

**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**

**АВТОРЕФЕРАТ**

**НА 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

Дисертационният труд съдържа 247 страници, 128 фигури и 26 таблици.

Цитирани са 672 литературни и 4 електронни източника.

Дисертационният труд е обсъден и насочен за защита от Катедра по Медицинска Химия и Биохимия, Медицински Университет – София.

Всички включени в дисертацията изследвания са извършени в Център по Молекулна Медицина, Катедра по Медицинска Химия и Биохимия, Медицински Университет – София.

Настоящото изследване е финансирано от проекти Договор №ДН03/14 от 19.12.2016., Договор № Д-138/02.05.2017, Договор № Д-138/02.05.2017, Договор № Д-№ 70/23.04.2019 г.

Дисертационният труд е представен на заседания на катедрен съвет на Катедра по Медицинска Химия и Биохимия, Медицински Университет – София на 22.03.2021 г.

Официалната защита на дисертационния труд ще се състои на 18 януари 2023 г. в Катедра по Медицинска Химия и Биохимия, Медицински Университет, ул. Здраве 2, гр. София пред научно жури в състав:

Председател:

Проф. Алаксей Славков Савов, дб

Членове:

Проф. д-р Иванка Истилианова Димова, дм

Проф. Сорен Бохос Хайрабемян, дбн

Проф. Татяна Иванова Влайкова, дб

Проф. Румяна Силвиева Миронова, дб

Резервни членове:

Проф. д-р Милена Иванова Иванова-Шиварова, дм

Проф. Ана Иванова Манева, дб

Номерацията на фигурите и таблиците не отговаря на тази в дисертационния труд.

Материалите по защитата са на разположение в Център по Молекулна Медицина, сградата на СБАЛАГ „Майчин дом“, ул. Здраве 2, ет. 14.

Съдържание	
ИЗПОЛЗВАНИ СЪКРАЩЕНИЯ:	5
1. ВЪВЕДЕНИЕ	6
2. ЦЕЛ И ЗАДАЧИ	7
2.1. ЦЕЛ	7
2.2. ЗАДАЧИ	7
3. МАТЕРИАЛИ И МЕТОДИ	8
3.1. Материали	8
3.2. Методи	8
4. РЕЗУЛТАТИ	9
4.1. Установяване на микроРНК с променена експресия при АК и ПК в сравнение с прилежащите им нормални тъкани	9
4.2. Резултати от йерархичен клъстерен анализ	9
4.3. Избор на микроРНК за валидация	10
4.4. Валидация на избраните микроРНК чрез RT-qPCR	11
4.5. Нива на експресия на 14-те изследвани микроРНК в тъканни проби от рак на белия дроб	11
4.6. Асоциация между експресионните нива на микроРНКите и клиникопатологичните характеристики на пациентите	15
4.7. Корелационен анализ между изследваните микроРНК	16
4.8. Анализ на гените – мишени на микроРНКите	20
4.9. Онтологичен анализ	20
4.10. Избор на мРНК, регулирани от показалите диференциална експресия микроРНК-и в изследваните АК и ПК и имащи важно клинично значение при РБД	24
4.11. 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	24
4.11.1. Корелационен анализ между изследваните иРНК и микроРНК	26
4.11.2. Асоциация между експресионните нива на иРНКите и клиникопатологичните характеристики на пациентите	27
4.12. Резултати от контролираното самообучение (Supervised Learning) за разграничаване на двата подтипа НДБРД	28
4.12.1.1. Logistic regression	28
4.12.1.2. Decision tree	29
4.12.1.3. Random forest (RF)	29
4.12.1.4. Метод на опорните вектори (Support Vector Machines; SVM)	30
4.12.2. Оценка на моделите	31
4.13. Резултати от неконтролирано самообучение (Unsupervised Learning) за разграничаване на морфологичните подтипове при АК и ПК	32
4.13.1. Йерархичен клъстерен анализ	32
4.14. Резултати от проведеното таргетно секвениране от ново поколение	35
4.15. Връзка между нивата на експресия на микроРНКите и иРНКите и наличието на мутации	36
5.1. Обсъждане на резултатите от микроРНК експресионното профилиране чрез микрочипове в туморни и нормални белодробни тъкани	37
5.2. Валидация на избраните микроРНК чрез RT-qPCR	39
5.3. Обсъждане на резултатите от корелационно-базираните мрежи и йерархичния клъстерен анализ	40
5.4. Обсъждане на резултатите от проведения онтологичен анализ и анализ на сигналните пътища	41
5.4.1. Поддържане на клетъчната пролиферация	42
5.4.2. Потискане на супресори на растежа	42

5.4.3. Инвазия и метастазиране.....	43
5.4.4. Геномна нестабилност.....	44
5.4.5. Потискане на клетъчната смърт. ....	44
5.4.6. Индуциране/достъп до кръвоносни съдове. ....	44
5.4.7. Препрограмиране на клетъчния метаболизъм. ....	45
5.4.8. Избягване на имунния отговор. ....	45
5.4.9. Немутационни промени водещи до трансформация на клетките. ....	45
5.5. Обсъждане на резултатите от RT-qPCR за мРНКите и асоциация с клинично-патологичните характеристики.....	47
5.6. Обсъждане на резултатите от проведеното машинно самообучение .....	50
5.7. Обсъждане на резултатите от проведеното секвениране от следващо поколение. ....	51
5.7.1. Мутации в <i>TP53</i> гена. ....	53
5.7.2. Мутации в <i>KRAS</i> гена. ....	56
5.7.3. Мутации в <i>STK11</i> гена. ....	56
6. ИЗВОДИ: .....	58
7. ПРИНОСИ: .....	60
8. ПУБЛИКАЦИИ, УЧАСТИЯ В НАУЧНИ ПРОЯВИ И ПРОЕКТИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД. ....	61
8.1. Публикации. ....	61
8.2. Участия в научни прояви: .....	61
8.3. Участие в проекти:.....	63
9. Използвана литература.....	64

## 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)  
туморната 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 (missence;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)

## 1. 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-дребно клетъчен белодробен карцином. От тях бяха избрани 14 специфични микроРНК с променена експресия, които имат отношение към туморогенезата. От проведенния 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-дребно клетъчен 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 молекулярни пътища и специфични регулаторни механизми би дало насоки и за търсенето на нови по-ефикасни лекарствени средства.

## **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.

### **3. МАТЕРИАЛИ И МЕТОДИ**

#### **3.1. Материали**

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. 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  $-80^{\circ}\text{C}$  до момента на изолиране на РНК и ДНК. Допълнително се отделяше материал от туморната част за имунохистологично изследване. На всички включени участници им беше предоставена информация за извършените изследвания и участието им бе потвърдено чрез подписване на информирано съгласие.

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. 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.

#### **3.2. Методи**

За постигането на поставените задачи бяха използвани следните молекулярно – генетични техники и методи.

- Изолиране на тотална РНК и ДНК от тъкан.
- Оценка на качеството и концентрацията на изолирана тотална РНК/ДНК чрез NanoDrop 2000, Qubit v2 и Agilent 2200 TapeStation Systems
- Обратна транскрипция на изолираната РНК за синтез на кДНК.
- микроРНК експресионно профилиране чрез микрочипове на туморна и нормална белодробна тъкан
- Детекция на изследваните микроРНКи и иРНКи посредством PCR в реално време (RT-qPCR).
- Секвениране от ново поколение, с използването на TSACP и секвенционна платформа MiSeq (Illumina).
- Биоинформатичен и статистически анализ да получените резултати
- Машинно самообучение

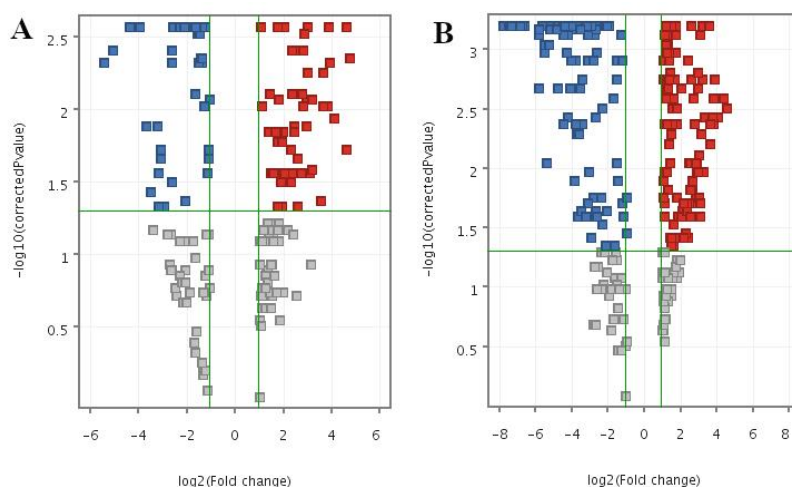


## 4. РЕЗУЛТАТИ

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

Бяха пуснати шест микрочипа – три с тумори от плоскоклетъчен белодробен карцином (12 тумора) и прилежащите нормални тъкани (12 норми) и три с тумори от аденокарцином (12 тумора) на белия дроб и прилежащите нормални тъкани (12 норми). Преди анализирането на резултатите бе извършена нормализация (90th percentile shift), контрол на качеството на резултатите (РСА) и филтриране. 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.

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 р стойност  $\leq 0.05$ . На Фигура 4.1 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. 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.



Фигура 4.1. Volcano plot, визуализиращ нивата на експресия на микроРНК-ите и статистическата значимост на получените резултати. А) аденокарцином В) плоскоклетъчен белодробен карцином

### 4.2. Резултати от йерархичен клъстерен анализ.

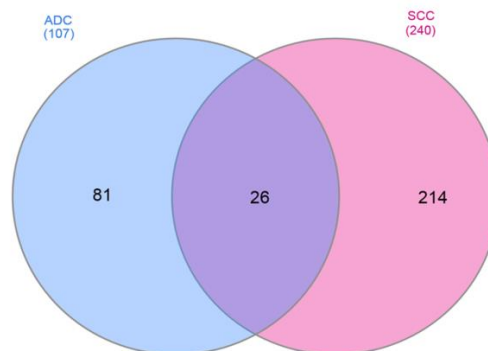
За да определим дали експресионните профили на туморните и нормалните тъкани в двете изследвани групи се разграничават, проведохме йерархичен клъстерен анализ, като анализът спомогна при подбора на микроРНК-и за последваща валидация в по-голяма извадка. При АК клъстерния анализ установи, че има добро разделяне на туморните и нормалните тъкани като се обособяват два отделни клъстера (групи). Това се дължи основно на микроРНК-ите с повишена експресия. При АК пациентите се разделят основно на 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 и пациентите с ранен туморен стадии, добре диференцирани или умерено диференцирани тумори, без нодални метастази и ацинарна, папилярна или смесена ацинарно-папилярна морфология.

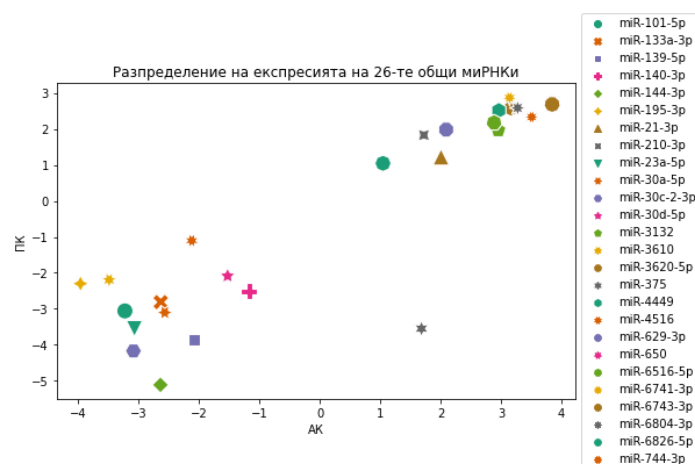
При ПК също наблюдавахме добро разделяне на туморните и нормалните тъкани на две групи. Първата група са пациентите в напреднал стадии, нискодиференцирани диференцирани тумори, с наличие на нодални и далечни метастази, които се подкълъстерираха според морфологията. Втората група са пациентите с ранен туморен стадии, добре диференцирани или умерено диференцирани тумори, без нодални метастази, като при един от тях не се наблюдава кератинизация.

### 4.3. Избор на микроРНКи за валидация.

На фигура 4.2 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 4.3.



Фигура 4.2. Вен диаграми с броя на общите и различните микроРНКи при АК и ПК (3)



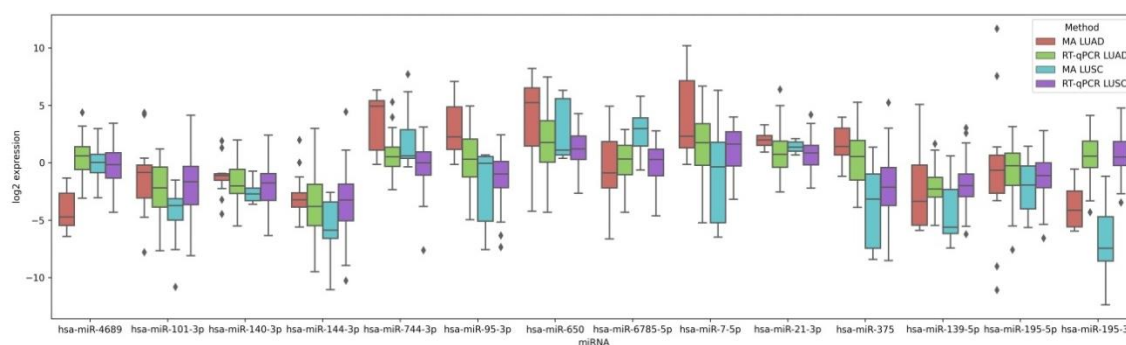
Фигура 4.3. Скатър графика, показваща разпределението на експресията на общите микроРНКи при двата подтипа белодробен карцином.

От 26-те общи микроРНКи за двата подтипа белодробен карцином бяха избрани 8 микроРНКи - 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.

#### 4.4. Валидация на избраните микроРНК чрез RT-qPCR.

Валидацията на 14-те избрани от микрочиповия анализ микроРНК 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 4.4.



Фигура 4.4. Сравнение на нивата на експресия на избраните за валидация 14 микроРНК, 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 x-оста са 14 микроРНК, а по y-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, проведохме 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 микроРНК - 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 микроРНК 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 микроРНК - 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 микроРНК 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 ( $p = 0.0001$ ).

#### 4.5. Нива на експресия на 14-те изследвани микроРНК в тъканни проби от рак на белия дроб.

Нивата на експресия на подбраните за валидация 14 микроРНК 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 4.1 и Таблица 4.2 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 4.1. 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, $\chi^2$
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 4.2. 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, $\chi^2$
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

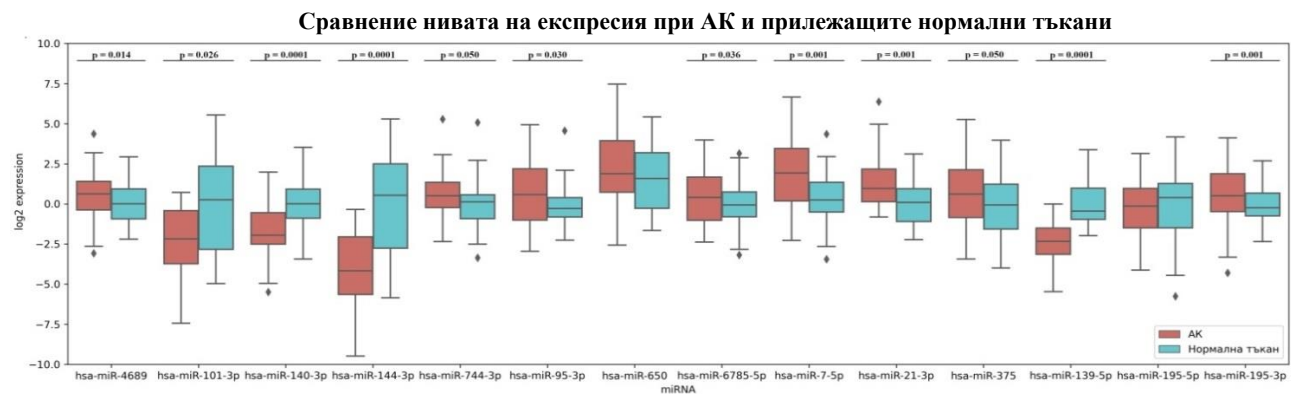
За да се определи, кои от изследваните микроРНКи са статистически значимо променени в туморна белодробна тъкан в сравнение с нормална белодробна тъкан, в зависимост от разпределението на данните, се проведе съответния параметричен или непараметричен статистически тест (Independent T-тест или Mann-Whitney 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 4.5.

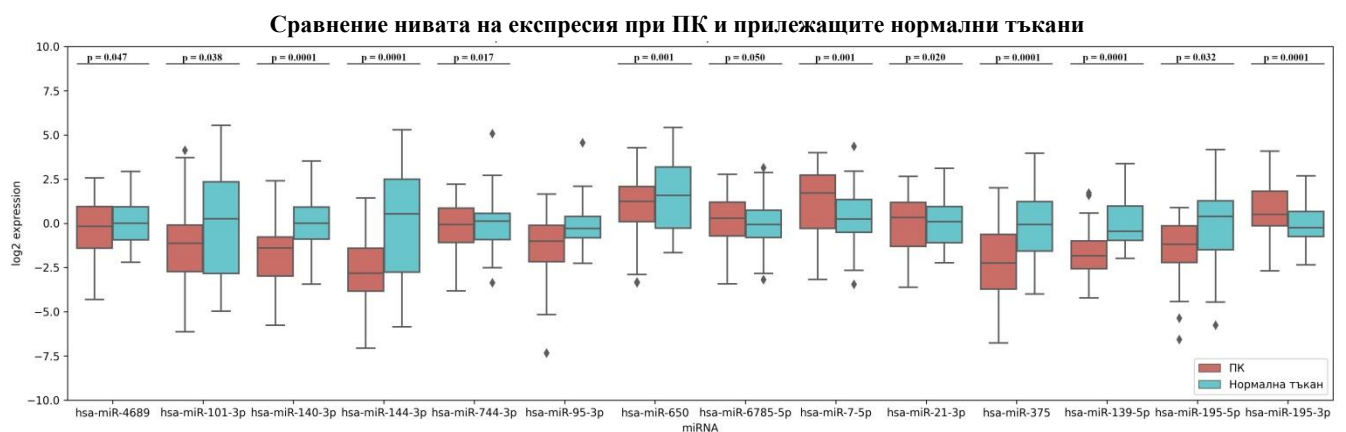
При ПК само за miR-95-3p ( $p = 0.382$ ) не се наблюдава статистически значима разлика в експресионните нива в туморите спрямо нормалните тъкани. Всички останали изследвани микроРНКи при ПК бяха със значимо променена експресия в

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 4.6.

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 4.7.

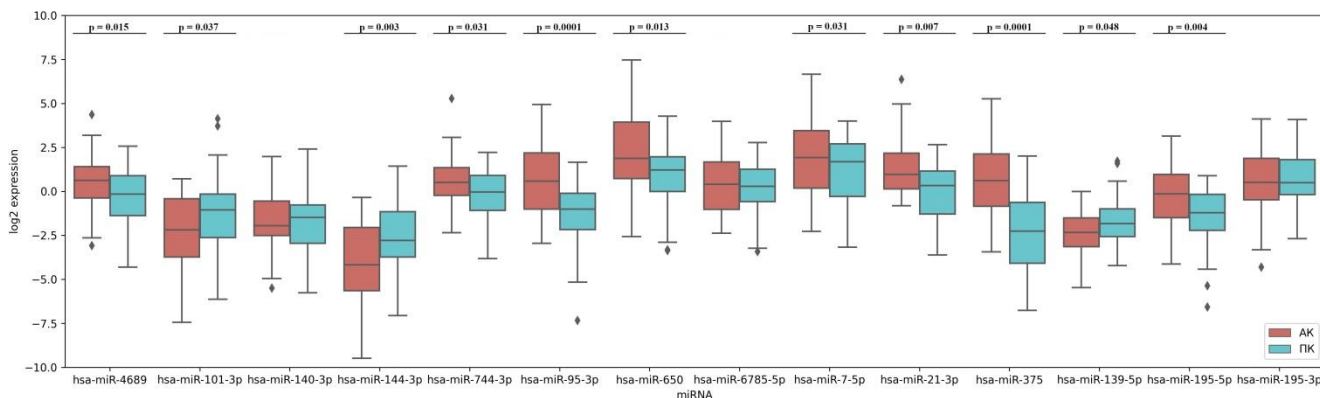


Фигура 4.5. Сравнение нивата на експресия при АК и прилежащите нормални тъкани. По x-оста са микроРНКите, по y-оста  $\log_2$  от експресионните нива. Беше използван Independent T-тест или Mann-Whitney unpaired тест, според разпределението на данните;  $p \geq 0.05$  е незначимо



Фигура 4. 6. 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 4.7. 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 test, според разпределението на данните;  $p \geq 0.05$  е незначимо;

### 4.6. Асоциация между експресионните нива на микроРНКите и клиникопатологичните характеристики на пациентите.

При АК получените RQ 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 АК 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).

При ПК всички микроРНКи показаха 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.

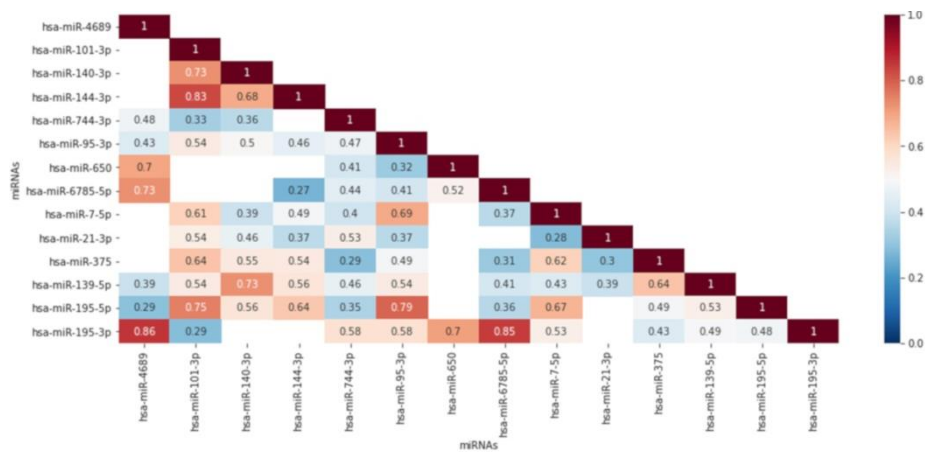
#### **4.7. Корелационен анализ между изследваните микроРНКи.**

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 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.

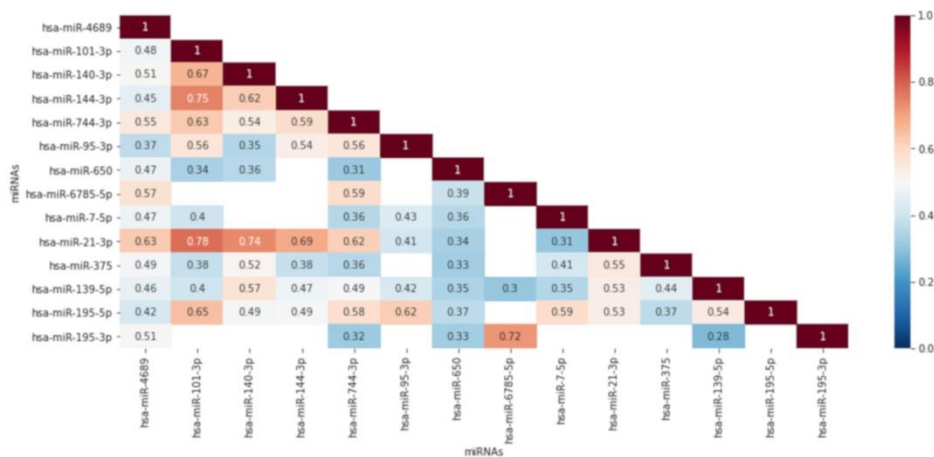
На следващата Фигура 4.8 и Фигура 4.9 са представени съответно матрица с корелационните коефициенти на всички възможни комбинации между микроРНКите. Наблюдава се само положителни корелации. В червената скала са отбелязани силните положителните корелации, докато в синята скала са отбелязани по-слабите 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 построяването на корелационно – базирани мрежи. Мрежите са представени на Фигура 4.10. за АК и Фигура 4.11. за ПК.

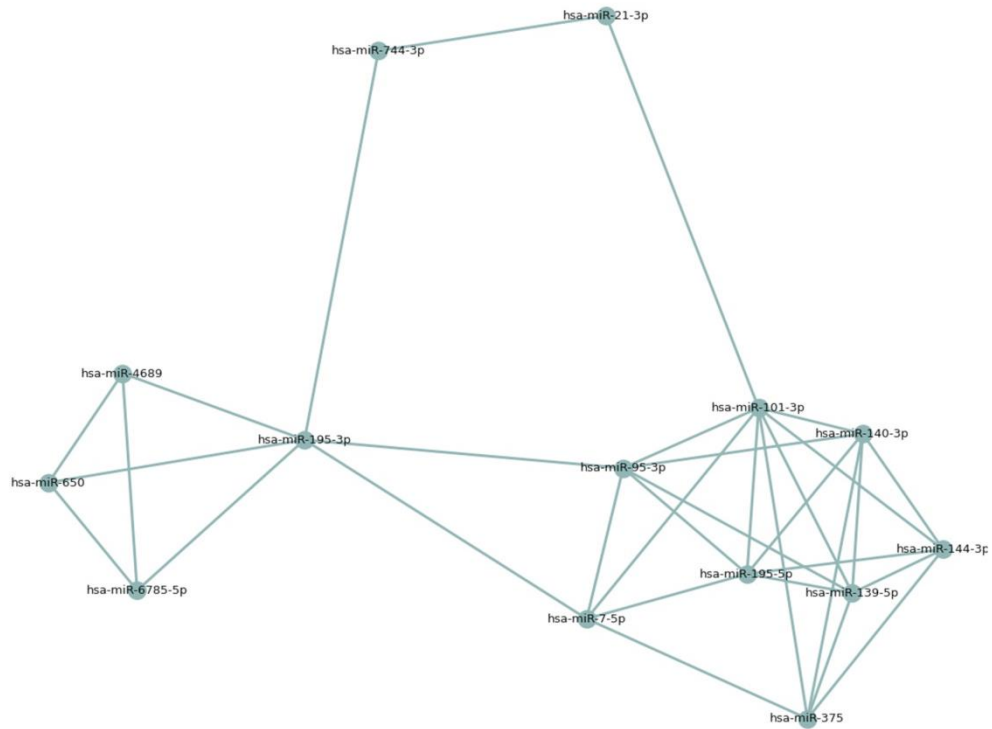




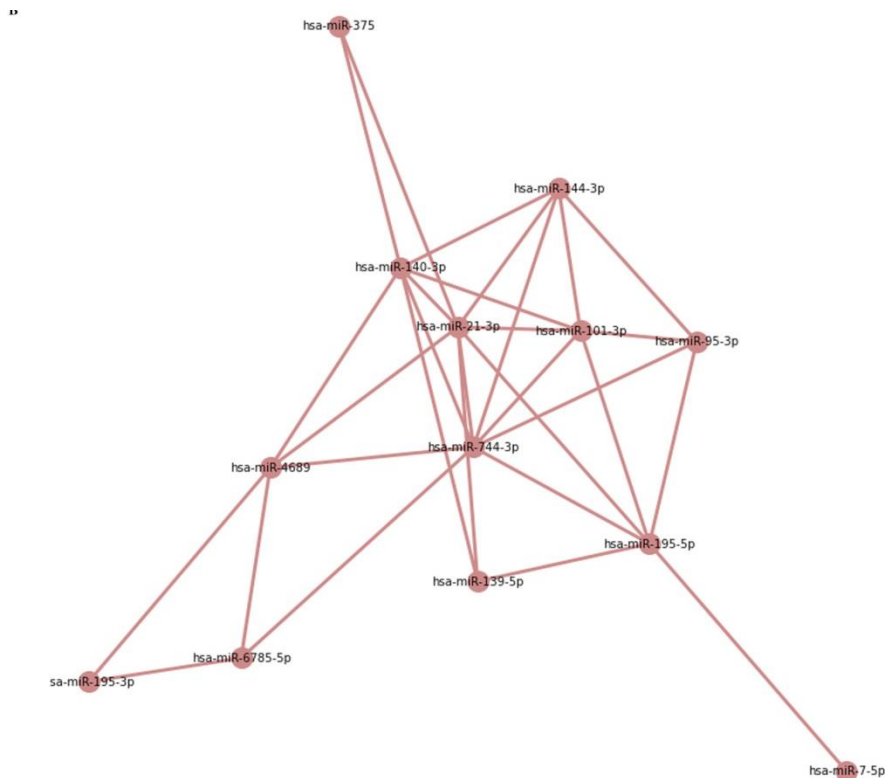
Фигура 4.8. Матрица с корелационните коефициенти на възможните комбинации между микроРНКи при АК. Визуализирани са само статистически значимите корелации с  $p \leq 0.05$ , без приложена корекция.



Фигура 4.9. 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 4.10. Корелационно – базирана мрежа със статистически значимите корелационни връзки между маркерите при АК. С добавена корекция (BH-FDR) и коефициент на корелация  $\geq 0.500$ .



Фигура 4.11. Корелационно – базирана мрежа със статистически значимите корелационните връзки между маркерите при ПК. С добавена корекция (BH-FDR) и коефициент на корелация  $\geq 0.500$ .

Ако две микроРНКите взаимодействат помежду си в мрежа е по-вероятно те да регулират пътищата и целеви гени с подобни функции (4–6). 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 (7). Характеристиките на получените мрежи ни позволява да идентифицираме най-важните възли в мрежата по отношение на броя на взаимодействията и оценим всеки възел в нея.

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 клъстера, свързани чрез miR-21-3p, miR-744-3p и miR-195-3p. В образуването на по-малкия Клъстер 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 Клъстер 1, с miR-744-3p и с микроРНКите miR-7-5p и miR-95-3p, които са част от Клъстер 2. Останалите микроРНКи от Клъстер 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 4.3 са представени характеристиките на мрежата при АК.

Таблица 4. 3. Характеристики на мрежата от микроРНКи при АК.

микроРНК	Брой връзки	Степен на централност	Коефициентът на клъстериране
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 клъстериране 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 4.4 са представени характеристиките на мрежата при ПК.

Таблица 4. 3. Характеристики на мрежата от микроРНКи при ПК.

микроРНК	Брой връзки	Степен на централност	Коефициентът на клъстериране
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

#### 4.8. Анализ на гените – мишени на микроРНКите.

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

#### 4.9. Онтологичен анализ.

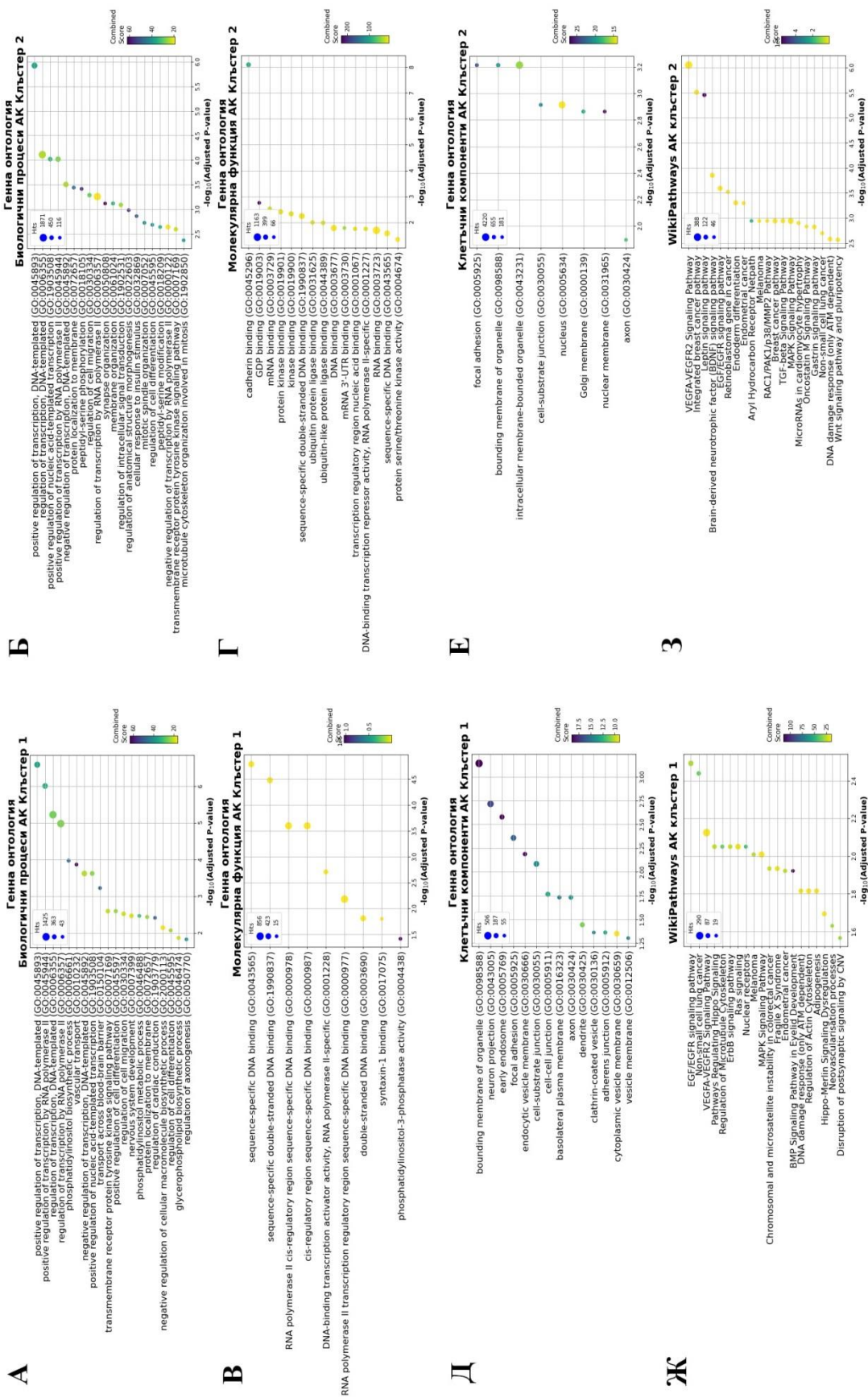
За да определим разпределението на генните-мишени проведохме онтологичен анализ и анализ на най-засегнатите сигнални пътища. При АК извършихме двата анализа за Клъстер 1 и Клъстер 2 като използвахме валидираните и предсказаните генни-мишени. За ПК в анализите участваха валидираните и предсказаните генни-мишени 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 онтологичния 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 4.12А,В,Д). Според WikiPathway 2021 тези гени, след корекция на р-стойността, обогатяват значимо 30 пътя. На Фигура 4.12Ж са представени двайсетте най-засегнати сигнални пътя. В тях обогатени са 910 генни-мишени.

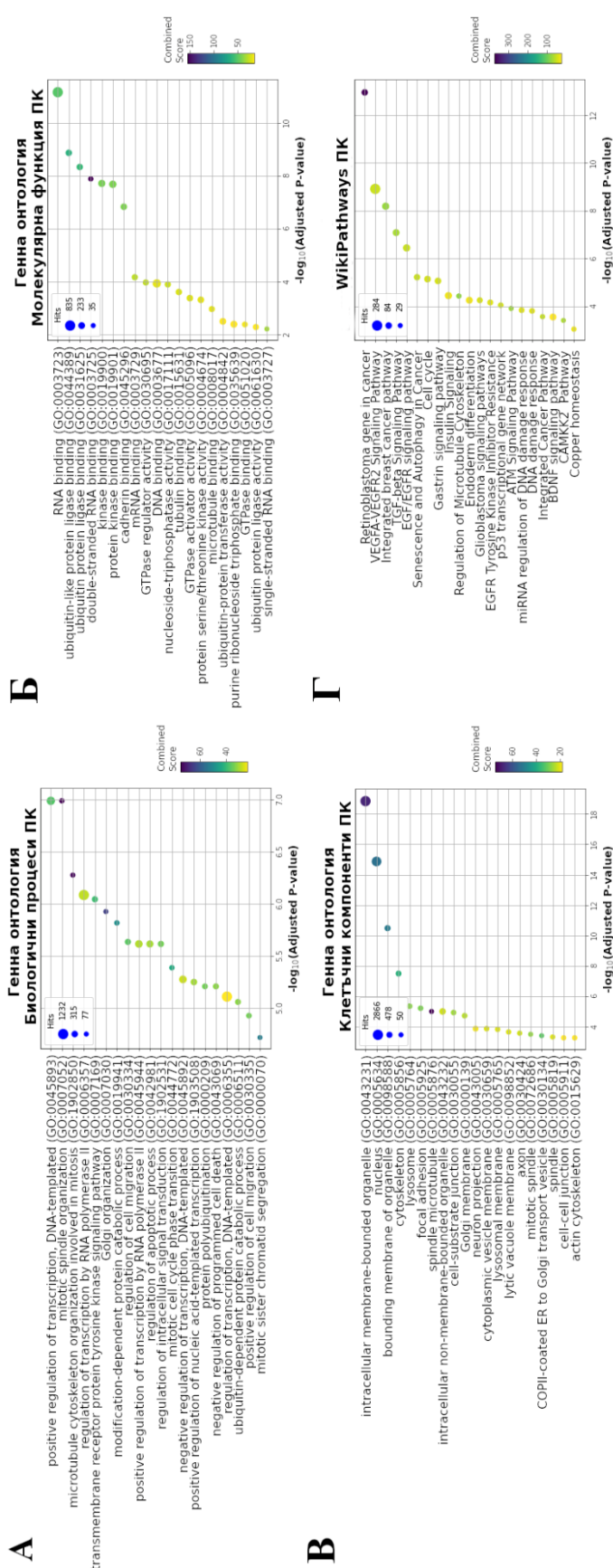
При АК Клъстер 2 валидирани и предсказани генни-мишени на осемте микроРНКи са 15808 гена. Според WikiPathway 2021, след корекция на р-стойността, тези гени обогатяват значимо 93 пътя. На Фигура 4.12З са представени двайсетте най-засегнати сигнални пътя. В тях обогатени са 1231 генни-мишени. Според онтологичния

анализ мишените на изследваните микроРНК 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 мРНК(Фигура 4.12Б,Г,Е).

При ПК валидирани и предсказани генни-мишени на четирите микроРНК са 9889 гена. Според онгологичния анализ генните-мишени на четирите микроРНК 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 мРНК, клетъчния транспорт, биват убиквитин лигази или се свързват с пуринов нуклеотид (Фигура 4.13А,Б,В). Според WikiPathway 2021, след корекция на р-стойността, тези гени обогатяват значително 57 пътя. На Фигура 4.13Г са представени двайсетте най-засегнати сигнални пътя. В тях обогатени са 915 генни-мишени.



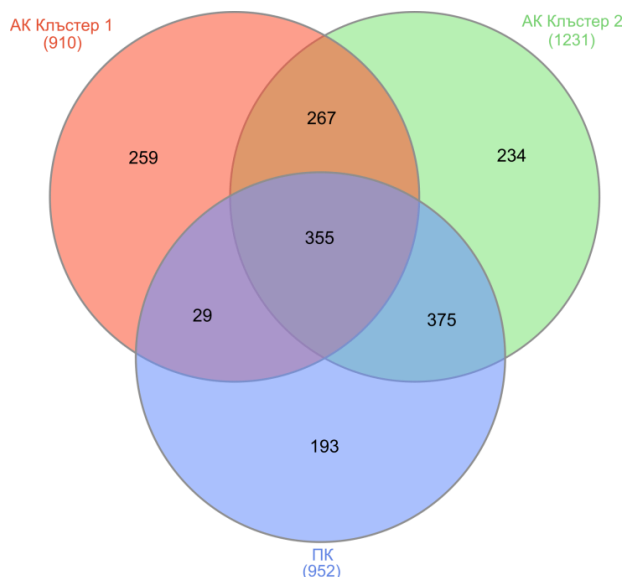
Фигура 4.12. Онтологичен анализ и анализ на сигналните пътища при АК. А, В, Д и Ж резултати за Клъстер 1; Б, Г, Е, З резултати за Клъстер 2



Фигура 4.13. Онтологичен анализ и анализ на сигналните пътища при ПК.

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

На фигура 4.14 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 4.14 Вен диаграма на общите и специфичните гени за двата клъстера на АК и ПК (3).

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 НДБРД. 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 50 пациента с АК и 50 пациента с ПК.

#### 4.11. РЕЗУЛТАТИ ОТ ЕКСПРЕСИОННИЯ АНАЛИЗ на *PD-L1*, *EGFR*, *STAT3*, *PTEN*, *MINA53* и *mTOR* при НДБРД.

Нивата на експресия на подбраните шест иРНКи бяха изследвани и анализирани при 50 пациента с АК и 50 пациента с ПК. На Таблица 4.4 и Таблица 4.5 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 4.4. 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, $\chi^2$
<b>PD-L1</b>	26 (52) 0.192 ± 0.027	17 (34) 1.257 ± 0.089	7 (14) 7.845 ± 2.175	0.002
<b>STAT3</b>	13 (26) 0.191 ± 0.043	24 (48) 1.139 ± 0.101	13 (26) 12.410 ± 6.780	0.088
<b>PTEN</b>	25 (50) 0.209 ± 0.030	14 (28) 0.989 ± 0.146	11 (22) 19.323 ± 6.911	0.058
<b>EGFR</b>	15 (30) 0.229 ± 0.037	22 (44) 0.928 ± 0.094	13 (26) 5.531 ± 0.905	0.422
<b>MINA53</b>	10 (20) 0.321 ± 0.056	19 (38) 1.286 ± 0.092	21 (42) 13.975 ± 3.239	0.048
<b>mTOR</b>	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 4.5. 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, $\chi^2$
<b>PD-L1</b>	19 (38) 0.219 ± 0.034	19 (38) 1.111 ± 0.107	12 (24) 6.293 ± 1.582	0.416
<b>STAT3</b>	22 (44) 0.271 ± 0.023	20 (40) 1.012 ± 0.094	8 (16) 3.284 ± 0.383	0.015
<b>PTEN</b>	24 (42) 0.223 ± 0.025	13 (22) 1.079 ± 0.073	13 (22) 5.193 ± 0.984	0.089
<b>EGFR</b>	14 (28) 0.166 ± 0.051	18 (36) 1.048 ± 0.112	18 (36) 5.599 ± 0.896	0.469
<b>MINA53</b>	20 (40) 0.192 ± 0.039	12 (24) 1.369 ± 0.111	18 (36) 5.763 ± 0.914	0.050
<b>mTOR</b>	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

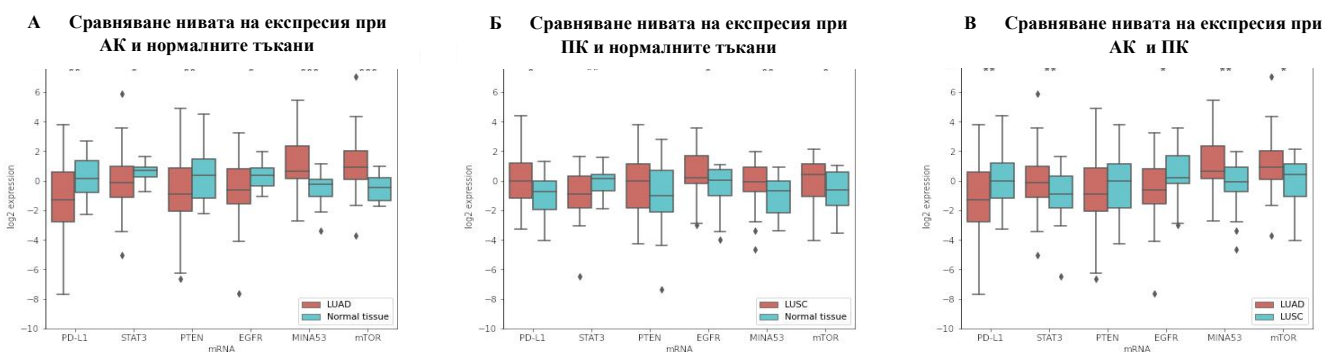
За да се определи, кои от изследваните иРНКи са статистически значимо променени в туморна белодробна тъкан в сравнение с нормална белодробна тъкан, в зависимост от разпределението на данните, се проведе съответния параметричен или непараметричен статистически тест (Independent T-тест или Mann-Whitney 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 4.15А.

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 4.15Б.

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 4.15В.

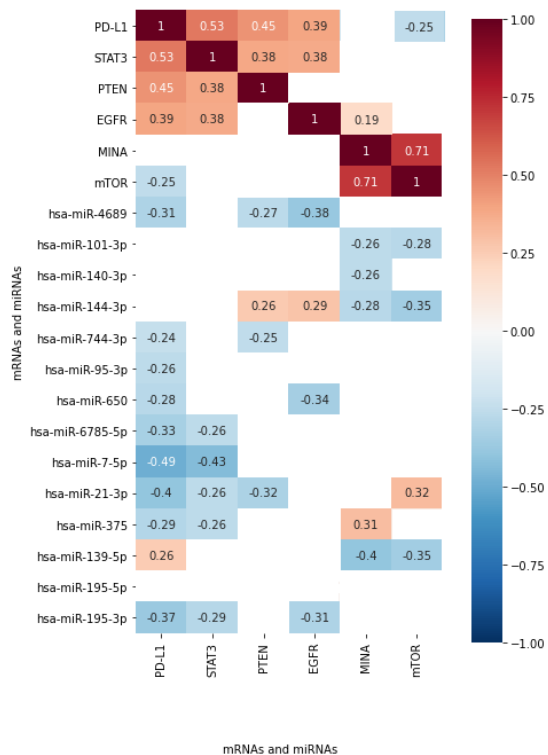


Фигура 4.15. (А) Сравнение нивата на експресия на изследваните иРНКи при АК и прилежащите нормални тъкани; (Б) Сравнение нивата на експресия изследваните иРНКи при ПК и прилежащите нормални тъкани; (В) Сравнение нивата на експресия при АК и ПК; По х-оста са иРНКите, по у-оста  $\log_2$  от експресионните нива. Беше използван Independent T-тест или Mann–Whitney unpaired тест, според разпределението на данните;  $p \geq 0.05$  е незначимо, \*:  $p \leq 0.05$ , \*\*:  $p \leq 0.01$ , \*\*\*:  $p \leq 0.001$ , LUAD – АК, LUSC – ПК

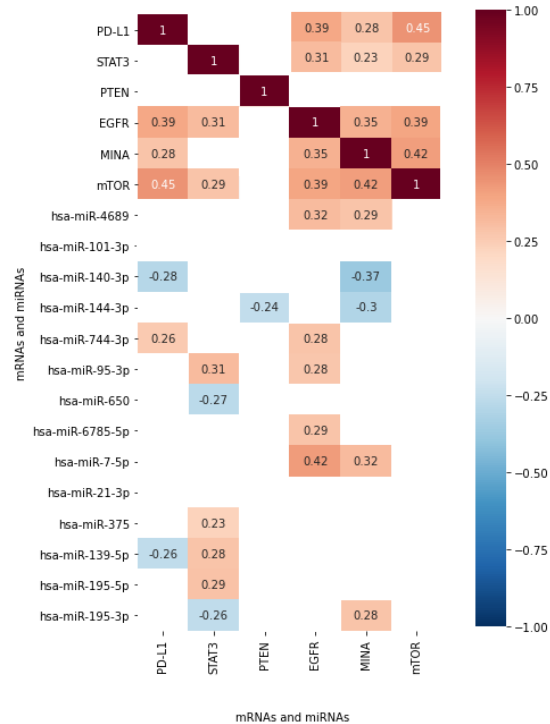
#### 4.11.1. Корелационен анализ между изследваните иРНКи и микроРНКи.

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 4.16 и Фигура 4.17 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  $p \leq 0.05$ .



Фигура 4.16. 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 4.17. 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.

#### 4.11.2. Асоциация между експресионните нива на иРНКите и кликопатологичните характеристики на пациентите

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 пониска при добре диференцираните тумори ( $p = 0.0001$ ). Нивата на *PTEN* са пониски 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.

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

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). Върху хетерогенните данни за експресията на 14 микроРНКи 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. Приложените четири вида алгоритми са:

- 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.

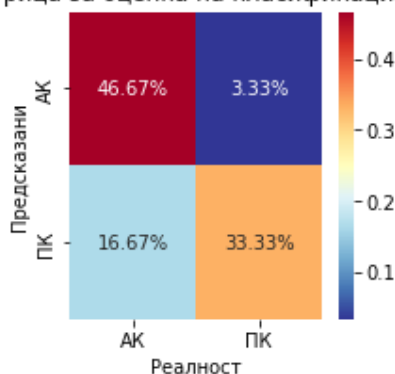
##### 4.12.1.1. Logistic regression.

Алгоритъм, при който променливата има две възможни стойности и използва само линейни функции (8). Моделът пресмята вероятността пациента да принадлежи към една от две категории (АК или ПК), в зависимост от експресията на микроРНКите.

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.

Фигура 4.18 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 4.18. 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

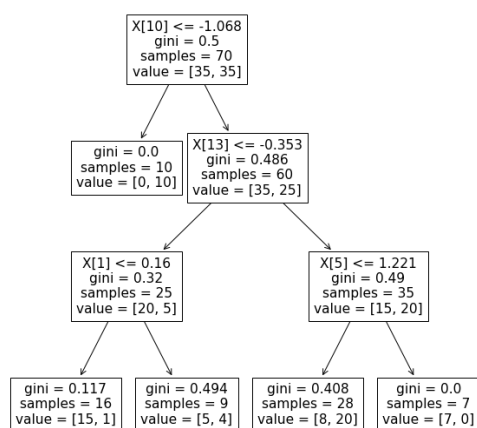
класифицирани от логистичната регресия. Неправилно класифицираните от логистичната регресия са 3.33% от АК (1 пациент) и 16.67% (5 пациента) от ПК, погрешно са класифицирани като АК.

#### 4.12.1.2. Decision tree.

Дърво за вземане на решение или Decision tree (DT) е съставено от следните елементи: корен, много възли, клонои и листа. Коренният възел прави разделянето въз основа на стойността на атрибута на класа, вътрешният възел взема атрибут за понататъшна класификация, клоните вземат решение за разделяне на възлите на листови възли, накрая, листовите възли ни дават крайния резултат (8). 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 (9). Поради тази причина е от съществено значение правилната му настройка. DT е подходящ за работа алгоритъм, когато проблемът е прост и данните са малко (10). Въпреки че алгоритмът 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 (11). Но DT е способен да картографира както линейни, така и нелинейни връзки (12).

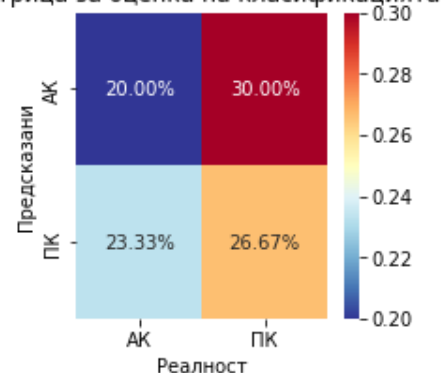
За правилната конфигурация на дърво на решенията е важно да бъдат определени точно: максимална дълбочина на дървото и минималният брой проби, които трябва да се намират в листов възел. Максимална дълбочина на дървото в случая е 3, а минималният брой проби, които трябва да се намират в листов възел са 7.

Моделът е съставен от набор от "въпроси", организирани йерархично във формата на дърво и представен схематично на фигура 4.19. На Фигура 4.20 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 4.19. Визуализация на Дърво на решенията.

Матрица за оценка на класификацията



Фигура 4.20. Матрицата за оценка на Дърво на решенията.

#### 4.12.1.3. 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) (13). RF и неговите варианти се

прилагат за различни биоинформатични проблеми, като например анализ на гена експресия, анализ на белтъци чрез мас-спектрометрия, откриване на биомаркери, аотиране на последователности и прогнозиране на белтъчно-белтъчно взаимодействия (13).

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.

Фигура 4.21 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).



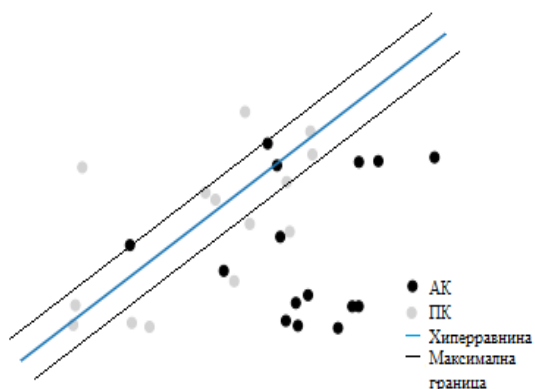
Фигура 4.21. Матрицата за оценка на случайната гора.

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

SVM алгоритмите за класификация са въведени за първи път от Vapnik (14). Те са особено подходящи и добри в случаите, когато броят на наличните данни за обучение е ограничен (15,16). 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 хиперравнина трябва да се избере по такъв начин, че да се намира възможно най-далеч от примерите и на двата класа (17). Докладвано, че те работят добре, когато данните са представени като вектори на непрекъснати променливи и има по-голяма точност в сравнение с другите алгоритми (12). 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 подтипизиране на ракови заболявания (15,18).

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.

На Фигура 4.22 е изобразена геометричната илюстрация на метода на опорните вектори. На Фигура 4.23 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 4.22. Геометрична илюстрация на SVM класификатор.



Фигура 4.23. Матрицата за оценка на SVM.

#### 4.12.2. 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 амрогитмите - точност (accuracy), прецизност (precision), чувствителност (recall), хармонична средна стойност (F1-Score), положителна прогностична стойност (PPV) и отрицателна прогностична стойност (NPV). На Таблица 4.6. са показани основните метрики за оценка на всеки от приложените модели.

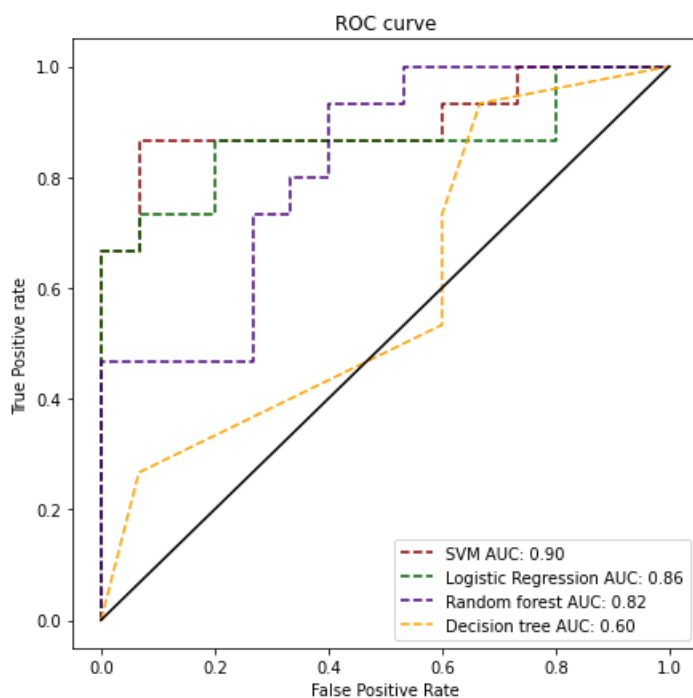
Таблица 4.6. Метрики за оценка на моделите за машино самообучение.

Алгоритъм	Точност на модела (accuracy)	Прецизност на модела (precision)	Чувствителност на модела (recall)	Средна хармонична стойност (F1-score)	PPV	NPV
Logistic regression	0.80	0.91	0.67	0.77	0.93	0.66
Decision Tree	0.47	0.47	0.53	0.50	0.40	0.73
Random forest	0.60	0.64	0.46	0.54	0.73	0.46
SVM	0.80	0.99	0.60	0.75	1.00	0.60

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 4.24 е представена ROC крива за логистичната регресия, дърво вземащо решения, случайна гора и SVM.



Фигура 4.24. ROC крива на логистичната регресия, дърво вземащи решения, случайна гора и SVM.

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

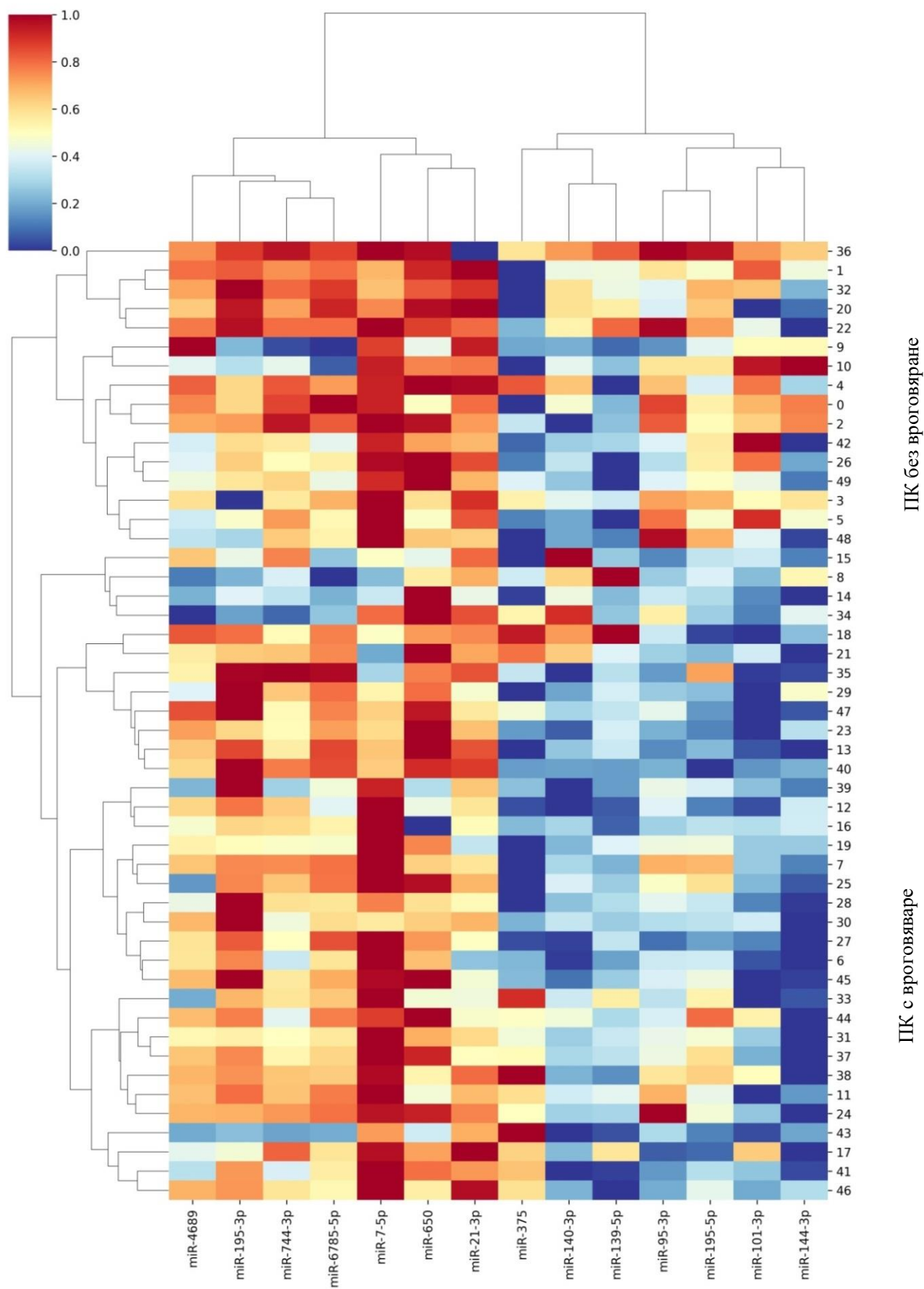
#### 4.13.1. Йерархичен клъстерен анализ.

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.

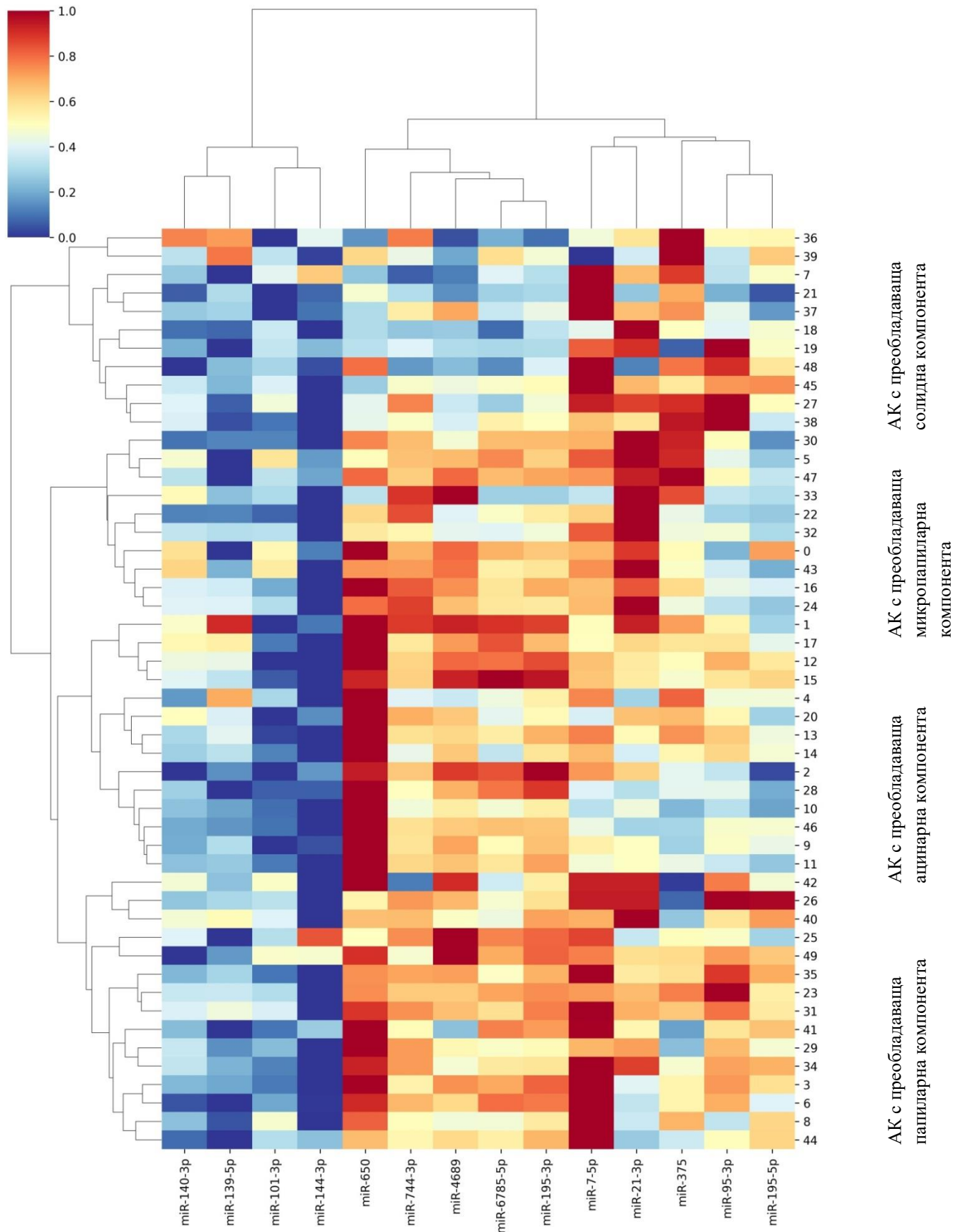
При клъстерния анализ се установи, че има добро разделяне на пациентите спрямо морфологичния тип на аденокарциномите и плоскоклетъчните карциноми (Фигури 4.25 и 4.26). При АК има добро разделяне на смесени АК с преобладаване на солидна компонента, АК с преобладаване на микропапиларна компонента, АК с 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 4.25. 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 4.26. 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.

#### 4.14. 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 раково 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 (19). 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 (missense) варианта, 9 варианта, променящи рамката на четене, и 1 вариант, засягащ сплайс място. При ПК всички наблюдавани варианти включват 33 missense 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%).

#### **4.15. 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*, *P TEN* 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. 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*, *P TEN* и *mTOR* значително по-висока при пациентите с мутация. При *STK11* положителните пациенти средното ниво на експресия на *miR-144-3p*, *miR-6785-5p*, *miR-375*, *miR-195-3p* и *P TEN* е по-високо, сравнение с негативните пациенти.

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. 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. 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.

## 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 (20), въпреки че точните механизми на регулиране на генната експресия все още не са установени (21). МикроРНКите участват в редица клетъчни процеси включително имунния отговор и туморната прогресия и могат да бъдат тумор-супресори или онкогени (22). Неотдавна беше показана ролята на микроРНКите в регулирането на специфичните профили на генната експресия при 15 типа епителни тумора от TCGA, включително РБД (23,24).

Аберантната експресия на микроРНКите се счита за важен диагностичен и прогностичен биомаркер при НДРБД (25). Все повече доказателства сочат, че различни генетични промени, включително хромозомни пренареждания, геномни амплификации, делеции, генни мутации и епигенетични промени допринасят за дерегулацията на микроРНКите (26,27). Около 80% от гените за микроРНКи са локализирани в интроните на гените, кодиращи белтъци и тяхната експресия корелира с експресията им (28,29). 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 (30).

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 (31). 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 (32). Друг подход е биоинформатичен анализ на свободно достъпните бази данни. Използвайки различни алгоритми учените идентифицират диференциално експресираните микроРНКи и след това да ги валидират (33).

В опит да се идентифицират надеждни биомаркери за диагностични и прогностични цели, са изследвани специфичните експресионни профили на микроРНКите при пациенти с НДКБК (31,34–36). Въпреки това, валидирането на характерни микроРНКи 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 (37,38). 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 (39). 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 (40–42). Докато микроРНКите, идентифицирани за разграничаване на АК и ПК от нормалната тъкан се различават значително в различните проучвания с малко припокриване (43,44).

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; фигура 1.1), отколкото с понижена (39 на брой; фигура 1.1). При ПК установихме 240 микроРНКите с диференциална експресия спрямо нормалните тъкани. Повечето от тях са с повишена експресия (144 на брой; фигура 1.1) и по малко с намалена експресия (96 на брой; фигура 1.1).

МикроРНКите с повишена експресия при АК показват промяна от 2,06 до 26 пъти. Осемте най-експресирани микроРНКи 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 (45). 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.

След нормализиране на суровите данни, филтриране и статистически анализ 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 PCA анализа.

От ранните етапи на тяхното откриване е известно, че много микроРНКите се експресират тъканно-специфично (46). 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.

## **5.2. 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 (47). Чувствителността на микрочипове е по-ниска от РНК секвенирането и RT-qPCR, тъй като при този метод няма стъпка на амплификация (48). Поради тази причина е необходимо резултатите, получени чрез този метод на скрининг за биомаркери, да се валидират чрез по-чувствителен метод (49). Един такъв метод за количествено определяне на РНК е RT-qPCR (50). Характеризира се с по-малък технически шум и по-голяма възпроизводимост от микрочипове. 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 небиологичните вариации между тестовите проби (51). . 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 микроРНК от микрочиповете изследвахме в нови 100 пациента чрез RT-qPCR.

След нормализиране на данните от микрочиповете и RT-qPCR, установихме че експресията на 8 микроРНК 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. 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.

### **5.3. Обсъждане на резултатите от корелационно-базираните мрежи и йерархичния клъстерен анализ.**

Съчетавайки сложността на иницирането, 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 микроРНК имат решаваща роля в регулирането на голямо разнообразие от биологични процеси, като клетъчна пролиферация, развитие, диференциация, апоптоза и метаболизъм (52). 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 (52,53).

При РБД са засегнати редица микроРНК, а не една и синергичното взаимодействие на множество микроРНК осигурява възможност за по-добро разбиране на сложните пост-транскрипционни механизми на регулации (54). 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 микроРНК (55–58). Тези изчислителни методи могат да бъдат разделени на три различни категории: (1) основани на нуклеотидна последователност (59–61), (2) базирани на корелация (62–64) и (3) базирани на причинно-следствена връзка (65).

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). 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 4.10 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 4.26 спомага за анализа на корелационно-базираната мрежа при АК.

Големият Клъстер 2 (Фигури 4.10 и 4.26) 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 Подкълъстер 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. За АК Кълъстер 1 основната микроРНКа е miR-195-3p, а за АК Кълъстер 2 това е miR-95-3p.

На Фигура 4.11 е представена корелационно-базираната мрежа при ПК, която не показва ясно изразено кълъстериране като при АК. Според дендрограмата на микроРНКите от йерархичния кълъстерен анализ на (Фигура 4.25), микроРНКите се разделят на 2 кълъстера, с 2 подкълъстера всеки. Първият кълъстер 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 микроРНКа молекулата, която има най-висок резултат за степента на централност. 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 кълъстер 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 4.25). 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 4.11). 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.

#### **5.4. Обсъждане на резултатите от 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 подбрани гени мишени.

Понастоящем осемте отличителни белега на раковите клетки според Hanahan и Weinberg 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 (66).

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.

#### **5.4.1. 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 (67).

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 свръхпредставени (Фигура 4.12). 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 анализиранияте белодробни аденокарциноми.

МикроРНКите от Клъстер 2 (АК с преобладаваща папиларна и солидна компонента) също таргетират иРНКи, имащи отношение за поддържане на пролиферацията (Фигура 4.12). Белтъците участват в сигналните пътища на MAPK, лептин, онкостатин М.

Четири микроРНКи при ПК, с най-голямо значение според проведения мрежови анализ, таргетират иРНКи, участващи в инсулиновия сигнален път (Фигура 4.13).

#### **5.4.2. 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 4.12). Прави впечатление, че при ПК GO анализът показва че термините са свързани с регулацията на клетъчния цикъл, включващи Rb1 и TP53 сигналните пътища (Фигура 4.13). 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 (68,69). Докато RB1 трансдуцира сигнали за инхибиране на растежа, произхождащи до голяма степен извън клетката, TP53 сигналния път се активира от вътреклетъчни сигнали (68,70). 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 (71).

### 5.4.3. Инвазия и метастазирание.

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 (72,73). При АК Клъстер 2 повечето от засегнатите белтъци са кадхерини (Фигура 4.12). Те представляват клас  $Ca^{2+}$  зависими трансмембранни гликопротеини, които медиат клетъчната адхезия и играят важна роля в поддържането на клетъчната полярност и адхезия (74).

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 метастазиранието в мозъка, клатрин-медирана ендцитоза, липиден метаболизъм и други (Фигура 4.12). 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 метастазиранието на туморните клетки (75).

GO анализът на 6508 гена на четирите микроРНКи при ПК показва, че GO термините са силно свързани с ремоделиране на клетъчния цитоскелет (Фигура 4.13). 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 (76). ПК на белия дроб се разпознава под светлинен микроскоп по наличието на кератинизация и междуклетъчни мостове (77). Кератините са специфични за епитела междинни филаменти, които имат важна роля в поддържането на структурната цялост на клетките и участват в клетъчната диференциация (76). Кератиноцитите следват уникална програма на терминална диференциация и апоптоична клетъчна смърт, което в крайна сметка води до образуването на кератиновия слой (78). Поради тези причини наличие на кератини в туморите се счита за маркер за добре диференциран плоскоклетъчен карцином на белите дробове (79).

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 4.13). Множество белтъци, участващи в ендцитозата се характеризират с аберантна експресия при НДРБД и регулират клетъчната миграция и инвазия.

#### **5.4.4. 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 (80,81)

#### **5.4.5. Потискане на клетъчната смърт.**

GO анализът на 6508 таргетни гени на четирите микроРНКи при ПК показва, че GO термините асоциират с апоптозата и автофагията (Фигура 4.13). Две важни физиологични дейности, които контролират оцеляването и смъртта на клетките и имат много важна роля по време на прогресията на белодробния рак (82). Автофагия е процес на „самоизяждане“, поглъщане на цитоплазмични белтъци, комплекси или органели в автофагозома (цитоплазмична структура с двойна мембрана). 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 (83). 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 4.13)

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 (84). 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 (85).

#### **5.4.6. 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 (86). 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 (71,87). При АК и при ПК процесът на ангиогенеза е засегнат от изследваните микроРНКи (Фигури 4.12 и 4.13).

#### **5.4.7. Препрограмиране на клетъчния метаболизъм.**

Една от основните промени, които настъпват в раковите клетки е метаболитното препрограмиране (71). 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 (88).

Проведеният онтологичен анализ показва, че липидният метаболизъм е повлиян при АК (Фигура 4.12). Липидите са хетерогенна група биомолекули, включваща фосфолипиди, мастни киселини, триглицериди и сфинголипиди и имат значителна роля в пролиферацията, инвазията и метастазирането на туморните клетки (89). 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 (89). Липидният метаболизъм също така има важна роля за туморната микросреда и имунотерапията (90). 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 имуносупресивно действие и избягване на тумора от действието на имунната система (91,92).

Убиквитинирането и фосфорилирането при злокачествени заболявания представляват двете най-важни посттранслационни белтъчни модификации (93). Все повече проучвания сочат, че убиквитинът играе критична роля в канцерогенезата и прогресията на РБД (94). Онгологичният анализ показва обогатяване на тези биологични процеси при АК и ПК (Фигури 4.12 и 4.13).

#### **5.4.8. 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 имунен 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 (95).

#### **5.4.9. 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 (66).

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 (96). 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 придобита резистентност (97). В повечето случаи е наблюдавана трансформация от АК към дребноклетъчен карцином, но само няколко проучвания са описали трансформация на ПК (98).

**Като заключение онтологичният анализ на 15808 гена-мишени на микроРНКите от АК Клъстер 2 показва, че GO 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 GO термини са силно свързани със специфични клетъчни процеси, улесняващи метастазирането в мозъка, клатрин-медирана ендоцитоза, липиден метаболизъм и др. (Фигура 4.12).**

**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 – медираните 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.**

## 5.5. 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-медиирани пътя. Измежду гените, обогатяващи се предимно в 20-те най-силно повлияни от мРНКите 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 инициирането и поддържането на карциногенезата при РБД (99). Такива аберации могат да бъдат предизвикани от гени на околната среда, активирани поради излагане на опасности вещества или канцерогени (100). Настоящото проучване изследва експресионните нива на един такъв ген - *MINA53*, в двата подтипа на НДРБД.

Нивата на експресия на PD-L1 бяха значително по-високи в ПК, отколкото при АК (Фигура 4.15В) и се асоциираха с ниско диференцираните тумори, без кератинизация и без далечни метастази. При АК по-високата експресия се наблюдава при пациентите с нодални и далечни метастази, ниско диференцираните тумори, при които преобладава солидна или микропапиларна морфологична компонента. Скорешни метаанализи показват неблагоприятната прогноза при PD-L1 положителните тумори (101).

STAT3 самостоятелно или заедно с NF-κB регулира експресията на много цитокини и хемокини (102). INFγ, кодиран от *IFNG* гена, е първият цитокин, който се секретира от макрофагите и другите тумор инфилтрирани лимфоцити. Той регулира транскрипцията на PD-L1 чрез активиране на JAK/STAT3 сигналния път (103). Така чрез тази положителна връзка, секретираният от раковата клетка цитокин 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 (104).

В изследваните от нас тумори STAT3 има значително по-висока експресия при ниско диференцираните ПК тумори. При АК се наблюдава силна положителна корелация между експресионните нива на STAT3 и PD-L1.

Генът PTEN (phosphatase and tensin homolog) е локализиран на дългото рамо на 10 хромозома (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 (105).

При изследваните ПК експресията му е значително по-ниска при кератинообразуващите карциноми и при пациентите, при които се откриват далечни метастази. 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 (106). Тези пациенти се характеризират с висок мутационен товар и хромозомни пренареждания (107).

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 (108). 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 (109).

Установено е че, по-високите нива на експресия на *mTOR* са предиктивен фактор за чувствителност към лечение с *mTOR* инхибитори при панкреасни НЕТ (110). В нашето изследване експресията е по-висока сред АК тумори на белия дроб (Фигура 4.15B), 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 и далечни метастази. *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 (111). Инхибирането на *mTOR* възстановява чувствителността на раковите клетки към цисплатина в 2D и 3D клетъчни култури и миши модели (111). 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 (112).

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).

*EGFR* е трансмембранен рецепторен белтък с тирозин киназна активност и се характеризира с повишена експресия в голям брой белодробни карциноми (113). Свързва се с туморния растеж, неопластогенеза, туморна инвазия и клетъчно безсмъртие (114). В изследваните от нас тумори е значимо по-ниска експресията му при АК сравнение с ПК (Фигура 4.15B). 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 (115,116). При 62% от НДБК *EGFR* показва повишена експресия, която се асоциира с по-лоша прогноза (117). Резултатите ни показват, че експресията на *EGFR* е значително по-висока при нискодиференцираните тумори и АК със солидна и микропапиларна компонента.



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. Известно е, че *MINA53* регулира експресията на EGFR при НДКБД (118,119). 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(120). *EGFR* участва в регулирането на експресията на PD-L1 и клетъчната пролиферация чрез сигналния път IL-6/JAK/STAT3 при НДРБД (121). **Настоящото проучване показва потенциала на комбинираната таргетна терапия с имунотерапия при лечението на АК.**

Някои фактори на околната среда като тютюнев дим, силициев диоксид и арсен, индуцират експресията на *MINA53*, чрез JNK-STAT3 сигналния път (122). Наблюдавахме положителна корелация между нивата на експресия на *STAT3* и *PD-L1* с *MINA53*. 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.

Корелационния анализ при двата подтипа НДРБД показва корелация между експресия на *EGFR* и *MINA53*. Корелацията между двата биомаркера е наблюдавана и в изследването на Ма и сътр. (123). В клетъчните линии свръхекспресията на *MINA53* повишава нивата на *EGFR* и стимулира туморната пролиферация чрез регулиране на фазовия преход G1 към S (123).

В изследваните от нас пациенти няма статистически значима разлика в експресията на *MINA53* между АК и ПК в ранен и напреднал стадий. 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 (119). Обратно, намалените нива на *MINA53* в напреднал стадий на НДРБД индуцира инвазивен фенотип (119). 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. Нашите резултати са в съгласие с предишните проучвания, илюстриращи ролята на *MINA53* като ранно събитие по време на белодробната карциногенеза (119,124,125). 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 (125).

От друга страна според Zhang и сътр. *MINA53* също така деметилюва 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 (126). По този начин двата колектива показват различната роля на *MINA53*, според етапа на карциногенезата.

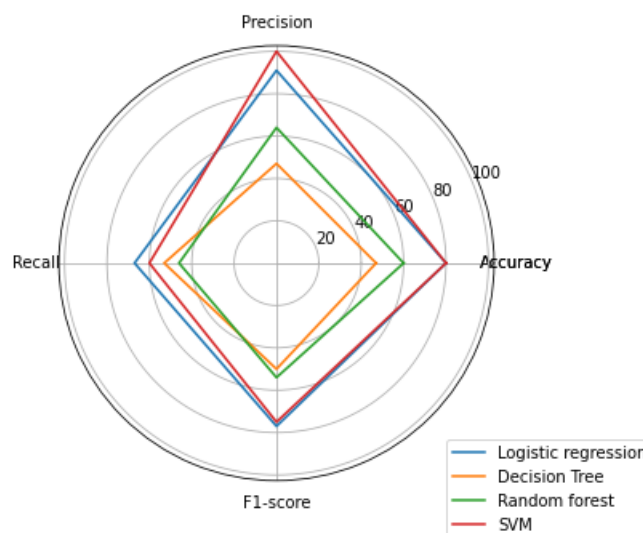
## 5.6. Обсъждане на резултатите от проведеното машинно самообучение

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 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. Хетерогеността 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 гена експресия и идентифицирането на модели, които не могат да бъдат точно уловени чрез традиционните математически подходи (127).

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 настоящото 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 5.1. 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 5.1. Сравнение на метриците за оценка на всеки от приложените модели за машинно самообучение .

Чрез методите за неконтролирано машинно самообучение, включващи йерархичния клъстерен анализ (Фигури 4.25 и 4.26) и метода на k-най-близките съседи, 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 (128).

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.**

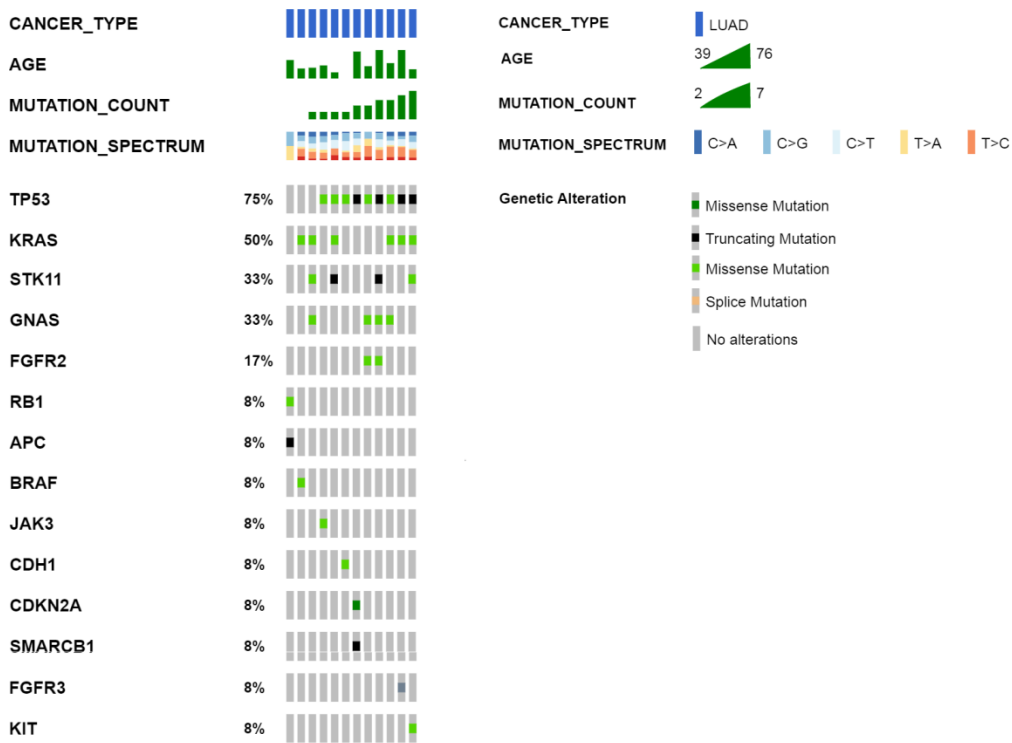
### 5.7. Обсъждане на резултатите от проведеното секвениране от следващо поколение.

Не-дребно клетъчният 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 (17).

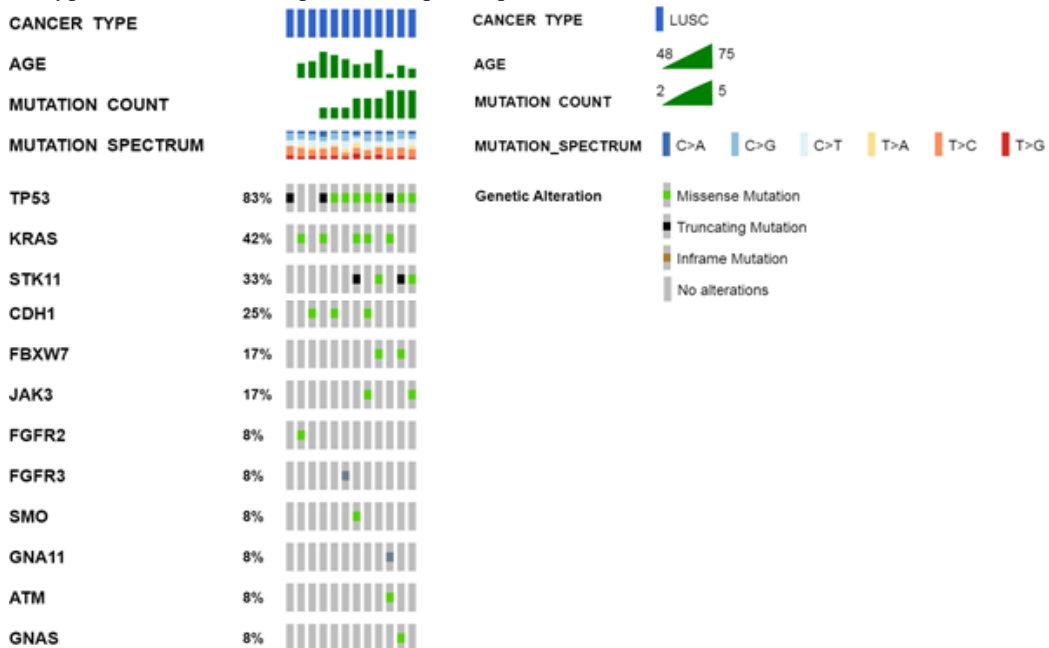
Част от поставените задачи в настоящия дисертационен труд е да се изследва соматичния спектър на мутации при РБД, събрани от 24 пациента с АК и ПК, посредством таргетно NGS секвениране, (TruSeq Amplicon-Cancer Panel) на платформа MiSeq, Illumina. Бяха избрани АК и ПК от всеки морфологичен подтип. Непосредствено 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.

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 5.2 и 5.3 са представени генетични варианти открити при АК и ПК.



Фигура 5.2. Генетични варианти открити при АК (129).

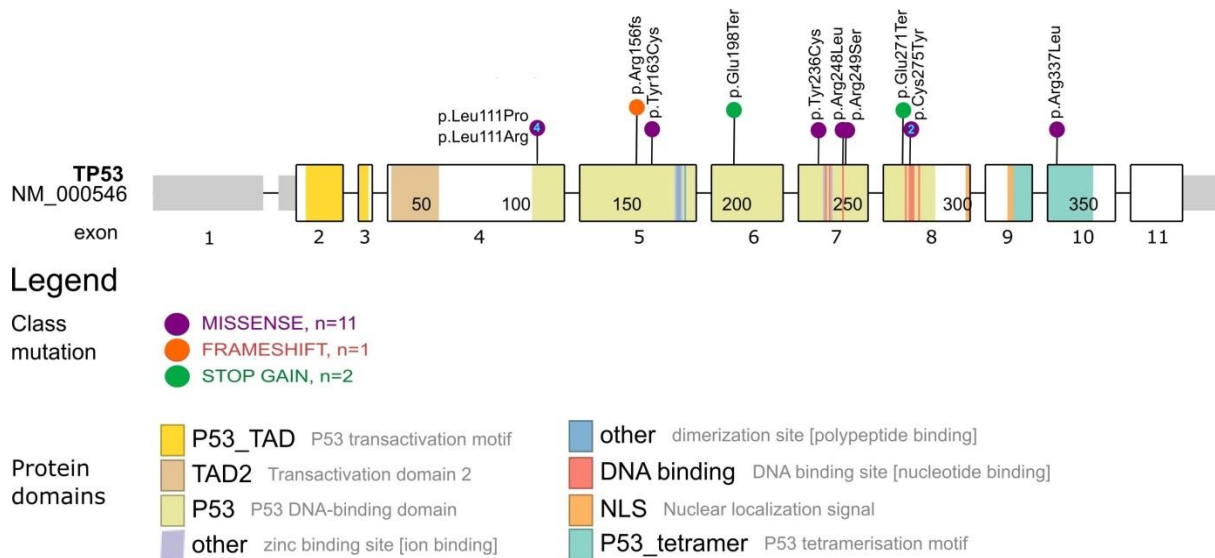


### 5.7.1. Мутации в *TP53* гена.

*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, съдържат 393 аминокиселини, четири белтъчни домена и характеризира се с високо ниво на консервативност между гръбначните животни (130). 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 *p53* 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 *p53* активността. С-терминалният домейн (АК остатъци 363-393) е включен в даун-регулацията на централния ДНК-свързващ домейн (131).

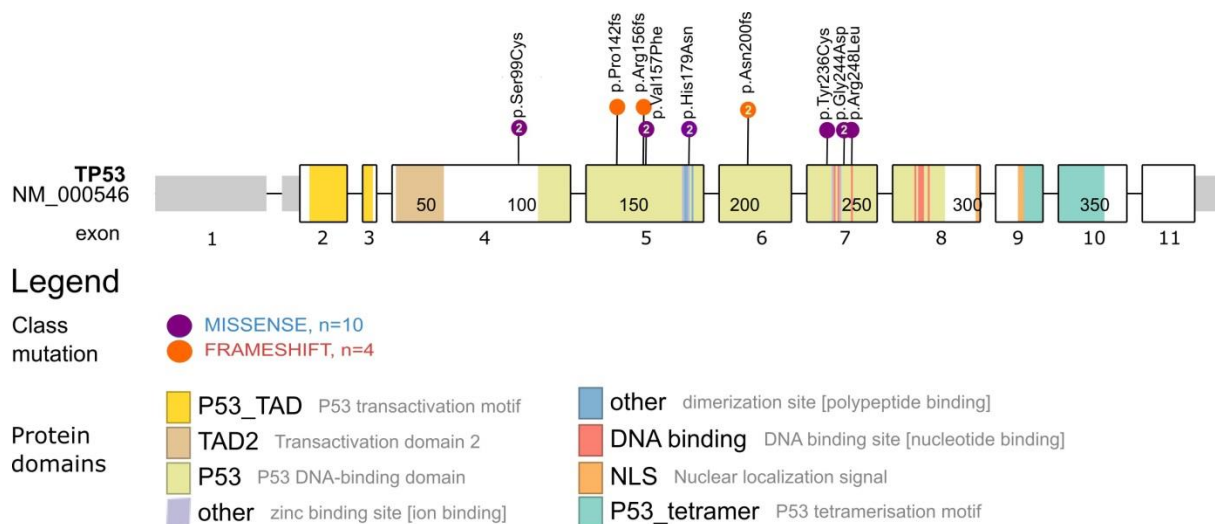
При изследваните пациенти с АК и ПК преобладават мутациите в *TP53* гена, което е в съгласие с твърдението, че *TP53* гена е един от най-често мутиралите гени при солидни тумори (132). Мутациите в *TP53* гена се появяват рано в процеса на канцерогенезата и са свързани с лоша обща преживяемост, резистентност към терапия и чести случаи на рецидиви. При нормални условия концентрацията на *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 CHK1 и CHK2, 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 (133). 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 хетерогенност 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 еднонуклеотидната 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 тютюнопушетето. Фигура 5.4. изобразява кодиращите екзони на *TP53* гена, белтъчните домени и откритите мутации.



Фигура 5.4. Кодиращите екзони на *TP53* гена, белтъчните домени и откритите мутации при ПК (134).

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 5.5 изобразява кодиращите екзони, белтъчните домени и откритите мутации на *TP53* гена.



Фигура 5.5. Кодиращите екзони на *TP53* гена, белтъчните домени и откритите мутации при АК (134).

. 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* (135). 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 (136). 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 (137). Обратно, наличието на *TP53* мутация не повлиява преживяемост без прогресия или общата преживяемост на пациентите, лекувани с 2-ро поколение EGFR-TKI (138).

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 (139). 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 замени (с изключение на стоп кодони) извън ДНК свързващия домен, които могат да запазят някои от функционалните свойства на p53 белтъка(139). Няколко поучавания показват, че недеструктивните мутации в *TP53* гена имат значима връзка със съкратена обща преживяемост, в сравнение с дивия тип (140–142).

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 (143). 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 (143). 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 (144,145). Досега не са публикувани проучвания при лечение на пациенти с EGFR-TKIs и химиотерапия при напреднал NSCLC с комутация 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 (146). Колективът открива, използвайки мултивариантна логистична регресия, че експресията на PD-L1 се асоциира с мутация в *TP53* гена (146). Освен това, Hao Sun и сътр. демонстрират, че пациенти с *TP53* миссенс мутации имат повишено ниво на PD-L1 и обогатени IFN- $\gamma$  сигнатури в сравнение с подгрупата от *TP53*-див тип (147). 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 (148). 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 (148). 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.

### 5.7.2. 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 клетки функции (149,150). 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.

Онкогенните мутации на гена *KRAS* възникват най-вече в екзон 2 в кодон 12, по-рядко в кодон 13 (3–5%) и още по-рядко в екзон 3 кодон 61 (по-малко от 1%) (150). 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 (151). Поради високата му честота в АК са проведени няколко предклинични и клинични изследвания, търсещи ефективни терапевтични подходи и насочени към мутацията в *KRAS* гена (152). 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 непущачи, докато мутациите G12C и G12V се срещат по-често сред бивши или настоящи пушачи (153). 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 (154).

Неотдавнашен ретроспективен анализ на пациенти с напреднал НДРБД установи, че мутацията в *KRAS* е предиктор за намалена обща преживяемост, при лечение с цитотоксична химиотерапия (155). Освен това съвместното детектиране на мутация в *TP53* гена предсказва по-лоша прогноза (156). Повечето публикувани данни, включително мета-анализ на 22 проучвания, предполагат, че *KRAS* мутационният статус е отрицателен предиктор за EGFR-ТКИ (157). 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* (158). Въпреки това, все още съществуват противоречиви резултати и не всички проучвания са стигнали до същите заключения (159–162). 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 (163).

### 5.7.3. 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 (164–167). 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 (168). 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 повлияят на експресията на гените, влияещи на имунното разпознаване и редица интерферони (169,170). 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 (171,172). 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 (167,173).

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 молекулярните 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 T-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 M-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 T-статуса показва – 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 M-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 показва 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 – медиирани 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, трансляцията, организацията на клетъчната мембрана, апарата на Голджи, включително регулация на клетъчната диференциация и морфогенеза. След засегнатите иРНКи преобладават кадхерини, протеин кинази, РНК-свързващи се белтъци, убиквитин лигази и транскрипционни фактори.
- При ПК корелационно-базираните мрежи от микроРНКи 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 иРНКи преобладават РНК-свързващите се белтъци, кадхерините, протеин кинази и транскрипционни фактори.

- На база ко-експресията на 14-те микроРНКи 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 микроРНКи 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 микроРНКа 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 недребноклетъчен 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 подтипове 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 подтипове 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 отговора 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. Kyurkchian, 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. Kyurkchian, 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. Използвана литература

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. 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).
4. 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.
5. 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.
6. Deng S, Qi J, Stephen M, Qiu L, Yang H. Network-based identification of reliable biomarkers for cancers. *J Theor Biol*. 2015;383:20–7.
7. 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.
8. 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.
9. 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.
10. 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.
11. Mingers J. An Empirical Comparison of Pruning Methods for Decision Tree Induction. *Mach Learn* 1989 42. 1989;4(2):227–43.
12. Jayatilake SMDAC, Ganegoda GU. Involvement of Machine Learning Tools in Healthcare Decision Making. *J Healthc Eng*. 2021;2021.
13. Qi Y. Random Forest for Bioinformatics. *Ensemble Machine Learning*. Springer US; 2012. p. 307–23.
14. Cortes C, Vapnik V, Saitta L. Support-vector networks. *Mach Learn* 1995 203. 1995 Sep;20(3):273–97.
15. 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.
16. 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.
17. Cai Z, Poulos RC, Liu J, Zhong Q. Machine learning for multi-omics data integration in cancer. *iScience*. 2022;25(2):103798.
18. 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.
19. 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.
20. 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.
  21. 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.
  22. Esquela-Kerscher A, Slack FJ. Oncomirs — microRNAs with a role in cancer. *Nat Rev Cancer* 2006 64. 2006 Apr;6(4):259–69.
  23. Peng Y, Croce CM. The role of microRNAs in human cancer. Vol. 1, *Signal Transduction and Targeted Therapy*. Springer Nature; 2016.
  24. 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.
  25. 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.
  26. Iorio M V, Croce CM. Causes and consequences of microRNA dysregulation. *Cancer J*. 2012;18(3):215–22.
  27. Di Leva G, Garofalo M, Croce CM. MicroRNAs in cancer. *Annu Rev Pathol*. 2013/09/25. 2014;9:287–314.
  28. Kim Y-K, Kim VN. Processing of intronic microRNAs. *EMBO J*. 2007/01/25. 2007 Feb 7;26(3):775–83.
  29. 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.
  30. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*. 2004;429(6990):457–63.
  31. 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.
  32. 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.
  33. 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.
  34. 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.
  35. 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.
  36. 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.
  37. 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.
  38. 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.
39. 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.
  40. 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.
  41. 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.
  42. 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.
  43. 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.
  44. 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.
  45. 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.
  46. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116(2):281–97.
  47. Quackenbush J. Microarray data normalization and transformation. *Nat Genet.* 2002;32(S4):496–501.
  48. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res.* 1996;6(10):986–94.
  49. Jaksik R, Iwanaszko M, Rzeszowska-Wolny J, Kimmel M. Microarray experiments and factors which affect their reliability. *Biol Direct.* 2015 Sep 3;10:46.
  50. 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.
  51. 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.
  52. Hill M, Tran N. MicroRNAs Regulating MicroRNAs in Cancer. *Trends in Cancer.* 2018;4(7):465–8.
  53. Hill M, Tran N. miRNA interplay: Mechanisms and consequences in cancer. *DMM Dis Model Mech.* 2021;14(4):1–24.
  54. 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.
  55. 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.
  56. 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.
  57. 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.
58. 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.
  59. 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.
  60. 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.
  61. 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.
  62. Boross G, Orosz K, Farkas JJ. Human microRNAs co-silence in well-separated groups and have different predicted essentialities. *Bioinformatics*. 2009;25(8):1063–9.
  63. 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.
  64. Sahu M, Mallick B. Deciphering synergistic regulatory networks of microRNAs in hESCs and fibroblasts. *Int J Biol Macromol*. 2018;113:1279–86.
  65. 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.
  66. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov*. 2022 Jan 1;12(1):31–46.
  67. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr LA, Kinzler KW. Cancer genome landscapes. *Science*. 2013 Mar 29;339(6127):1546–58.
  68. Burkhart DL, Sage J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer*. 2008 Sep;8(9):671–82.
  69. Deshpande A, Sicinski P, Hinds PW. Cyclins and cdks in development and cancer: a perspective. *Oncogene*. 2005;24(17):2909–15.
  70. Sherr CJ, McCormick F. The RB and p53 pathways in cancer. *Cancer Cell*. 2002;2(2):103–12.
  71. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 2011 Mar 4;144(5):646–74.
  72. 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.
  73. 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.
  74. Charalabopoulos K, Gogali A, Kostoula OK, Constantopoulos SH. Cadherin superfamily of adhesion molecules in primary lung cancer. *Exp Oncol*. 2004;26(4):256–60.
  75. 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.
  76. Aseervatham J. Cytoskeletal Remodeling in Cancer. *Biology (Basel)*. 2020 Nov 1;9(11):1–40.
  77. Suarez E, Knollmann-Ritschel BEC. Squamous Cell Carcinoma of the Lung. *Acad Pathol*. 2017;4.
  78. Eckert RL, Rorke EA. Molecular biology of keratinocyte differentiation. *Environ Health Perspect*. 1989;80:109.

79. 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.
80. 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.
81. Moynahan ME, Jasin M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol*. 2010 Mar;11(3):196–207.
82. 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.
83. 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.
84. Martinon F, Tschopp J. Inflammatory Caspases. *Cell*. 2004;117(5):561–74.
85. 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.
86. Hanahan D, Folkman J. Patterns and Emerging Mechanisms of the Angiogenic Switch during Tumorigenesis. *Cell*. 1996;86(3):353–64.
87. Fouad YA, Aanei C. Revisiting the hallmarks of cancer. *Am J Cancer Res*. 2017;7(5):1016.
88. Mendes C, Serpa J. Metabolic Remodelling: An Accomplice for New Therapeutic Strategies to Fight Lung Cancer. *Antioxidants*. 2019 Dec 1;8(12).
89. 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.
90. 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.
91. 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.
92. Xiang Y, Miao H. Lipid Metabolism in Tumor-Associated Macrophages. *Adv Exp Med Biol*. 2021;1316:87–101.
93. Wang Y, Wang F. Post-Translational Modifications of Deubiquitinating Enzymes: Expanding the Ubiquitin Code. *Front Pharmacol*. 2021 Jun 10;12:685011.
94. 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.
95. 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.
96. 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.
97. 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.
98. 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.
99. 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.
100. 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.
  101. 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.
  102. Igelmann S, Neubauer HA, Ferbeyre G. STAT3 and STAT5 activation in solid cancers. *Cancers (Basel)*. 2019;11(10).
  103. Mandai M, Hamanishi J, Abiko K, Matsumura N, Baba T, Konishi I. Dual Faces of IFN $\gamma$  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.
  104. Hu R, Han Q, Zhang J. STAT3: A key signaling molecule for converting cold to hot tumors. *Cancer Lett*. 2020;489(44):29–40.
  105. 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.
  106. 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.
  107. 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.
  108. 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.
  109. Moon YW, Hajjar J, Hwu P, Naing A. Targeting the indoleamine 2,3-dioxygenase pathway in cancer. *J Immunother Cancer*. 2015;3.
  110. 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.
  111. Liang S-Q, Bührer 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.
  112. 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.
  113. Nicholson RI, Gee JMW, Harper ME. EGFR and cancer prognosis. *Eur J Cancer*. 2001;37(SUPPL. 4):9.
  114. 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.
  115. 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.
  116. 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;
  117. Wee P, Wang Z. Epidermal growth factor receptor cell proliferation signaling pathways. *Cancers (Basel)*. 2017;9(5):1–45.
  118. 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.
119. 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.
  120. 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.
  121. 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.
  122. 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.
  123. 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.
  124. 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.
  125. 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 $\beta$ / $\beta$ -catenin signaling. *Int J Oncol*. 2017;51(6):1898–908.
  126. 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.
  127. Artificial Intelligence: A Modern Approach, 4th US ed. [Internet]. [cited 2022 Oct 30]. Available from: <http://aima.cs.berkeley.edu/>
  128. 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.
  129. 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).
  130. 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.
  131. 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.
  132. 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.
  133. 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.
  134. 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.
  135. Bykov VJN, Eriksson SE, Bianchi J, Wiman KG. Targeting mutant p53 for efficient cancer therapy. *Nat Rev Cancer*. 2017;18(2):89–102.

136. 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.
137. 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.
138. 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.
139. 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.
140. 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.
141. 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.
142. 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.
143. 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.
144. 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.
145. 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.
146. 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.
147. 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.
148. 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.
149. 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.
150. Tímár J. The clinical relevance of KRAS gene mutation in non-small-cell lung cancer. *Curr Opin Oncol*. 2014;26(2):138–44.
151. O’Bryan JP. Pharmacological targeting of RAS: Recent success with direct inhibitors. *Pharmacol Res*. 2018/10/23. 2019 Jan;139:503–11.
152. Ghimessy A, Radeckzy 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.
153. 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.
  154. 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.
  155. 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.
  156. 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.
  157. 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.
  158. 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.
  159. 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.
  160. 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.
  161. Garassino MC, Martelli O, Brogгинi 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.
  162. 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.
  163. 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.
  164. 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.
  165. Johnson JL, Pillai S, Chellappan SP. Genetic and biochemical alterations in non-small cell lung cancer. *Biochem Res Int*. 2012;2012.
  166. 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.



167. 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.
168. Shackelford DB, Shaw RJ. The LKB1-AMPK pathway: metabolism and growth control in tumour suppression. *Nat Rev Cancer*. 2009 Aug;9(8):563–75.
169. 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.
170. 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.
171. Koyama S, Akbay EA, Li YY, Aref AR, Skoulidis F, Herter-Sprrie 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.
172. 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.
173. 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.