinicopathological factors and diagnostic markers. *Int J Mol Sci.* 2019;20(4):1–15.
595. Lucibello G, Mograbi B, Milano G, Hofman P, Brest P. PD-L1 regulation revisited: impact on immunotherapeutic strategies. *Trends Mol Med.* 2021;27(9):868–81.
596. Zerdes I, Matikas A, Bergh J, Rassidakis GZ, Foukakis T. Genetic, transcriptional and post-translational regulation of the programmed death protein ligand 1 in cancer: biology and clinical correlations. *Oncogene.* 2018;37(34):4639–61.
597. Saigi M, Albuquerque-Bejar JJ, Sanchez-Cespedes M. Determinants of immunological evasion and immunecheckpoint inhibition response in non-small cell lung cancer: the genetic front. *Oncogene.* 2019;
598. Sabari JK, Leonardi GC, Shu CA, Umeton R, Montecalvo J, Ni A, et al. PD-L1 expression, tumor mutational burden, and response to immunotherapy in patients with MET exon 14 altered lung cancers. *Ann Oncol Off J Eur Soc Med Oncol.* 2018;29(10):2085–91.
599. Yamaguchi H, Hsu JM, Yang WH, Hung MC. Mechanisms regulating PD-L1 expression in cancers and associated opportunities for novel small-molecule therapeutics. *Nat Rev Clin*

- Oncol 2022 195. 2022 Feb 7;19(5):287–305.
600. Hirano T, Ishihara K, Hibi M. Roles of STAT3 in mediating the cell growth, differentiation and survival signals relayed through the IL-6 family of cytokine receptors. *Oncogene*. 2000 May 15;19(21):2548–56.
 601. Rébé C, Végran F, Berger H, Ghiringhelli F. STAT3 activation. <http://dx.doi.org/104161/jkst23010>. 2013 Jan;2(1):e23010.
 602. Mohrherr J, Uras IZ, Moll HP, Casanova E. STAT3: Versatile Functions in Non-Small Cell Lung Cancer. *Cancers* 2020, Vol 12, Page 1107. 2020 Apr 29;12(5):1107.
 603. Verhoeven Y, Tilborghs S, Jacobs J, De Waele J, Quatannens D, Deben C, et al. The potential and controversy of targeting STAT family members in cancer. *Semin Cancer Biol*. 2020;60(June 2019):41–56.
 604. Rébé C, Ghiringhelli F. STAT3, a Master Regulator of Anti-Tumor Immune Response. *Cancers (Basel)*. 2019 Sep 1;11(9).
 605. Carbognin E, Betto RM, Soriano ME, Smith AG, Martello G. Stat3 promotes mitochondrial transcription and oxidative respiration during maintenance and induction of naive pluripotency. *EMBO J*. 2016 Mar 15;35(6):618–34.
 606. Igelmann S, Neubauer HA, Ferbeyre G. STAT3 and STAT5 activation in solid cancers. *Cancers (Basel)*. 2019;11(10).
 607. Mandai M, Hamanishi J, Abiko K, Matsumura N, Baba T, Konishi I. Dual Faces of IFN γ in Cancer Progression: A Role of PD-L1 Induction in the Determination of Pro- and Antitumor Immunity. *Clin Cancer Res*. 2016;22(10):2329–34.
 608. Hu R, Han Q, Zhang J. STAT3: A key signaling molecule for converting cold to hot tumors. *Cancer Lett*. 2020;489(44):29–40.
 609. Carnero A, Blanco-Aparicio C, Renner O, Link W, Leal J. The PTEN/PI3K/AKT signalling pathway in cancer, therapeutic implications. *Curr Cancer Drug Targets*. 2008 May 6;8(3):187–98.
 610. Vidotto T, Melo CM, Castelli E, Koti M, dos Reis RB, Squire JA. Emerging role of PTEN loss in evasion of the immune response to tumours. *Br J Cancer*. 2020;122(12):1732–43.
 611. Taylor H, Laurence ADJ, Uhlig HH. The role of PTEN in innate and adaptive immunity. *Cold Spring Harb Perspect Med*. 2019;9(12):1–22.
 612. Moon YW, Hajjar J, Hwu P, Naing A. Targeting the indoleamine 2,3-dioxygenase pathway in cancer. *J Immunother Cancer*. 2015;3.
 613. Sun SY. Searching for the real function of mTOR signaling in the regulation of PD-L1 expression. *Transl Oncol*. 2020 Dec 1;13(12).
 614. Duran I, Kortmansky J, Singh D, Hirte H, Kocha W, Goss G, et al. A phase II clinical and pharmacodynamic study of temsirolimus in advanced neuroendocrine carcinomas. *Br J Cancer*. 2006 Nov 6;95(9):1148–54.
 615. Liang S-Q, Bühner ED, Berezowska S, Marti TM, Xu D, Froment L, et al. mTOR mediates a mechanism of resistance to chemotherapy and defines a rational combination strategy to treat KRAS-mutant lung cancer. *Oncogene*. 2018;38(5):622–36.
 616. Yokogami K, Wakisaka S, Avruch J, Reeves SA. Serine phosphorylation and maximal activation of STAT3 during CNTF signaling is mediated by the rapamycin target mTOR. *Curr Biol*. 2000;10(1):47–50.
 617. Nicholson RI, Gee JMW, Harper ME. EGFR and cancer prognosis. *Eur J Cancer*. 2001;37(SUPPL. 4):9.
 618. WC Chan L, F Wang F, CS Cho W. Genomic Sequence Analysis of EGFR Regulation by MicroRNAs in Lung Cancer. *Curr Top Med Chem*. 2012;12(8):920–6.
 619. T K, Y Y, H E, H K, T T, T M. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. *Cancer Res*. 2004 Dec 15;64(24):8919–23.
 620. Brabender J, Danenberg KD, Metzger R, Schneider PM, Park J, Salonga D, et al. Epidermal Growth Factor Receptor and HER2-neu mRNA Expression in Non-Small Cell Lung Cancer Is Correlated with Survival. 2001;
 621. Wee P, Wang Z. Epidermal growth factor receptor cell proliferation signaling pathways. *Cancers (Basel)*. 2017;9(5):1–45.
 622. Komiya K, Sueoka-Aragane N, Sato A, Hisatomi T, Sakuragi T, Mitsuoka M, et al. Mina53, a

- novel c-Myc target gene, is frequently expressed in lung cancers and exerts oncogenic property in NIH/3T3 cells. *J Cancer Res Clin Oncol*. 2010;136(3):465–73.
623. Ramaiah MJ, Kumar KR. mTOR-Rictor-EGFR axis in oncogenesis and diagnosis of glioblastoma multiforme. *Mol Biol Rep*. 2021 May 1;48(5):4813–35.
624. Zhang N, Zeng Y, Du W, Zhu J, Shen D, Liu Z, et al. The EGFR pathway is involved in the regulation of PD-L1 expression via the IL-6/JAK/STAT3 signaling pathway in EGFR-mutated non-small cell lung cancer. *Int J Oncol*. 2016;49(4):1360–8.
625. Sun J, Yu M, Lu Y, Thakur C, Chen B, Qiu P, et al. Carcinogenic metalloid arsenic induces expression of mdig oncogene through JNK and STAT3 activation. *Cancer Lett*. 2014;346(2):257–63.
626. Ma D, Guo D, Li W, Zhao H. Mdig, a lung cancer-associated gene, regulates cell cycle progression through p27KIP1. *Tumor Biol*. 2015;36(9):6909–17.
627. Geng F, Jiang Z, Song X, Zhou H, Zhao H. Mdig suppresses epithelial-mesenchymal transition and inhibits the invasion and metastasis of non-small cell lung cancer via regulating GSK-3 β / β -catenin signaling. *Int J Oncol*. 2017;51(6):1898–908.
628. Artificial Intelligence: A Modern Approach, 4th US ed. [Internet]. [cited 2022 Oct 30]. Available from: <http://aima.cs.berkeley.edu/>
629. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 1998 Dec 8;95(25):14863–8.
630. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013 Apr 2;6(269).
631. Monti P, Menichini P, Speciale A, Cutrona G, Fais F, Taiana E, et al. Heterogeneity of TP53 Mutations and P53 Protein Residual Function in Cancer: Does It Matter? *Front Oncol*. 2020 Oct 28;10.
632. Kamaraj B, Bogaerts A. Structure and Function of p53-DNA Complexes with Inactivation and Rescue Mutations: A Molecular Dynamics Simulation Study. *PLoS One*. 2015 Aug 5;10(8):e0134638.
633. Bradford CR, Zhu S, Poore J, Fisher SG, Beals TF, Thoraval D, et al. p53 mutation as a prognostic marker in advanced laryngeal carcinoma. Department of Veterans Affairs Laryngeal Cancer Cooperative Study Group. *Arch Otolaryngol Head Neck Surg*. 1997;123(6):605–9.
634. Rivlin N, Brosh R, Oren M, Rotter V. Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. *Genes Cancer*. 2011 Apr;2(4):466.
635. Zhou X, Edmonson MN, Wilkinson MR, Patel A, Wu G, Liu Y, et al. Exploring genomic alteration in pediatric cancer using ProteinPaint. *Nat Genet* 2015 481. 2015 Dec 29;48(1):4–6.
636. Bykov VJN, Eriksson SE, Bianchi J, Wiman KG. Targeting mutant p53 for efficient cancer therapy. *Nat Rev Cancer*. 2017;18(2):89–102.
637. Qin K, Hou H, Liang Y, Zhang X. Prognostic value of TP53 concurrent mutations for EGFR-TKIs and ALK-TKIs based targeted therapy in advanced non-small cell lung cancer: a meta-analysis. *BMC Cancer*. 2020 Apr 16;20(1):328.
638. Cheng Y, Ma L, Liu Y, Zhu J, Xin Y, Liu X, et al. Comprehensive characterization and clinical impact of concomitant genomic alterations in EGFR-mutant NSCLCs treated with EGFR kinase inhibitors. *Lung Cancer*. 2020;145:63–70.
639. Poeta ML, Manola J, Goldwasser MA, Forastiere A, Benoit N, Califano JA, et al. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. *N Engl J Med*. 2007 Dec 20;357(25):2552–61.
640. Molina-Vila MA, Bertran-Alamillo J, Gascó A, Mayo-de-las-Casas C, Sánchez-Ronco M, Pujantell-Pastor L, et al. Nondisruptive p53 Mutations Are Associated with Shorter Survival in Patients with Advanced Non-Small Cell Lung Cancer. *Clin Cancer Res*. 2014;20(17):4647–59.
641. Canale M, Petracci E, Delmonte A, Chiadini E, Dazzi C, Papi M, et al. Impact of TP53 Mutations on Outcome in EGFR-Mutated Patients Treated with First-Line Tyrosine Kinase Inhibitors. *Clin Cancer Res*. 2017;23(9):2195–202.

642. Song P, Zhang F, Li Y, Yang G, Li W, Ying J, et al. Concomitant TP53 mutations with response to crizotinib treatment in patients with ALK-rearranged non-small-cell lung cancer. *Cancer Med.* 2019/03/07. 2019 Apr;8(4):1551–7.
643. Jiao X-D, Qin B-D, You P, Cai J, Zang Y-S. The prognostic value of TP53 and its correlation with EGFR mutation in advanced non-small cell lung cancer, an analysis based on cBioPortal data base. *Lung Cancer.* 2018;123:70–5.
644. Fu Y, Wang A, Zhou J, Feng W, Shi M, Xu X, et al. Advanced NSCLC Patients With EGFR T790M Harboring TP53 R273C or KRAS G12V Cannot Benefit From Osimertinib Based on a Clinical Multicentre Study by Tissue and Liquid Biopsy. *Front Oncol.* 2021 Feb 24;11:621992.
645. Zheng C, Li X, Ren Y, Yin Z, Zhou B. Coexisting EGFR and TP53 Mutations in Lung Adenocarcinoma Patients Are Associated With COMP and ITGB8 Upregulation and Poor Prognosis. *Front Mol Biosci.* 2020 Feb 27;7:30.
646. Liu Y, Wu A, Li X, Wang S, Fang S, Mo Y. A retrospective analysis of eleven gene mutations, PD-L1 expression and clinicopathological characteristics in non-small cell lung cancer patients. *Asian J Surg.* 2022;45(1):367–75.
647. Sun H, Liu S-Y, Zhou J-Y, Xu J-T, Zhang H-K, Yan H-H, et al. Specific TP53 subtype as biomarker for immune checkpoint inhibitors in lung adenocarcinoma. *EBioMedicine.* 2020/09/11. 2020 Oct;60:102990.
648. Jin Y, Xue Q, Shen X, Zheng Q, Chen H, Zhou X, et al. PD-L1 Expression and Comprehensive Molecular Profiling Predict Survival in Nonsmall Cell Lung Cancer: A Real-World Study of a Large Chinese Cohort. *Clin Lung Cancer.* 2022;23(1):43–51.
649. Fernández-Trujillo L, Garcia-Robledo JE, Zúñiga-Restrepo V, Sua LF. Clinical characteristics and PD-L1 expression in primary lung squamous cell carcinoma: A case series. *Respir Med Case Reports.* 2020 Jan 1;30:101114.
650. Tímár J. The clinical relevance of KRAS gene mutation in non-small-cell lung cancer. *Curr Opin Oncol.* 2014;26(2):138–44.
651. O'Bryan JP. Pharmacological targeting of RAS: Recent success with direct inhibitors. *Pharmacol Res.* 2018/10/23. 2019 Jan;139:503–11.
652. Ghimessy A, Radeczky P, Laszlo V, Hegedus B, Renyi-Vamos F, Fillinger J, et al. Current therapy of KRAS-mutant lung cancer. *Cancer Metastasis Rev.* 2020 Dec 1;39(4):1159.
653. Riely GJ, Kris MG, Rosenbaum D, Marks J, Li A, Chitale DA, et al. Frequency and distinctive spectrum of KRAS mutations in never smokers with lung adenocarcinoma. *Clin Cancer Res.* 2008 Sep 15;14(18):5731–4.
654. Ferrer I, Zugazagoitia J, Herbertz S, John W, Paz-Ares L, Schmid-Bindert G. KRAS-Mutant non-small cell lung cancer: From biology to therapy. *Lung Cancer.* 2018 Oct 1;124:53–64.
655. Hames ML, Chen H, Iams W, Aston J, Lovly CM, Horn L. Correlation between KRAS mutation status and response to chemotherapy in patients with advanced non-small cell lung cancer☆. *Lung Cancer.* 2016 Feb 1;92:29–34.
656. Shepherd FA, Lacas B, Le Teuff G, Hainaut P, Jänne PA, Pignon J-P, et al. Pooled Analysis of the Prognostic and Predictive Effects of TP53 Comutation Status Combined With KRAS or EGFR Mutation in Early-Stage Resected Non-Small-Cell Lung Cancer in Four Trials of Adjuvant Chemotherapy. *J Clin Oncol.* 2017/04/28. 2017 Jun 20;35(18):2018–27.
657. Sun JM, Hwang DW, Ahn JS, Ahn MJ, Park K. Prognostic and Predictive Value of KRAS Mutations in Advanced Non-Small Cell Lung Cancer. *PLoS One.* 2013 May 28;8(5):e64816.
658. Linardou H, Dahabreh IJ, Kanaloupiti D, Siannis F, Bafaloukos D, Kosmidis P, et al. Assessment of somatic k-RAS mutations as a mechanism associated with resistance to EGFR-targeted agents: a systematic review and meta-analysis of studies in advanced non-small-cell lung cancer and metastatic colorectal cancer. *Lancet Oncol.* 2008 Oct 1;9(10):962–72.
659. Eberhard DA, Johnson BE, Amler LC, Goddard AD, Heldens SL, Herbst RS, et al. Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol.* 2005;23(25):5900–9.
660. Massarelli E, Varella-Garcia M, Tang X, Xavier AC, Ozburn NC, Liu DD, et al. 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.* 2007 May 15;13(10):2890–6.
661. Garassino MC, Martelli O, Broggin M, Farina G, Veronese S, Rulli E, et al. Erlotinib versus docetaxel as second-line treatment of patients with advanced non-small-cell lung cancer and wild-type EGFR tumours (TAILOR): a randomised controlled trial. *Lancet Oncol.* 2013 Sep 1;14(10):981–8.
662. Mao C, Qiu LX, Liao RY, Du FB, Ding H, Yang WC, et al. KRAS mutations and resistance to EGFR-TKIs treatment in patients with non-small cell lung cancer: A meta-analysis of 22 studies. *Lung Cancer.* 2010 Sep 1;69(3):272–8.
663. Reiniger L, Téglási V, Pipek O, Rojkó L, Glasz T, Vágvölgyi A, et al. Tumor necrosis correlates with PD-L1 and PD-1 expression in lung adenocarcinoma. *Acta Oncol (Madr).* 2019 Aug 3;58(8):1087–94.
664. Johnson JL, Pillai S, Chellappan SP. Genetic and biochemical alterations in non-small cell lung cancer. *Biochem Res Int.* 2012;2012.
665. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell.* 2012 Sep 14;150(6):1107–20.
666. Shire NJ, Klein AB, Golozar A, Collins JM, Fraeman KH, Nordstrom BL, et al. STK11 (LKB1) mutations in metastatic NSCLC: Prognostic value in the real world. *PLoS One.* 2020 Sep 1;15(9):e0238358.
667. Shackelford DB, Shaw RJ. The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nat Rev Cancer.* 2009 Aug;9(8):563–75.
668. Corte CM Della, Byers LA. Evading the STING: LKB1 Loss Leads to STING Silencing and Immune Escape in KRAS-Mutant Lung Cancers. *Cancer Discov.* 2019 Jan 1;9(1):16–8.
669. Xu C, Fillmore C, Koyama S, Wu H, Zhao Y, Chen Z, et al. Loss of Lkb1 and Pten Leads to Lung Squamous Cell Carcinoma with Elevated PD-L1 Expression. *Cancer Cell.* 2014;25(5):590–604.
670. Koyama S, Akbay EA, Li YY, Aref AR, Skoulidis F, Herter-Sprie GS, et al. STK11/LKB1 Deficiency Promotes Neutrophil Recruitment and Proinflammatory Cytokine Production to Suppress T-cell Activity in the Lung Tumor Microenvironment. *Cancer Res.* 2016 Mar 1;76(5):999–1008.
671. Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, et al. STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. *Cancer Discov.* 2018 Jul 1;8(7):822–35.
672. Bange E, Marmarelis ME, Hwang W-T, Yang Y-X, Thompson JC, Rosenbaum J, et al. Impact of KRAS and TP53 Co-Mutations on Outcomes After First-Line Systemic Therapy Among Patients With STK11-Mutated Advanced Non-Small-Cell Lung Cancer. *JCO Precis Oncol.* 2019 Dec;3(3):1–11.

Електронни източници

E1:<https://www.wikipathways.org>

E2: <https://cancer.sanger.ac.uk/cosmic>

E3: <https://www.ncbi.nlm.nih.gov/snp/>

E4:<https://string-db.org>