МЕДИЦИНСКИ УНИВЕРСИТЕТ – СОФИЯ
МЕДИЦИНСКИ ФАКУЛТЕТ
КАТЕДРА ПО МЕДИЦИНСКА ГЕНЕТИКА
РЪКОВОДИТЕЛ: ЧЛ. КОР. ПРОФ. Д-Р ДРАГА ТОНЧЕВА, Д-Р, ДБН

ЗОРА АКРАМ ХАМУДЕ

ОПРЕДЕЛЯНЕ НА ВАРИАНТИ, СВЪРЗАНИ С КАРЦИНОГЕНЕЗА И ФАРМАКОГЕНЕТИЧНИ ЕФЕКТИ ПРИ РАК НА БЯЛ ДРОБ И ЩИТОВИДНА ЖЛЕЗА

А В Т О Р Е Ф Е Р А Т

на дисертационен труд за присъждане на образователна и научна степен “ДОКТОР”
Област на висше образование: „Природни науки, математика и информатика ”,
Шифър 4.3.
Професионално направление: „Биологически науки ”
Докторска програма: „Генетика”
НАУЧНИ РЪКОВОДИТЕЛИ: ЧЛ. КОР. ПРОФ. Д-Р ДРАГА ТОНЧЕВА, ДБН,
ДОЦ. Д-Р САВИНА ХАДЖИДЕКОВА, ДМ

НАУЧЕН КОНСУЛТАНТ: ДОКТОР РАДА СТАНЕВА, ДБ

НАУЧНО ЖУРИ В СЪСТАВ:
ПРЕДСЕДАТЕЛ: Чл.-кор. Проф. Д-р Драга Иванова Тончева, дбн - становище
ЧЛЕНОВЕ: Проф. Д-р Стоян Ганчев Лалчев, дм – рецензент
Проф. Д-р Вили Кръстева Стоянова, дм – рецензент
Проф. Мариела Константинова Оджакова, дб – становище
Доц. Светослав Георгиев Димов, дб - становище

София, 2018 г.

Изследванията са извършени в:

Изследванията са извършени в Катедра по Медицинска генетика, Медицински Факултет, Медицински Университет – София, СБАЛГАР „Д-р Малинов“ и Ажибадем Сити Клиник Болница Токуда.

Дисертацията съдържа 125 стандартни страници и включва 30 фигури, 17 таблици. Оформена е в 6 глави, заключение, общи изводи, приноси, използвана литература и декларация за оригиналност. Библиографията обхваща 228 литературни източника.

Докторантът е задочен докторант към Катедра по Медицинска генетика, Медицински Факултет, Медицински Университет – София.

Дисертационният труд е обсъден на разширен Катедрен съвет в Катедрата по Медицинска генетика към Медицински Факултет, Медицински Университет – София на 16.07.2018 г. и е насочен за защита пред Научно жури.

Материалите по защитата са на разположение в Катедрата по Медицинска генетика, 6-ти етаж на СБАЛАГ „Майчин дом”, ул. „Здраве”2, София, 1431.

Защитата ще се състои на 03.12.2018 год. от 11.00 часа в семинарната зала на 14 етаж на СБАЛАГ „Майчин дом” на основание Заповед № РК 36 – 2403 / 08.10.2018г.

Автор: Зора Акрам Хамуде

Заглавие: Определяне на варианти, свързани с карциногенеза и фармакогенетични ефекти при рак на бял дроб и щитовидна жлеза.

Област на висше образование: „Природни науки, математика и информатика ”, Шифър 4.3.

Професионално направление: „Биологически науки”

Докторска програма: „Генетика”

Съдържание

Въведение 6

I. Цел и задачи 7

II. Материали и методи 8

1. Пациенти 8

2. Биологични проби 8

3. Използвани методи 8

3.1. Изолиране на ДНК 8

3.2. ДНК секвениране от ново поколение 10

3.3. Полимеразна верижна реакция в реално време (Real-time PCR) 23

III. Резултати 28

1. Данни от NGS секвениране с карциномен панел на туморни проби от НДКБК 28

1.1 Анализ на варианти свързани с туморогенеза при НДКБК 28

1.2. Фармакогенетични варианти при НДКБК, определени чрез цеквениране с карциномен панел 31

2. Мутационен спектър на *EGFR* гена определен чрез експресионен PCR анализ в реално време за избор на таргетна терапия на НДКБК 37

2.1. Клинико-патологична характеристика на изследваните пациенти 37

2.2. Мутации в *EGFR* гена за таргетна терапия на НДКБК 39

2.3. Анализ на фенотип/генотип корелации 41

3. Данни от NGS секвениране с карциномен панел на ДНК-и от кръвни проби на пациенти с фамилен папиларен тироиден карцином 43

3.1. Родословия на четири фамилии с ФПТК 43

3.3. Фармакогенетични варианти при ФПТК, определени чрез секвениране с карциномен панел 46

IV. Обсъждане 50

1. Обсъждане на патогенни варианти в гени, асоциирани с карциногенеза при НДКБК 50

2. Обсъждане на фармакогенетични варианти в гени за туморогенеза при НДКБК за избор на химиотерапия 54

3. Обсъждане на мутационния спектър на *EGFR* гена за избор на таргетна терапия при НДКБК 63

4. Обсъждане на патогенни варианти в гени, асоциирани с карциногенеза при ФПТК 66

5. Обсъждане на фармакогенетични варианти в гени за туморогенеза при ФПТК за избор на химиотерапия 69

V. Заключение 75

VI. Изводи 77

VII. Приноси 78

VIII. Научни публикации и съобщения по темата на дисертационния труд 79

IX. Благодарности 80
X. Summary 81

Използвани в текста съкращения

<u>Използвани в текста съкращения на кирилица</u>:

СГХ - микрочипова сравнителна геномна хибридизация
иРНК - информационна РНК
НДКБК - недребноклетъчен белодробен карцином
АТФ – аденозинтрифосфат
ТК - тироиден карцином
ПТК - папиларен тироиден карцином
ФПТК – фамилен папиларен тироиден карцином
РМЖ – рак на млечната жлеза
НЛР – нежелани лекарствени реакции

<u>Съкращения на латиница</u>:

DNMT - DNA methyltransferase
MGMT - O6-methylguanine DNA methyltransferase
miRNA - microRNA (ribonucleic acid)
CTCs – circulating tumor cells (Циркулиращи туморни клетки)
FISH - fluorescent in situ hybridization
CISH - chromogenic in situ hybridization
SISH - silver-enhanced in situ hybridization
Real-time PCR – Real-time Polymerase Chain Reaction
PCR – Polymerase Chain Reaction
NGS - Next Generation Sequencing
DNA - Deoxyribonucleic acid
Array-based CGH - Microarray based comparative genomic hybridization (микрочипова сравнителна геномна хибридизация)
ASO - antisense oligonucleotide
ODNs – oligodeoxynucleotides
mRNA - Messenger RNA
RNAi – RNA interference (РНК интерференцията)
siRNA - Small interfering RNA (малки интерферентни РНКи)
dsRNA – double-stranded RNA (двойно-верижна РНК)
CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats
gRNA - Guide RNAs
PAM - Protospacer Adjacent Motif
DSB - double-strand break
NHEJ - non-homologous end-joining
HDR - homology-directed DNA repair
HPV - Human papillomavirus
NSCLC - Non-small-cell lung cancer
EGFR – Epidermal growth factor receptor
EGFR TKIs – EGFR Tyrosine kinase inhibitors
ALK - Anaplastic lymphoma kinase (Анапластична лимфомна киназа)
NER - Nucleotide excision repair

Забележка: Номерата на таблиците и фигурите не съответстват на номерата в дисертационния труд.

Въведение

Настоящият дисертационен труд насочва вниманието към значим здравен проблем – търсенето на нови подходи при лечението на пациенти, страдащи от ракови заболявания. Раковите заболявания са широко разпространени, социално значими и заемат важна роля в човешката патология. Според Световната Здравна Организация (СЗО) туморните заболявания са втората причина за смъртност по света. Ефикасността и резултатите от лечението на злокачествените заболявания зависят от различни фактори. Известно е, че ефикасността на едно и също лекарство, прилагано в терапевтична доза по стандартен протокол, е различна между пациентите. Това може да се дължи на различни причини генетични фактори, физиологични/патофизиологични процеси - фактори на околната среда. Варианти в някои гени, могат да обуславят различия между индивидите в скоростта на лекарствената резорбция, транспорт, свързване с клетъчни рецептори, биотрансформация, образуване на активни субстанции и отстраняване от организма. Фармакогенетиката и фармакогеномиката изучават ролята на генетичните варианти за променения отговор на индивида (повишен, намален или липсващ) към действието на лекарства. Появата на секвенирането от ново поколение (NGS – next generation sequencing) съществено промени пейзажа на геномната медицина. NGS технологията има изключителен потенциал за приложение във фармакогеномиката, тъй като позволява анализ на целия геном, екзом или панел от голям брой гени в една реакция бързо, точно и евтино. Приложението на NGS за целите на фармакогеномиката в последните години преобрази областта на персонализираната медицина.

В научната разработка са анализирани два вида рак - недребноклетъчен белодробен карцином (НДКРБ) и папиларен тиреоиден карцином (ПТК), като са приложени два съвременни молекулярно-генетични подхода – секвениране от ново поколение (NGS – next generation sequencing) и PCR (polymerase chain reaction) в реално време.

Данните от NGS анализа бяха използвани за определяне на варианти, свързани с лекарствената чувствителност и нежелани лекарствени реакции (НЛР).

Данните от PCR в реално време бяха анализирани за детектиране на типове мутации в EGFR гена в туморни ДНК проби от пациенти с НДКБК за определяна на таргетна терапия.

Недребноклетъчният рак на белия дроб е най – често срещания вид белодробен карцином. В етиологията на НДКРБ стоят различни рискови фактори като тютюнопушене, фактори на околнатаne среда и генетични предразположеност. В зависимост от стадия на тумора и хистологичния тип, пациентите с белодробен карцином се насочват за различни видове терапии (конвенционална химиотерапия или прицелна терапия). За по-добър ефект от лечението при пациентите се търсят специфични биомаркери, напр.

мутации, които създават чувствителност в туморние клетки към определени медикаменти.

Папиларният карцином на щитовидната жлеза е най-честата форма на злокачествено заболяване на щитовидната жлеза, което обикновено има добра прогноза и представлява приблизително 80-85% от всички тиреоидни карциноми. Въпреки това, този вид рак може да причини далечни метастази и да бъде по-агресивен при възрастни пациенти. Изборът на терапия въз основа на определянето на биомаркери се очертава като ефикасен подход при лечението на заболяването.

Фармакогеномиката предлага нов подход за повишаване на ефекта от терапията и намаляване на страничните ефекти на лекарствата. Повишаването на ефекта от химиотерапията чрез предварително геномно тестване на пациентите и индивидуализиране на лечението ще подобри прогнозата за онкоболните и ще намали токсичните реакции.

I. Цел и задачи

Цел

Целта на дисертационния труд е да се характеризират генетични варианти при два типа карциноми – недребноклетъчен белодробен карцином (НДКБК) и фамилен папиларен тироиден карцином (ФПТК) с TruSight Cancer panel и да се определят фармакогенетични дефекти, свързани с ефективността на химиотерапията и прицелната терапия с тирозин киназни инхибитори.

Задачи

1. Да се секвенират с TruSight Cancer Sequence panel туморни ДНК проби от пациенти с НДКБК ;

2. Да се характеризира типа, честотата и клиничното значение на генетичните варианти при НДКБК и да се определят фармакогенетични варианти, свързани с ефективността на химиотерапията чрез използване на базата данните «PharmGKB»;

3. Да се секвенират с TruSight Cancer Sequence panel ДНК проби, изолирани от периферна кръв на пациенти с ФПТК;

4. Да се характеризира типа, честотата и клиничното значение на генетичните варианти при ФПТК и да се определят фармакогенетични варианти, свързани с ефективността на химиотерапията чрез използване на базата данните «PharmGKB»;

5. Да се характеризират с PCR в реално време типовете мутации в EGFR гена в туморни ДНК проби от пациенти с НДКБК за избор на подходяща таргетна терапия.

II. Материали и методи

1. Пациенти

В изследването са участвали общо 581 пациенти. От тях 569 пациенти са диагностицирани с НДКБК през последните пет години (от януари 2012 г. до октомври 2016 г.) и 12 пациенти от четири семейства с фамилна форма на папиларен карцином на щитовидната жлеза. Всички пациенти предоставиха подписано писмено информирано съгласие за извършване на генетичен анализ на техните проби.

2. Биологични проби

За извличане на геномна ДНК от пациентите с НДКБК са взети туморни проби, включени в парафинови блокчета, фиксирани с формалин (FFPE). Туморната проба е изследвана от патолог за откриване на райони, богати на туморни клетки повече от 30%. ДНК екстракцията е извършена чрез QIAamp DNA FFPE Tissue Kit за пречистване на геномно ДНК. Концентрациите на ДНК пробите е измерена с NanoDrop Lite спектрофотометър система.

При 12 пациента с папиларен карцином на щитовидната жлеза геномната ДНК е изолирана от кръвни проби с QIAGEN QIAamp DNA Blood Mini Kit. ДНК концентрациите е измерена на Qubit 2.0 Fluorometric Система с kit Qubit dsDNA BR (Thermo Fisher Scientific, Waltham, MA, САЩ). За по-нататъшен анализ са използвани проби със съотношение 260/280 от ~ 1.8, което показва достатъчна чистота на изолираната ДНК.

3. Използвани методи

В настоящата работа са използвани следните молекулно генетични методи – екстракция на геномна ДНК, ДНК секвениране от ново поколение (NGS) и полимеразна верижна реакция в реално време (Real-time PCR).

3.1. Изолиране на ДНК

Изолиране на ДНК от кръв

За изолиране на ДНК от кръв, беше използван кита на QIAGEN QIAamp DNA Blood Mini Kit, а за изолиране на ДНК от туморна тъкан в FFPE беше използван QIAGEN QIAamp DNA FFPE Tissue Kit.

За процеса на изолиране на ДНК от кръв се взима 200 µl от цялата кръв и се смесва с 20 µl протиназа К и 200 µl лизис буфер AL в 1.5 ml епруветка тип епендорф. Реакцията се вортексира за 10 секунди и се инкубира за 10 минути на 56°C. След кратко центрофугиране за да се съберат капките от стените на епруветката се добавят 200 µl етанол (96-100%), следва вортексиране за 10 секунди и прехвърляне на сместа в QIAamp Mini spin колона, поставена в събирателна епруветка. Колоната се центрофугира за 1 мин на 6000 x g на стайна температура (15-25°C) и се прехвърля в нова събирателна епруветка, а старата събирателна епруветка с течността се изхвърля. Към колоната се добавят 500 µl буфер AW1,

центрофугира се за 1 мин на 6000 x g, след което отново се слага в нова събирателна епруветка, а старата събирателна епруветка с течността се изхвърля. Добавят се 500 µl буфер AW2, след което пробата се центрофугира за 3 мин на пълна скорост (20 000 x g). Елуатът се изхвърля. Следва ново центрофугиране за още 1 мин на максимална скорост за пълно изсушаване на колоната, след което колоната се слага в нова епендорф епруветка и се инкубира за 5 мин на стайна температура с елуиращ буфер AE. Следва центрофугиране за 1 мин на 6000 x g, при което се отделя пречистената ДНК. Концентрациите на ДНК бяха измервани с помощта на Qubit 2.0 флуорометрична система за количествено определяне, използвайки Qubit dsDNA BR Assay Kit.

Изолиране на ДНК от туморна тъкан в FFPE

За процеса на изолиране на ДНК от туморна тъкан в FFPE, тъканата проба се нарязва много тънко на специален апарат (микротом) с дебелина на срезовете до 2-3 микрометра, взима се 8 - 10 среза от пробата и се поставят в микроцентрофужна епруветка тип епендорф от 1.5 ml. Добавя се 1 ml ксилен към пробата, със затворени капачки на епруветките се вортексират за 10 секунди. Пробите се центрофугират за 2 мин на пълна скорост (20 000 x g) на стайна температура (15-25°C). Супернатантата се отстранява чрез пипетиране, без да се отстранява утайката. Добавя се 1 ml етанол (96-100%) към утайката и се вортексиране за 10 секунди. Следва центрофугират за 2 мин на пълна скорост (20 000 x g) на стайна температура. Отново супернатантата се отстранява чрез пипетиране, без да се отстранява утайката. Пробите се инкубират на стайна температура или до 37°C с отворени епруветки за 10 мин или докато цялото количество остатъчен етанол се изпари. Утайката се ресуспендира в 180 µl буфер ATL, добавя се 20 µl протеиназа К и се вортексират за 10 секунди. Реакцията се инкубира за 1 час на 56°C (или докато пробата е напълно лизирана). Следва още една инкубация за 1 час на 90°C. След кратко центрофугиране за да се съберат капките от стените на епруветката се добавят 200 µl лизиращ буфер (AL), следва вортексиране за 10 секунди и се добавят 200 µl етанол (96-100%), следва вортексиране за 10 секунди. Сместа се прехвърля в QIAamp MinElute колона, поставена в събирателна епруветка. Колоната се центрофугира за 1 мин на 6000 x g и се прехвърля в нова събирателна епруветка, а старата събирателна епруветка с течността се изхвърля. Към колоната се добавят 500 µl буфер AW1, центрофугира се за 1 мин на 6000 x g, след което отново се слага в нова събирателна епруветка, а старата събирателна епруветка с течността се изхвърля. Добавят се 500 µl буфер AW2, след което пробата се центрофугира за 1 мин на 6000 x g. Елуатът се изхвърля. Следва центрофугиране за още 3 мин на максимална скорост (20000 x g) за пълно изсушаване на колоната, след което колоната се слага в нова епендорф епруветка и се инкубира за 5 мин на стайна температура с елуиращ буфер ATE. Следва центрофугиране за 1 мин на максимална скорост (20 000 x g), при което се отделя пречистената ДНК. Концентрацията на ДНК, както и качеството бяха измервани с NanoDrop 2000c (ThermoScientific) след изолирането.

3.2. ДНК секвениране от ново поколение

Общо 30 проби (от 12 пациента с папиларен карцином на щитовидната жлеза и 18 пациенти с НДКРБД) бяха секвениране с TruSight Cancer панел на MiSeq секвенатор на Illumina. За ДНК секвениране е използван протокол на фирмата. Използван е секвенционен панел - TruSight Cancer на Illumina.

TruSight Cancer (Illumina) панелът включва 94 гена за диагностика на мутации, свързани с предразположеност към различни видове ракови заболявания. Списък на гените в панела е показан в Таблица 1. В панела също така са включени и 284 SNPs, за които големи проучвания са показали, че са асоциирани с рак. Броят на таргетните екзони е >1700 и общият таргетен район на панела е 255 Kb.

Таблица 1: Списък на гените включени в TruSight Cancer панел

Таблица 1: Списък на гените включени в TruSight Cancer панел				
AIP	DDB2	FANCL	PALB2	SLX4
ALK	DICER1	FANCM	PHOX2B	SMAD4
APC	DIS3L2	FH	PMS1	SMARCB1
ATM	EGFR	FLCN	PMS2	STK11
BAP1	EPCAM	GATA2	PRF1	SUFU
BLM	ERCC2	GPC3	PRKAR1A	TMEM127
BMPR1A	ERCC3	HNF1A	PTCH1	TP53
BRCA1	ERCC4	HRAS	PTEN	TSC1
BRCA2	ERCC5	KIT	RAD51C	TSC2
BRIP1	EXT1	MAX	RAD51D	VHL
BUB1B	EXT2	MEN1	RB1	WRN
CDC73	EZH2	MET	RECQL4	WT1
CDH1	FANCA	MLH1	RET	XPA
CDK4	FANCB	MSH2	RHBDF2	XPC
CDKN1C	FANCC	MSH6	RUNX1	
CDKN2A	FANCD2	MUTYH	SBDS	
CEBPA	FANCE	NBN	SDHAF2	
CEP57	FANCF	NF1	SDHB	
CHEK2	FANCG	NF2	SDHC	
CYLD	FANCI	NSD1	SDHD	

Приготвяне на библиотеки за секвениране

За приготвянето на библиотеки от таргетни райони за всеки пациент е следван протокола TruSight Enrichment DNA Sample Preparation Guide (Revision B).

Геномната ДНК проба се разрежда до 2.5 ng/µl. След това точната концентрация на двойноверижната ДНК в пробата се определя на Qubit 2.0 Fluorometric Quantitation System, използвайки Qubit dsDNA BR Assay Kit.

Прекачване на праймерни секвенции към ДНК

По време на тази стъпка геномната ДНК минава през процес наречен тагментация (тагване или маркиране и фрагментация) от Nextera транспозоми. Транспозомите на Nextera едновременно фрагментират геномната ДНК и добавя адаптерни секвенции към краищата, позволявайки амплификация чрез PCR в следващите етапи.

За реакцията се ползват 96-ямкови TCY плаки "NET1" (Nextera Enrichment Tagmentation Plate 1). От всяка проба се взимат 20 µl с концентрация 2.5 ng/µl (50 ng) начално количество ДНК и се накапват в отделна ямка. Към тях се добавят 25 µl Tagment DNA Buffer (TD) и 5 µl Tagment DNA Enzyme (TDE1), който съдържа Nextera транспозомите. С помощта на многоканална пипета леко се пипетира няколко пъти, за да се смесят реагентите. Плаката се запечатва със залепящо фолио Microseal 'B' (тази стъпка се повтаря винаги преди плаката да се вортексира, центрофугира или инкубира, след което фолиото се маха и изхвърля). Следва центрофугиране за 1 мин на 280 x g на 20oC. Плаката NET1 се поставя в PCR апарат като се уверим, че капакът на апарата е нагрят по време на инкубацията и се пуска следната програма:

- 55°C в продължение на 5 минути
- Съхранение на режим от 10 ° C

Продължаваме с пречистването на тагментираната ДНК.

Пречистване

Тагментираната ДНК се пречиства от Nextera транспозомите, защото Nextera транспозомите може да се свържат плътно с ДНК краищата и да компроментират по-нататъшната реакцията ако не бъдат премахнати. Към пробата се добавят 15 µl Stop Tagment Buffer (ST) към всяка ямка. Сместа се пипетира няколко пъти и се инкубира на стайна температура (20oC – 25oC) за 5 мин за стопиране на реакцията. След това плаката NET1 се центрофугира на 280x g на стайна температура за 1 мин.

Използвайки многоканална пипета, се прехвърля целия разтвор от всяко ямка на плаката NET1 към чиста 96 ямкова плака, означена "NET2" (Nextera Enrichment Tagmentation Plate 2). Към всяка ямка от NET2 плаката се добавят 52 µl добре разтворени AMPure XP beads, които преференциално се свързват само с ДНК фрагментите. Плаката се инкубира за 10 мин на стайна температура, покрива се с фолио и се центрофугира на 280 xg на 20°C в продължение на 1 минута. След което се маха фолиото и се поставя за 2 мин на магнитна стойка, която утаява пречиствателните мъниста (beads), свързани с ДНК. Супернатантата се премахва от всяка ямка, без да се нарушава интактността на утаените мъниста. Докато плаката е на магнита, се добавят бавно 200 µl пресно приготвен 80% етанол и се изчаква 30 секунди, след което етанолът се премахва от ямките с пипета. Повтаря се накапването с нови 200 µl етанол, последвани от 30 секунди инкубация, след което от ямките се премахва целият остатъчен етанол. Плаката се оставя незапечатана на магнита за 10 мин на стайна температура, за да изсъхне напълно от етанола. След това се маха от

магнита и внимателно се добавят 22.5 µl RSB буфер (Resuspension Buffer), разбъркваме суспензията с пепитиране няколко пъти внимателно. Плаката се поставя за 2 мин на магнитната стойка, докато разтворът се избистри. Супернатантата съдържа отделената от магнитните частици ДНК. 20 µl от супернатантата на всяка ямка се прехвърлят в NEA1 плака (Nextera Enrichment Amplification Plate 1), без да се влиза в контакт с утаените мъниста.

PCR амплификация

В този етап, пречистената тагментирана ДНК се амплифицира чрез PCR програма с лимитиран брой цикли. Стъпката PCR амплификация добавя към краищата на тагментираната ДНК уникална за всяка проба комбинация от индекс 1 (i7) и индекс 2 (i5) с цел мултиплексиране на секвенционната реакция, както и адапторни секвенции (P5 и P7), служещи по-късно за свързването на фрагментите с flow клетка и генерирането на клъстери. Изключително важно е да се използва пълното препоръчано количество ДНК, както и да не се добавят допълнителни цикли на PCR циклите за осигуряване на висококачествени библиотеки, които произвеждат висококачествени резултати за секвениране (Фигура 1A).

Комплектът TruSight Enrichment DNA Sample Preparation kit е проектиран да работи само с индекс 2 праймери с префикс "E". По време на секвенирането MiSeq апарата отчита базите гуанин (G) и тимин (T) със зелен лазер, а аденин (A) и цитозин (C) с червен лазер. Подбира се уникална комбинация от индексови праймери i7 (N701, N702, N703 до N712) и i5 (E501, E502, E503 до E512) за всяка проба от една секвенционна реакция. Всеки индекс се състои от 8 нуклеотида, които служат като баркод за всички секвенции на даден пациент. Двойките индекси за пробите, които впоследствие ще се комбинират в един пул, се избират така, че на всяка една от осемте позиции на индекс i7 и i5 да има поне един нуклеотид, четен от другия лазер (Таблица 2).

Таблица 2: Поддържане на цветови баланс при лазерното отчитане на базите за всяка индексова позиция

Добър индекс				**Лощ индекс**			
	Index 1		Index 2		Index 1		Index 2
705	GGACTCCT	503	TATCCTCT	705	GGACTCCT	502	CTCTCTAT
706	TAGGCATG	503	TATCCTCT	706	TAGGCATG	502	CTCTCTAT
701	TAAGGCGA	504	AGAGTAGA	701	TAAGGCGA	503	TATCCTCT
702	CGTACTAG	504	AGAGTAGA	702	CGTACTAG	503	TATCCTCT
	√√√√√√√√		√√√√√√√√		√√√√√√√√		√√√√xxxx

Към всяка ямка на NEA1 плаката, съдържаща проба, се добавя: 5 µl Index 1 (i7, N7xx) прайпер, 5 µl Index 2 (i5, E5xx) праймер и 20 µl PCR Master Mix (PMM), за краен общ обем на реакцията 50 µl. Внимателно се пипетира за да се разбърка сместа хубаво и NEA1 плаката се покрива с фолио (Microseal 'A') и се уплътнява с гумена ролка след това се центрофугира на

280 x g за 1 мин. След което се слага в PCR апарата с предварително загрят до 100°C капак и се пуска следната PCR програма:

- 72°C за 3 минути
- 98°C за 30 секунди
- 10 цикъла по:

- 98°C за 10 секунди

- 60°C за 30 секунди

- 72°C за 30 секунди

- 72°C за 5 минути
- Изчакване на 10°C

Пречистване след PCR

Тази стъпка използва AMPure XP мъниста за пречистване на ДНК в библиотеката и осигурява избор за размер на фрагментите, като премахва много къси фрагменти от населението.

Изваждат се AMPure XP мъниста и се размразяват на стайна температура за около 30 минути. Изваждат се NEA1 плаката от PCR апарата и се центрофугират на 280 x g за 1 мин. Целият обем (50 µl) от всяка ямка се прехвърля в отделна ямка на нова "NEA2" плака (Nextera Enrichment Amplification Plate 2). След много добро разбъркване на AMPure XP мъниста към всяка ямка се добавят 45 µl от пречиствателните мъниста. Внимателно се пипетира за да се разбърка хубаво сместа след което се инкубира на стайна температура за 10 мин. NEA2 плаката се поставяне на магнитната стойка за 2 мин или докато супернатантата се избистри. Внимателно супернатантата се отстранява, след което следват две последователни измивания с 200 µl прясно приготвен 80% етанол с по 30 сек изчакване след всяко от тях. След премахване на остатъчния етанол плаката отново се оставя за сушене за 15 мин на стайна температура, след което се маха от магнита и към всяка ямка се добавят 40 µl RSB буфер. Микса се пипетира внимателно за да се ресуспендират мънистата, инкубира се за 2 мин на стайна температура на магнита. От всяка ямка се прехвърлят 38 µl от супернатанта към NEL плаката (Nextera Enrichment Library Plate).

Концентрацията на всяка проба беше измервана посредством Qubit 2.0 Fluorometric Quantitation System.

Първа хибридизация

За хибридизацията се приготвя обща ДНК библиотека (pool), съдържаща различно индексирани проби, която впоследствие да се пусне в една секвенционна реакция. От всяка проба се взимат 500 ng ДНК и се добавят в общия пул. Ако общият обем на смесените проби надвишава 40 µl, пулът се концентрира с помощта на вакуумен концентратор без прилагане на топлина. Ако обемът на пула е под 40 µl, се добавя необходимото количество RSB до 40 µl. Капацитетът на секвенатора MiSeq

и уникалното индексиране на всяка проба позволява в едно секвениране да се комбинират подготвените библиотеки на различни индивиди. Броят на пробите, които могат да се секвенират едновременно, зависи от размера на таргетния район и желаната дълбочина на секвениране. В едно секвениране могат да се комбинират до 12 различни библиотеки (12 различни пациента) подготвени с панела TruSight Cancer.

Преди приготвянето на хибридизационната реакция NCT1 (Nextera Capture Target Buffer 1) се вортексира продължително до пълното ресуспендиране на кристалите в него. След това към всеки приготвен пул (40 µl) се добавят 50 µl NCT1, както и 10 µl от специфичния за всеки панел TruSight Content Set CSO (Custom Selected Oligos), който съдържа набора от олигосонди, комплементарни на таргетните райони, включени в панела. След внимателно пипетиране за да се разбърка сместа, тя се прехвърля във всяка ямка на новите 300 µl 96-ямкова PCR плака или двойна 96-ямкова PCR плака означена като CTP1. Центрофугира се CTP1 плаката на 280 x g за 1 мин, след което се поставят в PCR апарата на следната програма:

- 95°C за 10 минути (денатурация)
- 18 цикъла по една минута, започващи от 93°C, като всеки следващ цикъл температурата спада с 2°C
- 58°C за 16 - 20 часа

За оптимизиране на ефикасността на хибридизацията продължителността на последната стъпка при първата хибридизация за всеки пул беше между 15 и 20 часа. След денатурацията (Фигура 1Б) комлементарните олигосонди се свързват само с ДНК фрагментите, съдържащи таргетни райони (Фигура 1В). Всяка олигосонда е биотинилирана - т.е. представлява нуклеотидна секвенция, към която ковалентно е прикрепена молекулата биотин (витамин B7).

Първо пречистване

Измиването е процедура на извличане на таргетните райони (Capture). По време на извличането на таргетните райони към хибридизационната реакция се добавят магнитни частици, към повърхността на които са прикрепени тетрамерни форми на белтъка стрептавидин. Стрептавидинът и биотинът образуват една от най-силните нековалентни връзки в природата с дисоциационна константа (Kd) ≈ 10−14 mol/L. Така стрептавидиновите магнитни частици се свързват с биотинилираните сонди, които от своя страна са се хибридизирали с таргетните райони (Фигура 1Г).

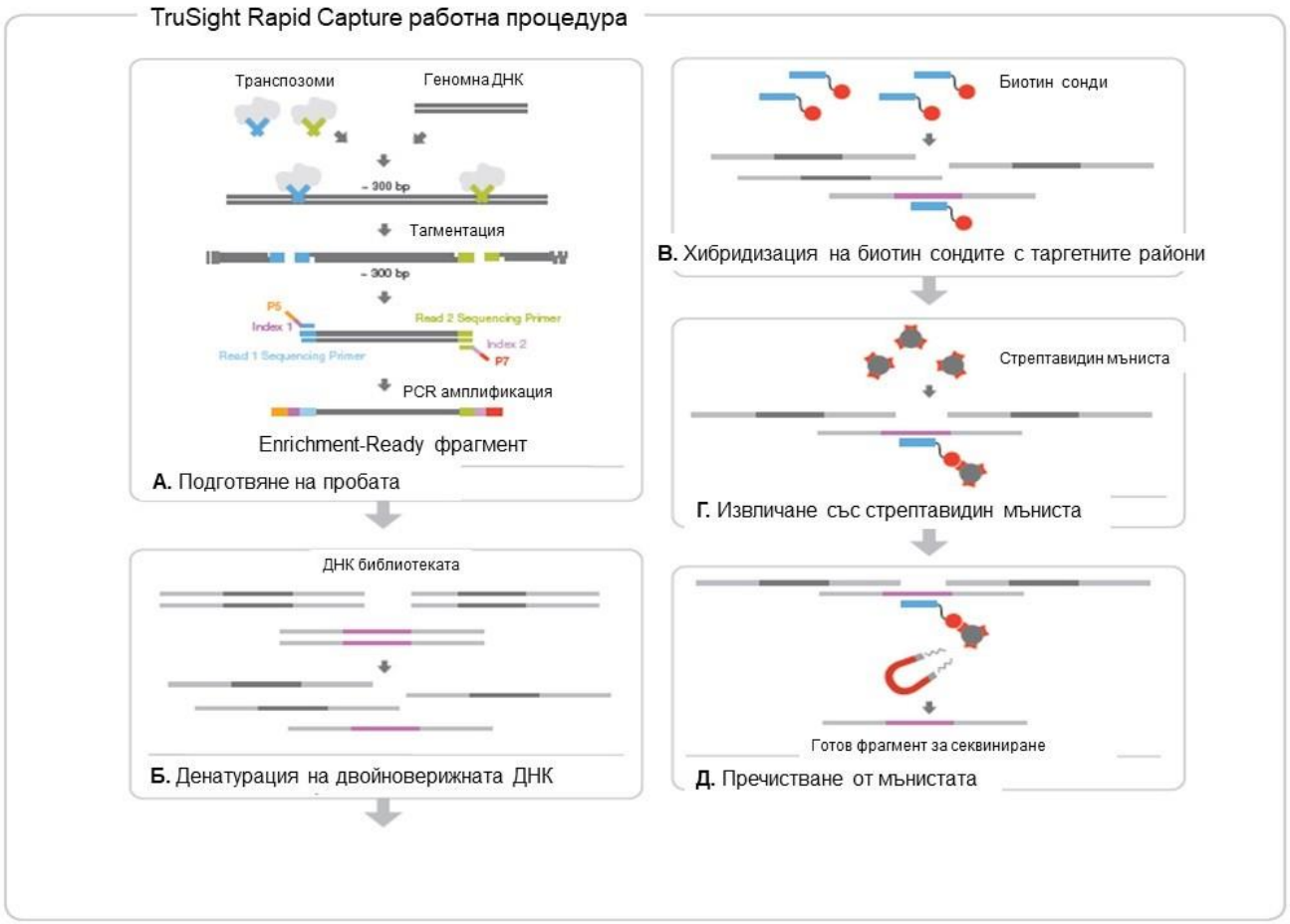
След края на хибридизацията CTP1 плаката се центрофугира на 280 x g за 1 мин и съдържанието се прехвърля в нова 96 ямкова MIDI плака с надпис WTP1. След като бъдат добре вортексирани, към всяка ямка на WTP1 се добавят 250 µl Streptavidin Magnetic Beads (SMB). Микса в плаката се пипетира внимателно, плаката се покрива и се инкубира на стайна температура за 30 мин. По време на тази стъпка се осъществява взаимодействието между стрептавидина и биотина. След инкубацията плаката се центрофугира (280 x g, 1 мин) и се поставя за 2 мин на

магнитната стойка. Супернатантата се премахва, без да се влиза в контакт с магнитните мъниста, след което плаката се сваля от магнита.

Следват две последователни измивания за отстраняване на неспецифични взаимодействия на стрептавидиновите мъниста и изолиране само на таргетните райони. Епруветката с промиващ разтвор (Wash Solution 1 - WS1) се темперира до стайна температура и се вортексира продължително. Добавя се 200 µl от промиващ разтвор WS1 към утаените мъниста във всяка ямка от плаката, библиотеките се ресуспендират чрез пипетиране на целия им обем 10-20 пъти. Плаката WTP1 се поставя на магнитната стойка за 2 минути на стайна температура или докато течността се избистри. Супернатантата се маха и плаката се сваля от магнита.

След това наново се накапват 200 µl промиващ разтвор WS2 (Wash Solution 2) към утаените частици във всяка ямка от плаката, библиотеките се ресуспендират чрез пипетиране на целия им обем 10-20 пъти. Плаката WTP1 се поставя на магнитната стойка за 2 минути на стайна температура или докато течността се избистри. Супернатантата се маха и плаката се сваля от магнита. Добавя се отново WS2 към всяка ямка на WTP1 плаката и библиотеките се ресуспендират много добре чрез пипетиране на целия им обем 10-20 пъти. Целия обем на реакцията от всяка ямка на WTP1 плаката се прехвърля в съответната ямка на новата 96 ямкова PCR плака наречена IWP1. Плаката се затваря и се инкубира на PCR апарат за 30 мин на 42°C със загрят капак до 100oC. Плаката IWP1 се изважда от термален сайклър и се поставя веднага върху магнитната стойка за 2 мин или докато течността се избистри. Супернатантата се маха и плаката се сваля от магнита, след което наново се накапват 200 µl WS2 и отново се повтарят същите последователни стъпки.

Добавя се 200 µl от промиващ разтвор WS3 към утаените частици във всяка ямка от плаката, наново библиотеките се ресуспендират чрез пипетиране на целия им обем 10-20 пъти. Плаката IWP1 се поставя на магнитната стойка за 2 минути на стайна температура или докато течността се избистри. Супернатантата се маха и отново се повтарят същите последователни стъпки.



Фигура 1: Работната процедура за подготвяне на библиотеката за секвениране по протокола TruSight Rapid Capture Sample Preparation guide

След повторното премахване на супернатантата с WS3 следва елуирането на таргетните райони от стрептавидин-биотиновия комплекс. За всеки пул се приготвя елуционен премикс (elution pre-mix), съдържащ 28.5 µl Elute Target Buffer 1 (ET1) и 1.5 µl 2N NaOH (HP3) прави общ обем за всяка реакция 30 µl. След кратко центрофугиране се добавят 23 µl от премикса към утаените стрептавидинови мъниста на всеки ямка от IWP1 плаката. Следва ресуспендиране много добре чрез пипетиране 10-20 пъти, затваряме плаката и инкубираме за 5 мин на стайна температура, при която става самото елуиране на ДНК фрагментите, и центрофугиране плаката на 280 x g за 1 мин. Плаката отново се поставя за 2 мин на магнит, като този път ДНК фрагментите не са свързани с утаените мъниста, а се намират в супернатантата. От нея се взимат 21 µl и се прехвърлят в нова 96 ямкова PCR плака надписана CTP2. Към тях се добавят 4 µl Elute Target Buffer 2 (ET2) за неутрализиране на елуирането. Обогатеният на таргетни райони пул се пипетира внимателно няколко пъти и плакта се покрива.

Втора хибридизация

Този процес смесва първото елуиране на ДНК библиотеката със сондите за улавяне на таргетните райони. Втората хибридизация гарантира, че таргетните райони са допълнително амплифицирани.

Към елуирана библиотеката (25 µl) се добавят 50 µl NCT1 (Nextera Capture Target Buffer 1), след вортексиране и ресуспендиране на кристалите, 10 µl от TruSight Content Set CSO (Custom Selected Oligos), съдържащ олигосондите, и 15 µl PCR Grade Water, което прави общ обем от 100 µl, целия обем се пипитира внимателно в CTP2 плака. Следва центрофугиране на затворената плака на 280 x g за 1 мин, след което CTP2 плаката се поставя в PCR апарата на следната програма:

- 95°C за 10 минути (денатурация)
- 18 цикъла по 1 минута, започващи от 93°C, като всеки следващ цикъл температурата спада с 2°C
- 58°C за 16 – 20 часа

Второ пречистване

Този процес използва частици от стрептавидин за улавяне на сондите, съдържащи таргетните райони. Три стъпки на измиване премахват неспецифично свързваните фрагменти от частиците. След което амплифицираната библиотека се елуира от частиците и се приготвя за секвениране.

След края на хибридизацията CTP2 плаката се центрофугира на 280 x g за 1 мин и съдържанието се прехвърля в нова 96 ямкова MIDI плака с надпис WTP2. След като бъдат добре вортексирани, към всяка ямка на WTP2 се добавят 250 µl Streptavidin Magnetic Beads (SMB). Микса в плаката се пипетира внимателно, плаката се покрива и се инкубира на стайна температура за 30 мин. По време на тази стъпка се осъществява взаимодействието между стрептавидина и биотина. След инкубацията плаката се центрофугира (280 x g, 1 мин) и се поставя за 2 мин на магнитната стойка. Супернатантата се премахва, без да се влиза в контакт с магнитните частици, след което плаката се сваля от магнита.

Следват две последователни измивания за отстраняване на неспецифични взаимодействия на стрептавидиновите частици и изолиране само на таргетните райони. Епруветката с Wash Solution 1 (WS1) се темперира до стайна температура и се вортексира продължително. Добавя се 200 µl от WS1 към утаените частици във всяка ямка от плаката, библиотеките се ресуспендират чрез пипетиране на целия им обем 10-20 пъти. Плаката WTP2 се поставя на магнитната стойка за 2 минути на стайна температура или докато течността се избистри. Супернатантата се маха и плаката се сваля от магнита.

След това наново се накапват 200 µl WS2 (Wash Solution 2) към утаените частици във всяка ямка от плаката, библиотеките се ресуспендират чрез пипетиране на целия им обем 10-20 пъти. Плаката WTP2 се поставя на магнитната стойка за 2 минути на стайна температура или докато течността се избистри. Супернатантата се маха и плаката се сваля от магнита. Добавя се отново WS2 към всяка ямка на WTP2 плаката и библиотеките се ресуспендират много добре чрез пипетиране на целия им обем 10-20 пъти. Целия обем на реакцията от всяка ямка на WTP2 плаката

се прехвърля в съответната ямка на новата 96 ямкова PCR плака наречена IWP2. Плаката се затваря и се инкубира на PCR апарат за 30 мин на 42°C със загрят капак до 100oC. Плаката IWP2 се изважда от PCR апарат и се поставя веднага върху магнитната стойка за 2 мин или докато течността се избистри. Супернатантата се маха и плаката се сваля от магнита, след което наново се накапват 200 µl WS2 и отново се повтарят същите последователни стъпки.

Добавя се 200 µl от WS3 към утаените мъниста във всяка ямка от плаката, наново библиотеките се ресуспендират чрез пипетиране на целия им обем 10-20 пъти. Плаката IWP2 се поставя на магнитната стойка за 2 минути на стайна температура или докато течността се избистри. Супернатантата се маха и отново се повтарят същите последователни стъпки.

След повторното премахване на супернатантата с WS3 следва елуирането на таргетните райони от стрептавидин-биотиновия комплекс (Фигура 5). За всеки пул се приготвя елуционен премикс (elution pre-mix), съдържащ 28.5 µl Elute Target Buffer 1 (ET1) и 1.5 µl 2N NaOH (HP3) прави общ обем за всяка реакция 30 µl. След кратко центрофугиране се добавят 23 µl от премикса към утаените стрептавидинови мъниста на всеки ямка от IWP2 плаката. Следва ресуспендиране много добре чрез пипетиране 10-20 пъти, затваряме плаката и инкубираме за 5 мин на стайна температура, при която става самото елуиране на ДНК фрагментите, и центрофугиране плаката на 280 x g за 1 мин. Плаката отново се поставя за 2 мин на магнит, като този път ДНК фрагментите не са свързани с утаените частици, а се намират в супернатантата. От нея се взимат 21 µl и се прехвърлят в нова 96 ямкова PCR плака надписана TTP. Към тях се добавят 4 µl Elute Target Buffer 2 (ET2) за неутрализиране на елуирането. Обогатеният на таргетни райони пул се пипетира внимателно няколко пъти и плакта се покрива.

PCR амплификация

По време на този PCR се амплифицират само специфичните таргетни райони, извлечени след двете хибридизации. След прехвърляне на 20 µl от библиотеката от TTP плаката към нова 96 ямкова PCR плака наречена TAP1 се добавят 5 µl PCR Primer Coctail (PPC) и 25 µl PMM (PCR Master Mix). Следва центрофугиране на 280 x g за 1 мин и пускане на следната PCR програма:

- 98°C за 30 секунди
- 12 цикъла по:

- 98°C за 10 секунди

- 60°C за 30 секунди

- 72°C за 30 секунди

- 72°C за 5 минути
- Изчакване на 10°C

Пречистване след PCR амплификация

В тази стъпка се извършва финално пречистване на амплифицираната библиотека, съдържаща таргетните райони, преди секвенирането.

TAP1 плаката се вади от PCR апарата, към съдържанието й (50 µl) се добавят добре вортексирани 90 µl AMPure XP Beads. Суспензията в плаката внимателно се пипетира за да се разбърка хубаво и се инкубира за 15 мин на стайна температура. След това плаката се поставяне на магнитната стойка за 5 мин. Супернатантата (140 µl) от TAP1 плаката се премахва, след което утаените частици се промиват два последователни пъти с 200 µl прясно приготвен 80% етанол с по 30 сек изчакване след всяко от тях. След премахване на етанола плаката се инкубира за 15 мин на стайна температура, за да изсъхнат утаените частици, след което се маха от магнита. Добавят се 30 µl RSB хубаво се разбърква и се инкубира за 2 мин на стайна температура. Плаката се поставя за 5 мин на магнитната стойка. Прехвърля се от всяка ямка на TAP1 плаката 28 µl от супернатантата, съдържаща таргетните ДНК фрагменти в нова 96 ямкова PCR плака надписана CAP. С тази стъпка завършва процесът на приготвяне на библиотеките за секвениране.

Определяне на концентрацията на библиотеките

KAPA Library Quantification Kit е използван за определяне на концентрацията на всяка от приготвените библиотеки преди зареждане на секвенционната реакция. Използван е протоколът KAPA Library Quantification Kit за Illumina platforms на KAPA Biosystems.

Приготвят се две разредки на приготвената библиотека - 1:1000 и 1:2000. 1 µl от библиотеката се смесва с 999 µl RSB. След 15 сек вортексиране се взимат 100 µl от първата разредка (1:1000) и се смесват с 100 µl RSB, за да се получи втората разредка от 1:2000. Реагентът 10X Primer Premix съдържа праймери, коплементарни на адапторните райони в краищата на всеки фрагмент от библиотеката. 1 ml 10X Primer Premix се смесва с 5 ml 2X KAPA SYBR FAST qPCR Master Mix, за да се приготви финалния мастър микс. В реакционната смес за определяне на концентрацията се смесват 12 µl мастър микс, 4 µl PCR-grade H2O и 4 µl от разредките на библиотеката или от ДНК стандартите. Използват се 6 ДНК стандарти със следните концентрации: 20 pM/µl, 2 pM/µl, 0.2 pM/µl, 0.02 pM/µl, 0.002 pM/µl и 0.0002 pM/µl. В 48-ямкова плака се накапват по три реакции за всяка от двете разредки на библиотеката и за всеки от шестте стандарта. Пуска се и една негативна контрола, в която вместо ДНК се добавя RSB. Плаката се поставя в qPCR система Illumina Eco и се пуска следната програма:

- 95°C за 5 минути
- 35 цикъла по:

- 95°C за 30 секунди

- 60°C за 45 секунди

За изчисляване на концентрацията на библиотеката бе ползвана по-малката разредка от 1:1000. Концентрацията се изчислява по формулата: C x 452/500 x 1000, където C е концентрацията на разредката, установена от qPCR-а, 452 е дължината в бази на стандартите, а 500 бази е средната дължина на фрагментите от приготвените библиотеки.

Денатурация и разреждане на фрагментите за секвениране

За всяко от секвенираният беше използван MiSeq Reagent Kit v2, включващ стандартна поточна (flow) клетка за 500 цикъла (цикъла на секвениране) (Фигура 2). При всяко секвениране flow клетката беше зареждана с концентрация в диапазона 15-25 pM/µl. Подобни концентрации дават оптимална гъстота на клъстъри по време на секвенирането. След определяне на концентрацията всяка от библиотеките е разреждана до 3-5 nM/µl. За денатурация бе ползвана 0.2 N NaOH. 5 µl от приготвената библиотека се смесва с 5 µl 0.2N NaOH. Сместа се инкубира за 5 мин на стайна температура, след което към нея се добавя 990 µl хибридизационен буфер HT1. Полученият 1 ml се накапва в предварително размразената на лед реагентна касета съдържаща необходимите реагенти за секвенирането (Фигура 3). Използвани бяха MiSeq Reagent Kit v2. След изплакване поточната клетка, на която се извършва клъстър генерирането и разчитането на базите, се поставя на мястото си в MiSeq апарата. Поставят се също и PR бутилката, съдържаща буфер за изплакване, и реагентната касета, след което се дава начало на секвенирането от контролния софтуер.



Фигура 2: Стандартна поточна клетка (flow cell), на която се извършва целият процес на секвениране



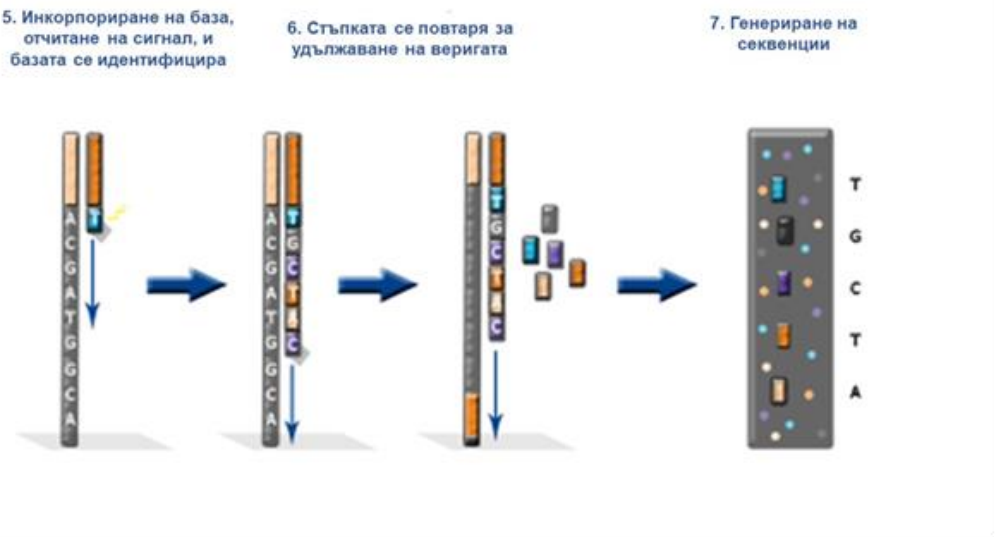
Фигура 3: Касетата с резервоара за реагентите, в който се добавят ДНК библиотеките от таргетните секвенции

Секвениране с Illumina

Процесът на секвениране започва като подготвените ДНК фрагменти навлизат в поточната клетка чрез микроканалчета. По повърхността ѝ ковалентно са фиксирани огромен брой олигонуклеотиди, комплементарни на адапторните секвенции, добавени по време на приготвянето на библиотеките. ДНК фрагментите се закрепят за поточната клетка (Фигура 4А - 1), след което чрез така наречената мостова (bridge) амплификация (Фигура 4А - 2) се генерират клъстъри. Така се създават голям брой клъстъри, всеки съдържащ хиляди копия на един специфичен ДНК фрагмент (Фигура 4А - 3). След генерирането на клъстърите започва самото разчитане на нуклеотидната последователност. Секвенционни праймери, комплементарни на част от адапторната секвенция, се хибридизират с намножените фрагменти (Фигура 4А - 4). Секвенирането се извършва по време на синтез (sequencing-by synthesis) чрез използването на dNTP-та, всеки от които е свързан с флуресцентно белязан терминатор. При инкорпориране на дадена база (Фигура 4Б - 5) флуресцентният сигнал се отчита от лазер и базата се идентифицира. След това терминаторът се отстранява, за да позволи инкорпорирането на следващата флуресцентно белязана база и удължаването на веригата (Фигура 4Б - 6). Крайният резултат е генерирането на секвенции с висока степен на прецизност (Фигура 4Б - 7).



Фигура 4А: Първите стъпки в процеса на секвениране с технологията на Illumina



Фигура 4Б: Пподължение на стъпките в процеса на секвениране с технологията на Illumina

След приключване на разчитането на базите започва първичният анализ на получените секвенции от самия секвенатор чрез софтуера MiSeq Reporter, който извършва демултиплексиране и разпределяне на секвенциите за всеки един от пациентите в отделна директория. Освен това получените секвенции се подравняват с референтния човешки геном (версия GRCh37) и несъответствията се отчитат като варианти. За подравняване на последователностите спрямо референтния геном е ползван алгоритъма на BurrowsWheeler (Li H, Durbin R. 2009) или Isaac, а за идентифициране на вариантите (еднонуклеотидни замени, делеции / инсерции с дължина до 25 бази) – приложението GATK Variant Caller или Isaac Variant Caller. Данните от всяко секвениране бяха също така качвани и в облака на Illumina - BaseSpace, където впоследствие могат да се извършват различен тип допълнителни анализи.

Анализ на данните

След завършване на секвенирането, анализът на данните се извършва с Softgenetics NextGene Software (версия 2.3.3). Софтуерът автоматично обработва данните, които извършват подравняването на последователностите с човешкия референтен геном (GRCh37), филтрирането и variant calling. Намерените варианти бяха допълнително анализирани чрез софтуерите за прогнозиране на мутация PROVEAN, SIFT и PolyPhen-2 (версия 2.2.2), за да се установи какъв е техният възможен ефект върху протеиновата функционалност. Вариантите бяха проверени и анализирани допълнително с използване на базата данни PharmGKB (https://www.pharmgkb.org/) за фармакогенетични SNP варианти, свързани с чувствителността към определени лекарства.

3.3. Полимеразна верижна реакция в реално време (Real-time PCR)

Пречистените ДНК проби от 551 туморни проби на пациенти с НДКРБД бяха амплифицирани с therascreen *EGFR* RGQ PCR Kit за Rotor-Gene Q апарат на QIAGEN. Therascreen *EGFR* RGQ китът се използва за определяне на 29 соматични мутации в екзони 18-21 на EGFR гена и е базиран на две технологии: ARMS и Scorpions. ARMS (Amplification Refractory Mutation System) техногията извършва алел-специфична амплификация на фрагмент, съдържащ мутацията, докато Scorpions технологията амплифира дивия тип фрагмент, който не съдържа мутация.

Китът therascreen *EGFR* RGQ PCR позволява откриване на следните типове мутации:

- 19 делеции в екзон 19 (но не ги диференцира);
- Точкови мутации: T790M в екзон 20; L858R в екзон 21; L861Q в екзон 21; S768I в екзон 20;
- Точкова мутация G719X (без да разграничава следните мутации - G719S, G719A или G719C);
- 3 инсерции в екзон 20 (открива тяхното наличие, но не ги разграничава).

Амплификация на ДНК

Амплифицирането на ДНК пробите беше извършено чрез спазване на протокол therascreen® *EGFR* RGQ PCR Kit Handbook (Version 1). Крайната концентрацията на ДНК беше доведена до 10 ng/µl и измерена на спектрофотомет NanoDrop Lite.

PCR процедура

Реакционните смеси в кита *therascreen® EGFR*, както и отрицателната контрола (NTC – no template control) и положителната контрола (PC) за *EGFR* мутации предварително се размразяват при стайна температура (15 - 25°C). Епруветките се бъркат чрез обръщане около 10 пъти,за да се избегне утаяване на концентрирани соли, след това се центрофугират за кратко, за да се събере съдържанието на епруветките.

Реакционните смеси се приготвя като се смесват 19.5 µl Reaction mix (съответно Ctrl, T790M, Del, L858R, L861Q, G719X, S768I и Ins) и 0.5 µl ДНК Taq полимераза. Сместа се разбърква добре, като внимателно се пипетира 10 пъти (Таблица 3).

Таблица 3: Разпределението на тестваните и контролните проби в реакционите смеси за Rotor-Gene Q

	Контроли		**Номерата на тестваните проби**						
Реакционни смеси	**PC**	**NTC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Ctrl - Control Reaction Mix	1	9	17	25	33	41	49	57	65
T790M	2	10	18	26	34	42	50	58	66
Deletions	3	11	19	27	35	43	51	59	67
L858R	4	12	20	28	36	44	52	60	68

L861Q	5	13	21	29	37	45	53	61	69
G719X	6	14	22	30	38	46	54	62	70
S768I	7	15	23	31	39	47	55	63	71
Insertions	8	16	24	32	40	48	56	64	72

Епруветка No 1 съдържа 20 µl реакционна смес Ctrl, към която се добавя 5 µl PC.

Епруветка No 2 съдържа 20 µl реакционна смес T790M, към която се добавя 5 µl PC.

Епруветка No 3 съдържа 20 µl реакционна смес Deletions, към която се добавя 5 µl PC.

Епруветка No 4 съдържа 20 µl реакционна смес L858R, към която се добавя 5 µl PC.

Епруветка No 5 съдържа 20 µl реакционна смес L861Q, към която се добавя 5 µl PC.

Епруветка No 6 съдържа 20 µl реакционна смес G719X, към която се добавя 5 µl PC.

Епруветка No 7 съдържа 20 µl реакционна смес S768I, към която се добавя 5 µl PC.

Епруветка No 8 съдържа 20 µl реакционна смес Insertions, към която се добавя 5 µl PC.

Епруветка No 9 съдържа 20 µl реакционна смес Ctrl, към която се добавя 5 µl NTC.

Епруветка No 10 съдържа 20 µl реакционна смес T790M, към която се добавя 5 µl NTC.

Епруветка No 11 съдържа 20 µl реакционна смес Deletions, към която се добавя 5 µl NTC.

Епруветка No 12 съдържа 20 µl реакционна смес L858R, към която се добавя 5 µl NTC.

Епруветка No 13 съдържа 20 µl реакционна смес L861Q, към която се добавя 5 µl NTC.

Епруветка No 14 съдържа 20 µl реакционна смес G719X, към която се добавя 5 µl NTC.

Епруветка No 15 съдържа 20 µl реакционна смес S768I, към която се добавя 5 µl NTC.

Епруветка No 16 съдържа 20 µl реакционна смес Insertions, към която се добавя 5 µl NTC.

Епруветка No 17 съдържа 20 µl реакционна смес Ctrl, към която се добавя 5 µl проба №1.

Епруветка No 18 съдържа 20 µl реакционна смес T790M, към която се добавя 5 µl проба №1.

Епруветка No 19 съдържа 20 µl реакционна смес Deletions, към която се добавя 5 µl проба №1.

Епруветка No 20 съдържа 20 µl реакционна смес L858R, към която се добавя 5 µl проба №1.

Епруветка No 21 съдържа 20 µl реакционна смес L861Q, към която се добавя 5 µl проба №1.

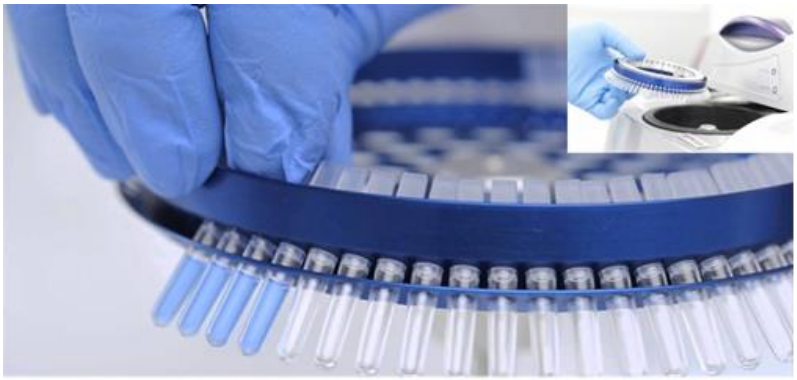
Епруветка No 22 съдържа 20 µl реакционна смес G719X, към която се добавя 5 µl проба №1.

Епруветка No 23 съдържа 20 µl реакционна смес S768I, към която се добавя 5 µl проба №1.

Епруветка No 24 съдържа 20 µl реакционна смес Insertions, към която се добавя 5 µl проба №1.

Добавя се 5 µl от всяка проба към епруветките за PCR (епруветки номер 17-72 при максимален брой за апарата Rotor Gene - 7 проби за едно пускане) и се затварят с капачките.

Всяка ДНК проба трябва да бъде тествана, както за Ctrl, така и за всички останали мутации чрез амплификация в реално време. PCR епруветките се поставят в съответните позиции на ротора (Фигура 5).



Фигура 5: Зареждане на ротора на Rotor-Gene Q апарата с проби

Роторът се поставя в Rotor-Gene Q апарата и се пуска следната програма:

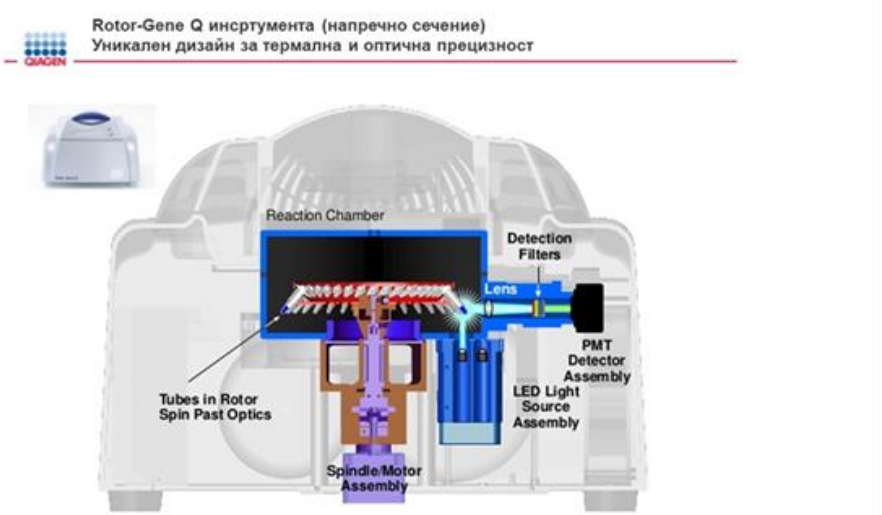
- 95°C за 15 минути
- 40 цикъла по:

- 95°C за 30 секунди

- 60°C за 60 секунди

PCR амплификация

Уникалният центрофужен ротационен дизайн на Rotor-Gene Q го прави най-точният и гъвкав PCR апарат в реално време (Фигура 6). Всяка епруветка се завърта в камера на движещ се въздух, като всички проби са изложени на една и същата температура по време на бърза термична циркулация. Когато всяка епруветка премине през оптичен детектор, пробите са осветяват и се възбужда флуоресцентен сигнал, който преминава бързо през един единствен, къс оптичен път. Този механизъм води до чувствителен, прецизен и бърз PCR анализ в реално време.



Фигура 6: Напречно сечение на Rotor-Gene Q апарата. Отопление и охлаждане се постига чрез бърз въздушен поток в реакционната камера. Епруветките преминават през оптиката на всеки 150 милисекунди, което позволява високоскоростно заснемане на данните. До 6 отделни LED светлинни източници могат да се използват в комбинация с 6 различни високо чувствителни филтъра за детекция (file:///C:/Users/zorah/Downloads/Rotor-Gene-Q--Pure-Detection-IT%20(1).pdf)

Rotor-Gene Q работи с различни видове багрила като SYBR® Green, сонди за хидролиза (TaqMan), хибридизация (FRET) и др. С до 6 канала, обхващащи от ултравиолетовите (UV) до инфрачервените (IR) дължини на вълните, циклера осигурява най-широкия оптичен диапазон (Таблица 4). Освен това, софтуерът позволява да се създадат излъчвания с различни комбинации от дължини на вълните, което означава, че Rotor-Gene Q е съвместими с бои, които потенциално могат да се използват в бъдеще.

Таблица 4: Каналите на Rotor-Gene Q апарата за оптичното детектиране

Каналите	Възбуждане (nm)	Детектиране (nm)	Примери за открити флуорофори
Синьо (Blue)	365±20	460±20	Marina Blue®, Edans, Bothell Blue, Alexa Fluor® 350

Зелено (Green)	470±10	510±5	FAM™, SYBR Green I, Fluorescein, EvaGreen®, Alexa Fluor 488
Жълто (Yellow)	530±5	557±5	JOE™, VIC®, HEX, TET™, CAL Fluor® Gold 540, Yakima Yellow®
Оранжево (Orange)	585±5	610±5	ROX™, CAL Fluor Red 610, Cy®3.5, Texas Red®, Alexa Fluor 568
Червено (Red)	625±5	660±10	Cy5, Quasar® 670, LightCycler® Red640, Alexa Fluor 633
Пурпурно (Crimson)	680±5	712 high pass	Quasar 705, LightCycler Red705, Alexa Fluor 680
HRM	460±20	510±5	SYBR Green I, SYTO®9, LC Green®, LC Green Plus+, EvaGreen

Анализ на EGFR мутации

След завършване на PCR теста, ДНК пробите се анализират за наличие или отсъствие на EGFR мутации със софтуер на Rotor-Gene Q версия (2.0.2).

Ако има амплификация за мутация в определена проба, разликата между прага на цикъла на детектираната мутация (CT) и контролната амплификация (CT) на същата проба се използва за изчисляване на стойноста ΔCT; ΔCT = mutation (мутация) CT – control (контрола) CT. След това сравняваме стойността на ΔCT за пробата с прага (cut-off) съгласно указанията на производителя (Qiagen) Таблица 5. Резултатът от амплификацията се счита за валиден ако, FAM CT стойността попада в диапазона от 15.00-40.00 цикъла. Ако FAM CT стойноста попада над зададения диапазон или няма амплификация, това е FAM отрицателна амплификация. Ако ΔCT стойността на пробата е по-висока от cut-off стойността, тя се класифицира като "не открита мутация" или извън границите на детекция на кита. Ако ΔCT стойността на пробата е в или под cut-off стойноста, пробата се счита за открита положителна мутация с този анализ.

Таблица 5: Приемливи стойности на реакционната мутация (FAM), за therascreen® EGFR RGQ PCR Kit

Таблица 5: Приемливи стойности на реакционната мутация (FAM), за *therascreen*® EGFR RGQ PCR Kit		
Реакционен микс	**Приемлива C_T стайност**	**Cut-off ΔC_T стойност**
T790M	15.00–40.00	6.38
Deletions	15.00–40.00	9.06
L858R	15.00–40.00	8.58
L861Q	15.00–40.00	9.26
G719X	15.00–40.00	9.31
S768I	15.00–40.00	9.26
Insertions	15.00–40.00	7.91

Статистически метод

Използвахме chi-square тест за оценка на връзката между мутациите в EGFR гена и пола, статуса на пушене (пушачи и не-пушачи) и хистологичния туморен фенотип (аденокарцином, сквамозноклетъчен карцином и едро клетъчен карцином). P-стойностите по-малки от 0,05 и се считат за статистически значими.

P- стойността се използва при тестове за значимост на нулева хипотеза, при която първо се избира модел (нулевата хипотеза) и прагова стойност за p, наречена нивото на значимост на теста (5% или 1%). Ако p-стойността е по-малка от избраното ниво на значимост, това предполага, че използваните данни са несъвместими с нулевата хипотеза, и че нулевата хипотеза може да бъде отхвърлена. За анализа праговата стойност се приема за 0.05 (cutoff), като нулевата хипотеза се отхвърля, когато $p < 0.05$ (статистически значима разлика), и не се отхвърля, когато $p > 0.05$ (няма статистически значима разлика).

III. Резултати

1. Данни от NGS секвениране с карциномен панел на туморни проби от НДКБК

1.1 Анализ на варианти свързани с туморогенеза при НДКБК

Отговарящите на критериите за включване, 18 туморни проби от недребноклетъчен белодробен карцином (НДКБК) са подложени на NGS-базирано таргетно секвениране с панел от 94 гени, асоциирани с онкогенезата (TruSight Cancer Sequencing Panel, Illumina). Установиха се 21 патогенни SNP-варианти свързани с туморогенезата в гените *BRCA1*, *PMS2*, *MLH1* и *PALB2* (Табл.6). Всички открити варианти са анотирани по сенс веригата.

Оценката на патогенността на намерените варианти е базирана на SIFT и PROVEAN стойността, генерирани от специално разработени софтуери за оценка на очаквания ефект на мутацията върху функцията на протеина (http://provean.jcvi.org/about.php#about_1).

PROVEAN стойността се изчислява, като се взима средната стойност на резултатите в рамките на и между получените клъстери. Ако стойността на PROVEAN е равна или по-ниска от предварително заложения праг (cutoff -2.5), се очаква вариантът да влияе негативно върху функцията на протеина. При PROVEAN коефицент над зададения праг се приема, че вариантът няма функционален ефект.

SIFT е програма, която класифицира вариантите като „толерантни“ (tolerated) или „патогенни“ (damaging) в зависимост от това дали вариантът се намира в еволюционно консервативен участък и оценката се базира на хипотезата, че еволюцията на протеините корелира с функция им. Варианти в генни сегменти със стойност <0,05 се очаква да бъдат „патогенни“. Тези, при които SIFT коефициентът е ≥0,05, се очаква да бъдат „толерантни“.

Установиха се 7 патогенни варианти в четири гени *BRCA1*, *PMS2*, *MLH1* и *PALB2*, свързани с туморогенезата (Табл.6). Всички открити варианти са анотирани по сенс веригата.

Таблица 6: Варианти свързани с онкогенеза в пациенти с НДКБК

Пациенти	Гени	Варианти	Генотип на пациента	PROVEAN (cutoff=-2.5)	SIFT (cutoff=0.05)
№1	*BRCA1*	rs16941	CC	Патогенен	Патогенен
	BRCA1	rs80356964	GA	Патогенен	Патогенен
№4	*BRCA1*	rs16941	TC	Патогенен	Патогенен
	BRCA1	rs1799950	TC	Патогенен	Толерантен
№5	*BRCA1*	rs16941	TC	Патогенен	Патогенен
	BRCA1	rs1799950	TC	Патогенен	Толерантен
№6	*BRCA1*	rs1799950	TC	Патогенен	Толерантен
	MLH1	rs63751049	CT	Патогенен	Патогенен
№7	*BRCA1*	rs16941	TC	Патогенен	Патогенен
	BRCA1	rs80356964	AG	Патогенен	Патогенен
№10	*BRCA1*	rs16941	CC	Патогенен	Патогенен
	PMS2	rs146176004	CG	Патогенен	Патогенен
№11	*BRCA1*	rs16941	CC	Патогенен	Патогенен
	PALB2	rs75023630	CA	Патогенен	Патогенен
№15	*BRCA1*	rs80356964	AG	Патогенен	Патогенен
	PMS2	rs146176004	CG	Патогенен	Патогенен
№16	*BRCA1*	rs80356964	AG	Патогенен	Патогенен
	PALB2	rs45551636	CT	Патогенен	Патогенен
№17	*BRCA1*	rs1799950	TC	Патогенен	Толерантен
№18	*BRCA1*	rs16941	CC	Патогенен	Патогенен
	BRCA1	rs1799950	CT	Патогенен	Толерантен

BRCA1

При анализа на данните се установиха три патогенни SNP варианти в *BRCA1* (rs16941, rs80356964 и rs1799950).

Алелните варианти на rs16941 се срещат, като хомозиготи по C-алел са 16,7% (n=3), а хетерозиготи – TC генотип са 16,7% (n=3), и хомозиготи по TT – генотип са 66,6% (n=12) от изследваните пациенти (Фиг. 7).

Вариантът rs1799950 се среща в хетерозиготно състояние -TC генотип, при 27,8% (n=5) от пациентите и хомозиготи по TT генотип 72,2% (n=13) (Фиг. 8).

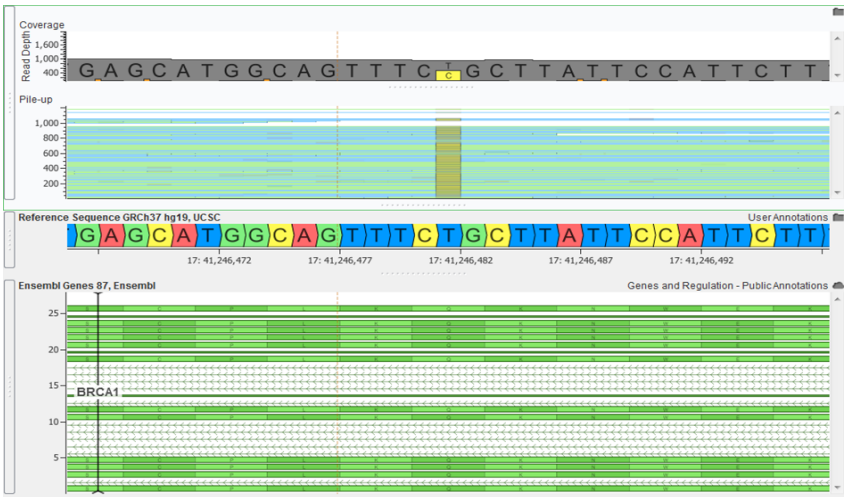
Вариантът rs80356964 се среща в хетерозиготно състояние – GA генотип, при 22,2% (n=4) от пациентите, хомозиготи носители на GG – алела са 66,6% (n=12) (Фиг. 9).

Според стойността на PROVEAN коефициента, трите варианта се класифицират като увреждащи функцията. Според SIFT анализа един от

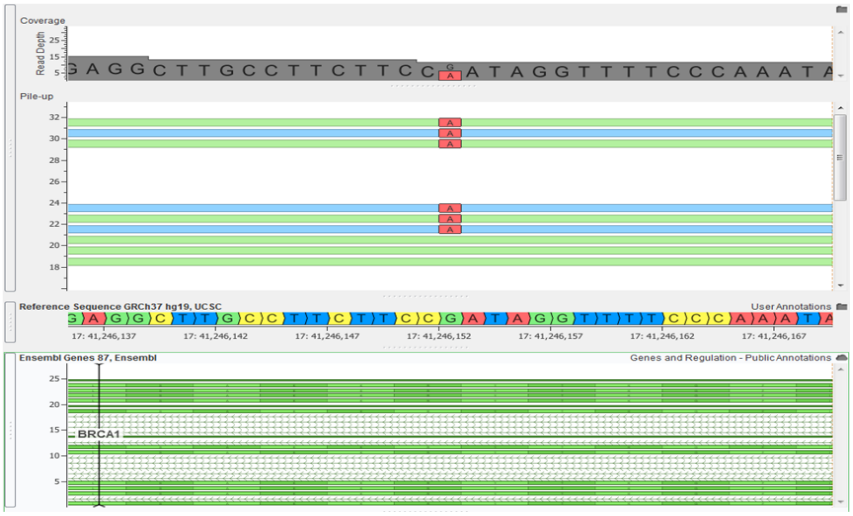
вариантите (rs1799950) се определя като „толерантен,“ а другите два варианта rs16941 и rs80356964 се класифицират като „вредни“.



Фигура 7: SNP вариантът rs16941 в хомозиготно състояние CC генотип



Фигура 8: SNP вариантът rs1799950 в хетерозиготно състояние TC генотип



Фигура 9: SNP вариантът rs80356964 в хетерозиготно състояние GA генотип

PMS2

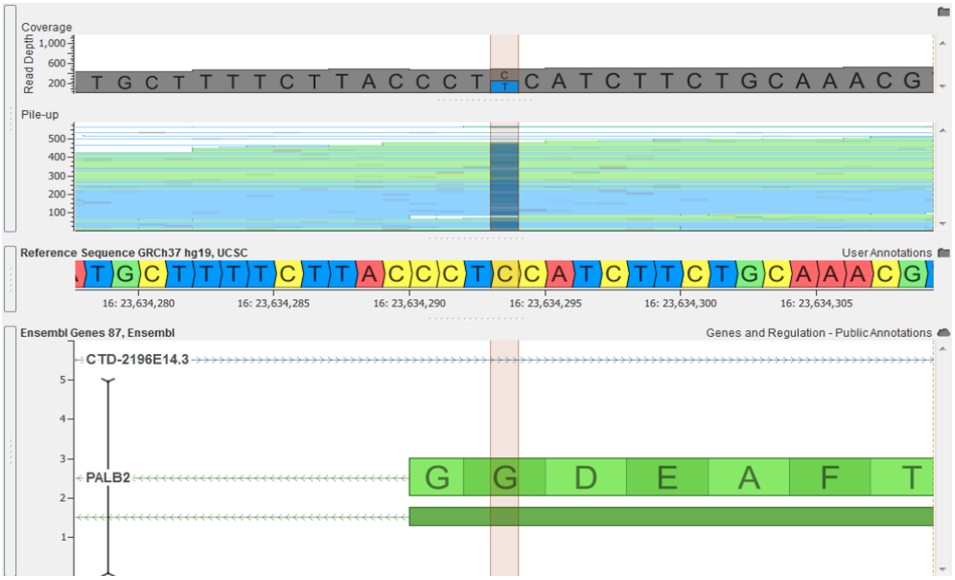
Вариантът rs146176004 се среща при 11,1% (n=2) от туморните проби в хетерозиготно състояние – CG генотип. И двата софтуера за предикция - PROVEAN и SIFT, определят този вариант като „вреден“, влияещ негативно върху функцията .

PALB2

Два SNP варианта (rs75023630 и rs45551636) са открити в *PALB2* гена, които са „вредни“ и влияещи негативно върху функцията, според двата предиктора.

Вариантът rs75023630 се среща при 5,6% (n=1) от туморните проби в хетерозиготно състояние – CA генотип, останалите туморни проби са хомозиготи CC генотип 94,4% (n=17) (Фиг. 10).

Вариантът rs45551636 също така се среща при 5,6% (n=1) в хетерозиготно състояние - CT генотип, останалите туморни проби са хомозиготи CC генотип 94,4% (n=17).



Фигура 10: SNP вариантът rs75023630 в хетерозиготно състояние CT генотип

MLH1

Вариантът rs63751049 се среща при 5,6% (n=1) в хетерозиготно състояние - CT генотип. Според предикцията по PROVEAN, този вариант се определя като увреждащ функцията, а според SIFT се класифицира като „вреден“.

1.2. Фармакогенетични варианти при НДКБК, определени чрез цеквениране с карциномен панел

Установените при секвенирането варианти бяха подложени на анализ за потенциалната роля като фармакогенетични дефекти при химиотерапия.

Всички открити варианти са анотирани по сенс веригата. За референтна база данни, включително и за алелните честотите в европейска популация, беше използвана PharmGKB.

Установихме 5 фармакогенетични варианта (rs1799793, rs13181, rs1042522, rs2228001, rs2227983) в 4 гена (*ERCC2*, *TP53*, *XPC* и *EGFR*), в туморите на изследваните пациенти, които имат отношение към ефекта на различни противотуморни препарати.

Приложихме PROVEAN и SIFT алгоритмите за определяне на ефекта на вариантите върху биологичната функция на протеина (Табл.7).

Таблица 7: Фармакогенетични варианти установени в туморите на пациенти с НДКБК

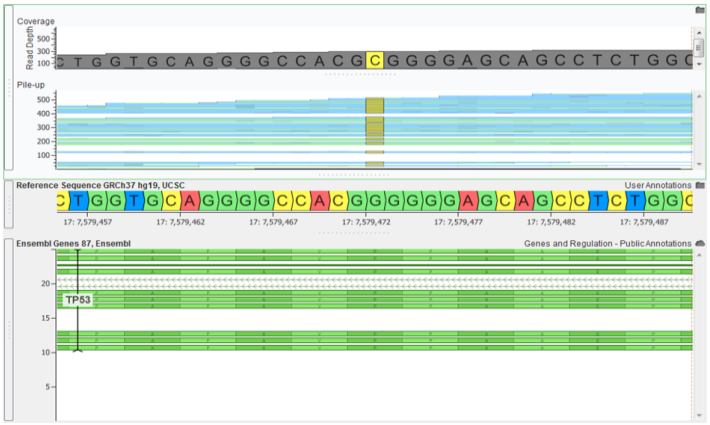
Пациенти	**Гени**	**Варианти**	**Генотип**	**PROVEAN cut off = -2.5**	**SIFT cut off = 0.05**
№1	*ERCC2*	rs1799793	CT	Патогенен	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№2	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	TP53	rs1042522	GC	Неутрален	Толерантен
	XPC	rs2228001	TT	Неутрален	Толерантен
№3	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	ERCC2	rs1799793	CT	Патогенен	Толерантен
	ERCC2	rs13181	GG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№4	*ERCC2*	rs1799793	TT	Патогенен	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№5	*ERCC2*	rs1799793	CT	Патогенен	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№6	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
№7	*ERCC2*	rs13181	TG	Неутрален	Толерантен
	TP53	rs1042522	GC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№8	*EGFR*	rs2227983	AA	Неутрален	Толерантен
	ERCC2	rs13181	TG	Неутрален	Толерантен
	TP53	rs1042522	CG	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№9	*TP53*	rs1042522	CC	Неутрален	Толерантен
№10	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	XPC	rs2228001	TT	Неутрален	Толерантен
№11	*ERCC2*	rs13181	TG	Неутрален	Толерантен

	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	TT	Неутрален	Толерантен
№12	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	ERCC2	rs13181	TG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
№13	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	ERCC2	rs1799793	TT	Патогенен	Толерантен
	ERCC2	rs13181	GG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№14	*ERCC2*	rs1799793	CT	Патогенен	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	TT	Неутрален	Толерантен
№15	*EGFR*	rs2227983	AA	Неутрален	Толерантен
	ERCC2	rs13181	TG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
№16	*EGFR*	rs2227983	AG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
№17	*ERCC2*	rs13181	GG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен
№18	*ERCC2*	rs13181	TG	Неутрален	Толерантен
	TP53	rs1042522	CC	Неутрален	Толерантен
	XPC	rs2228001	GT	Неутрален	Толерантен

Според предикцията по SIFT, всички открити фармакогенетични варианти са класифицирани като „толерантни“, докато PROVEAN определи един от вариантите (rs1799793 в *ERCC2*) като „вреден“, останалите варианти са категоризирани като неутрални.

TP53

SNP вариантът rs1042522 в TP53 гена се среща в 88,9% (n=16). Присъствието на C-алел (честота на C-алел в европейска популация - 0.72) в TP53 е открит при 11,1% (n=2) в хетерозиготно състояние - генотип GC, а 77,8% (n=14) са хомозиготи по C-алел (Фиг. 11). Този SNP-вариант асоциира с ефекта от химиотерапия с паклитаксел, циклофосфамид, флуроурацил и цисплатина. Генотип CC е асоцииран с намален риск от токсичност при лечение с циклофосфамид и флуроурацил. Наличието на G-алел се свързва с повишаване на риска за токсичност при тези терапии.

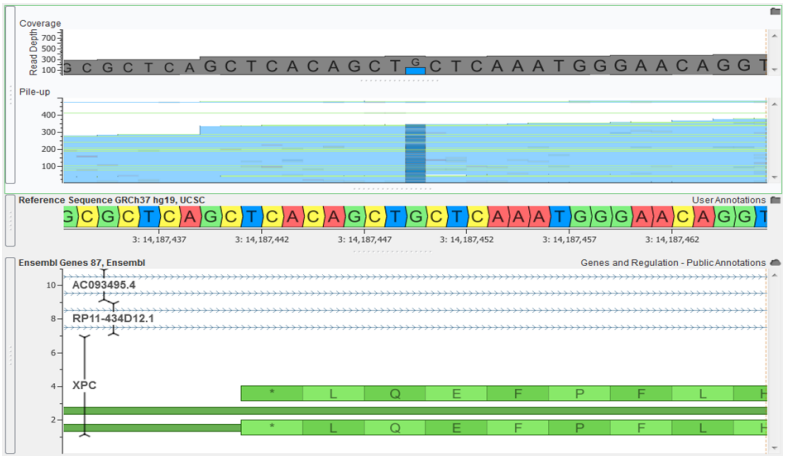


Фигура 11: SNP вариантът rs1042522 в хомозиготно състояние CC генотип

XPC

SNP вариантът rs2228001 в *XPC* гена присъства в 72,2% (n=13) от анализираните туморни проби. Присъствието на T-алел (честота в европейската популация 0,60)

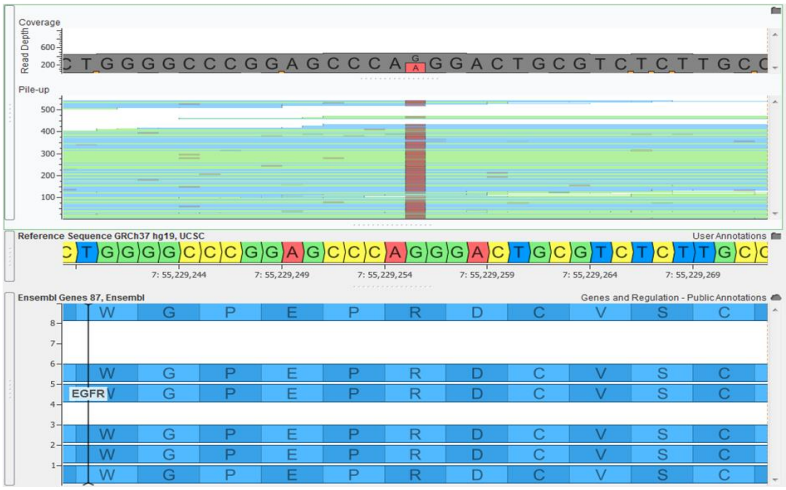
в *XPC* гена е открито при 50% (n=9) от пациентите в хетерозиготна форма – генотип GT и при 22,2% (n=4) се наблюдава хомозиготен TT генотип (Фиг. 12). Този вариант асоциира с ефекта при химиотерапия с цисплатина. Генотип TT е асоцииран с намален риск от токсичност при лечение с цисплатина спрямо хомозиготите по G-алел и хетерозиготите - GT.



Фигура 12: SNP вариантът rs2228001 в хетерозиготно състояние GT генотип

EGFR

В *EGFR* гена е установен фармакогенетичен вариант rs2227983 в хомозиготно и хетерозиготно състояние на A-алела (честота в европейската популация - 0.28). Туморите с генотип на хетерозиготи AG са 38,9% (n=7), а хомозиготите по A-алела (AA) са 11,1% (n=2) от пациентите (Фиг. 13). Генотип AA и генотип AG се асоциират с намален риск от нежелани лекарствени реакция свързани с кожна токсичност при лечение с *EGFR* TKIs.



Фигура 13: SNP вариантът rs2227983 в хетерозиготно състояние AG генотип

ERCC2

В *ERCC2* гена бяха открити два варианта: rs1799793 и rs13181.

Т-алел на rs1799793 (честота 0.36 в европейска популация) се открива при 33,3% (n=6) от туморите- в 11,1% (n=2) е в хомозиготно състояние - ТТ и в 22,2% (n=4) е в хетерозиготно състояние – СТ. При пациентите след лечение с цисплатин, които са с туморен генотип ТТ и СТ за rs1799793 се наблюдава удължен период без прогресия на заболяването в сравнение с тези с генотип СС.

При 50% (n=9) от пациентите се открива вариант rs13181 с G алел (честота в европейската популация - 0,36), като при 33,3%(n=6) е в хетерозиготно състояние – TG генотип и при 16,6% (n=3) в хомозиготно състояние – GG генотип (Фиг. 14). Пациенти с туморен генотип GG за rs13181 имат намалена преживяемост при лечение с платинови съединения спрямо пациентите с ТТ генотип.



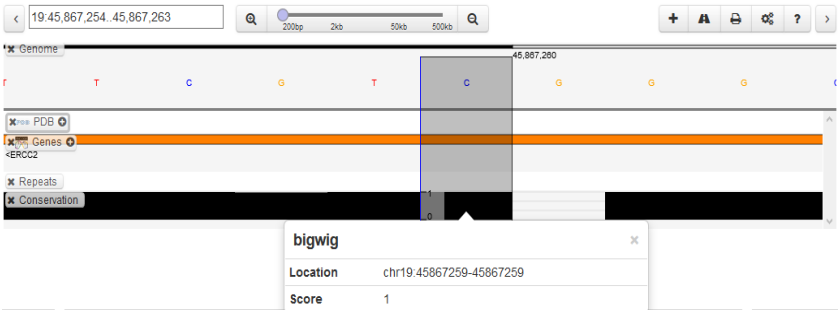
Фигура 14: SNP вариантът rs13181 в хомозиготно състояние GG генотип

In silico анализ на варианта rs1799793

Кодът на варианта според dbSNP база данни е NC_000019.9; NM_000400.3:c.934G>A, с позиция в генома g.45867259, алели C>T, протеин

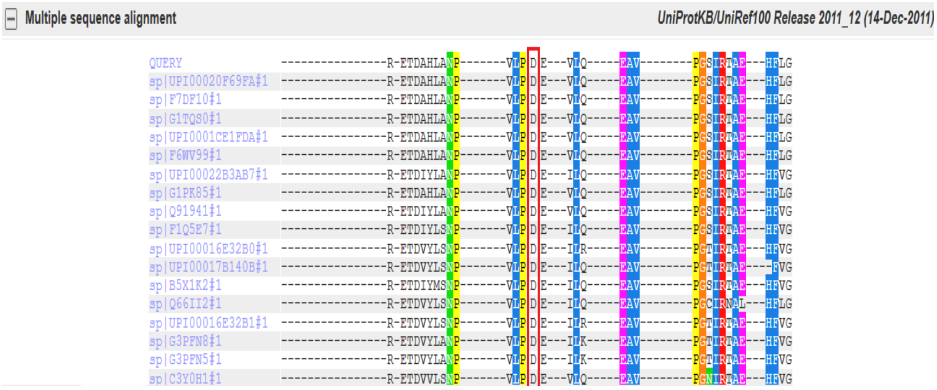
NP_000391.1:p.Asp312Asn; код на протеина според UniProt база данни P18074, D312N (rs1799793).

Цитозинът е еволюционно високо консервативна нуклеотидна база (PhastCons показател – 1) (Фиг. 15).



Фигура 15: Секвенцията на *ERCC2* гена в хромозома 19; Т алелът е заменен със С алел

Откритият миссенс вариант води до замяна на еволюционно високо консервативната аспарагинова аминокиселина (Asp) с аминокиселина аспарагин (Asn) (Фигура 16).



Фигура 16: Замяната на аспарагинова киселина (D) с аспарагин (N) на 312 позиция

Според различни предиктори (SIFT и PROVEAN), за този вариант има противоречива информация (Табл. 8).

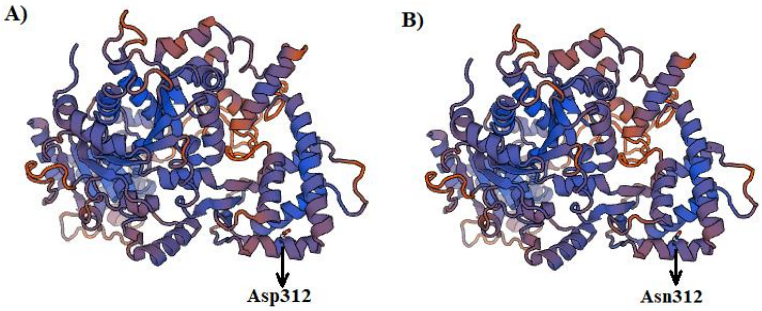
Таблица 8: Вариантът (rs1799793 в *ERCC2*) е толерантен (Tolerated) според предикцията по SIFT; според предикцията PROVEAN се определя като deleterious

	Prediction	Score
UniProt	Non-disease (reviewed)	NA
PROVEAN	Deleterious	-2.80
SIFT	Tolerated	0.174
PolyPhen-2 HumDiv	benign	0.053
PolyPhen-2 HumVar	benign	0.051

Вариантът (c.934G>A) води до потенциално нарушение на нативна нуклеотидна последователност (5′—TGCCCGA—3′) в екзон 10, която е необходима за свързване на пре-мРНК сплайсинг-фактор SF2/ASF (IgM-BRCA1) и създава мутантен екзон сплайс репресорен мотив (ESS:

5′–CTGCCCAA–3′). Евентуалното нарушение на нормалния сплайсинг може да доведе до изпускане на екзон 10 от *ERCC2* транскрипта (ENST00000391945.8). Като ефект, мутацията (c.934G>A) води до изместване на рамката на четене (фреймшифт) и преждевременно формиране на стоп кодон в екзон 12 (p. Arg272ArgfsTer376). При in silico моделиране установихме, че това нарушение не води до промяна в третичната структура на протеина (Фигура 17), но води до промяна в афинитета на протеина към други молекули в клетката, поради смяната на заряда на аминокиселината.

Полиморфизмът *ERCC2* Asp312Asn оказва влияние върху репарацията на ДНК чрез NER (Nucleotide excision repair) пътя. Увреденият механизъм за ДНК репарация е рисков фактор за развитието на рак.



Фигура 17: Структурата на домейна на *ERCC2*; полиморфизмът *ERCC2* Asp312Asn

2. Мутационен спектър на *EGFR* гена определен чрез експресионен PCR анализ в реално време за избор на таргетна терапия на НДКБК

2.1. Клинико-патологична характеристика на изследваните пациенти

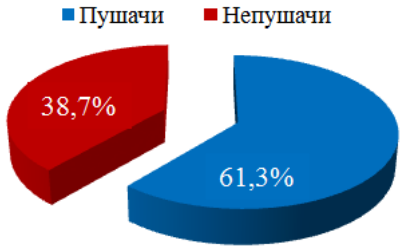
За период от 5 години са събрани 551 FFPE-туморни проби от пациенти диагностицирани с НДКБД в IIIB и IV стадий след предоставено подписано писмено информирано съгласие.

Средната възраст на пациентите в изследваната кохорта е 63 години, като възрастта към момента на диагностициране на заболяването варира между 24 и 84 години. В тази извадка се наблюдава превес на мъжкия пол - мъжете включени в анализа са 73,8% (n=407), а жените 26,2% (n=144). Съотношението мъже: жени в тестваната група е 2,83:1 (Фигура 18).



Фигура 18: Разпределение на пациентите по пол

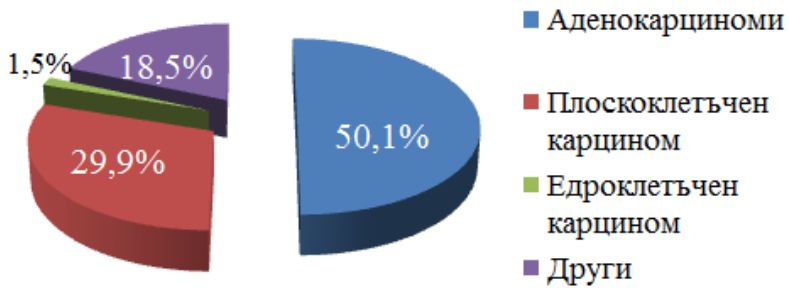
В изследваната кохорта се наблюдава предоминиране на пациентите, които са или са били дългогодишни пушачи - 61,3% (n=338). Непушачите са съответно 38,7% (n=213) (Фиг.19).



Фигура 19: Разпределението по групи на пушачи и непушачи

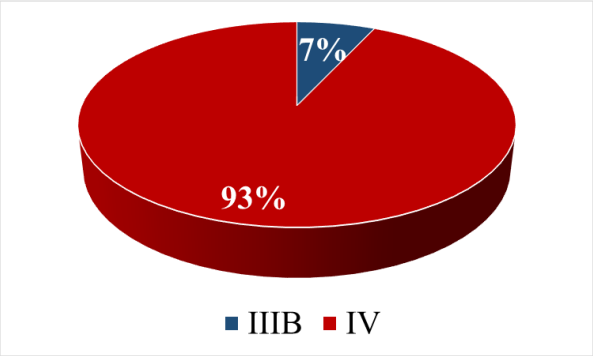
Всички туморни проби са преминали през хистологичната оценка и са класифицирани от клиничен патолог в 4 групи.

На фиг. 20 е представено разпредлението на туморните проби въз основа на хистологичния тип. Аденокарциномите представляват 50,1% (n=276) от всички тумори в изслседваната кохорта. Делът на плоскоклетъчни карциноми е 29,9% (n=165). Едроклетъчни карциноми имат най-малък дял - 1,5% (n=8). Останалите хистологични варианти съставляват групата „други“ - 18,5% (n=102).



Фигура 20: Хистологична класификиция на пробите

Разпределението по стадии е както следва: стадий IIIB - 6,9% (n=38); стадий IV- 93,1% (n= 513) (Фиг. 21).



Фигура 21: Разпределението на туморите по стадий

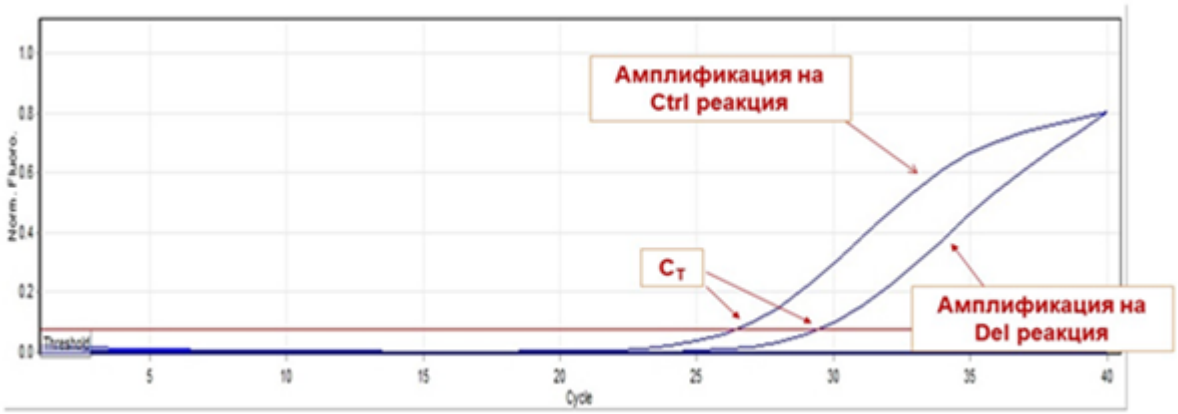
2.2. Мутации в *EGFR* гена за таргетна терапия на НДКБК

Мутациионният статус на *EGFR* в туморните проби беше определен чрез теста „therascreen *EGFR* assay“, с който се изследват 29 мутации в гена. Установени са генни мутации в *EGFR* в 10% от изследваните туморни проби с НДКБК (n=55) (табл. 9).

Таблица 9: Типовете на откритите мутации в *EGFR* гена

Открити Мутации	Пациенти	Процент
Делеции в екзон 19	31	56.40%
L858R	20	36.40%
Инсерции в екзон 20	2	3.60%
Делеции в екзон 19 и T790M	1	1.80%
Делеции в екзон 19 и S768I	1	1.80%

EGFR делеции в екзон 19 се откриват при 23 белодробни аденокарциноми, 3 плоскоклетъчни карциноми и 5 едроклетъчни карциноми (Фигура 22).



Фигура 22: PCR амплификацията на проба с положителна *EGFR* мутация (делеции в екзон 19 – Deletions Reaction Mix) и амплификация на контролата за същата проба (Control Reaction Mix);

Стойностите на ΔCT за пробата се изчисляват като разликата между CT стойността на амплифицираната мутация и контролен CT стойност същата проба:

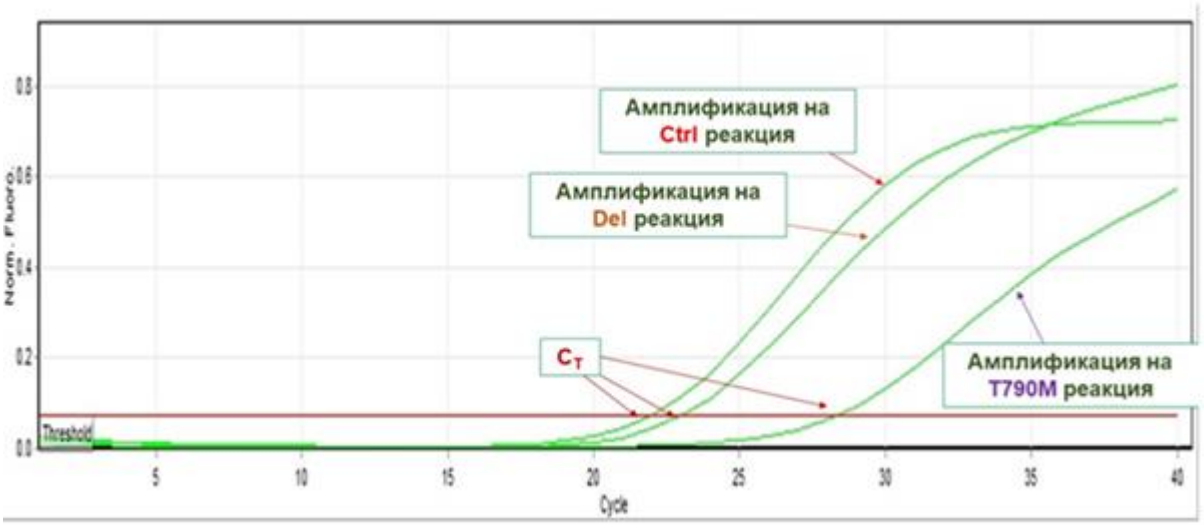
ΔCT = mutation (мутация) CT – control (контрола) CT

ΔCT (2.97) = 29.42 (мутация Del) CT – 26.45 (контрола Ctrl) CT

Стойността на ΔCT за пробата се сравнява с cut-off стойността за тази мутация съгласно указанията на производителя (Qiagen) (Табл. 5 на страница 70); При ΔCT стойност под cut-off (9.06) пробата се счита за положителна мутация за делеции в екзон 19;

Мутация L858R в *EGFR* гена е установена при 17 белодробни аденокарциноми, 3 плоскоклетъчни карциноми и 2 едроклетъчни карциноми. Мутация с инсерции в екзон 20, определяща нечувствителност към *EGFR* тирозин-киназни инхибитори е открита в два тумора – един аденокарцином и един плоскоклетъчен карцином.

При двама от пациентите се откриват две мутации едновременно. При единия пациент с плоскоклетъчен карцином са определени: делеции в екзон 19 и точкова мутация T790M в екзон 20, която е нечувствителна към тирозин-киназни инхибитори (Фигура 23). При втория пациент с аденокарцином се откри делеции в екзон 19 и точковата мутация S768I в екзон 20, която е с намалена чувствителност към тирозин-киназните инхибитори.



Фигура 23: PCR амплификацията на проба с два вида мутации в *EGFR* гена (делеции в екзон 19 – Deletions Reaction Mix и T790M точкова мутация) и амплификация на контролата за същата проба (Control Reaction Mix)

Стойностите на ΔCT за пробата се изчисляват като разликата между CT стойността на амплифицираната мутация и контролна CT стойност на същата проба:

ΔCT = mutation (мутация) CT – control (контрола) CT

ΔCT (1.02) = 23.07 (мутация Del) CT – 22.05 (контрола Ctrl) CT

ΔCT (6.38) = 28.43 (мутация T790M) CT – 22.05 (контрола Ctrl) CT

Стойността на ΔCT за пробата се сравнява с cut-off стойността за тази мутация в таблица съгласно указанията на производителя (Qiagen) (Табл. 5 на страница 70); При ΔCT стойност на пробата под cut-off стойността (9.06), пробата се счита за открита положителна мутация за делеции в екзон 19; ΔCT стойността на пробата за втората мутация е в cut-off стойността (6.38), пробата се счита за открита положителна мутация за T790M точкова мутация;

С най-висока честота са делециите в екзон 19 - 56.40% (n=31), следвани от мутацията L858R в екзон 21 - 36.40% (n=20). Общо те представляват около 92.7% от всички EGFR активиращи мутации. В нашата кохорта тези два вида мутации са установени само в аденокарциноми. Всички делеции в екозон 19 са в проби на пациенти в IV-ти стадий. Мутацията L858R предоминиращо се открива в IV стадий в нашата извадка – 90% (n=18). Останалите мутации са по-редки: инсерции в екзон 20 - 3,6% (n=2), двойна мутация - T790M, комбинирана с делеция в екзон 19 - 1,8% (n=1) и двойна мутация - S768I с делеция в екзон 19 - 1,8% (n=1). Не са

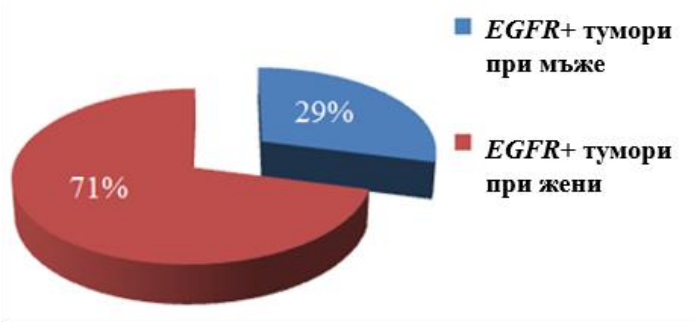
доказани останалите мутации, включени в теста: L861Q, S768I, G719A, G719S, G719C.

Проведен е анализ за честотата на мутациите в туморните проби в зависимост от:

- Пол;
- Туморна хистология ;
- Тютюнопушене.

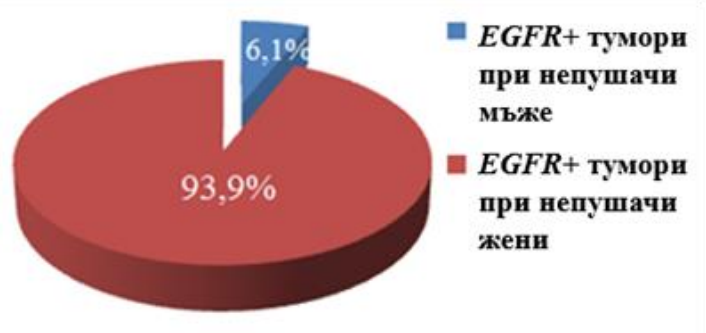
2.3. Анализ на фенотип/генотип корелации

EGFR положителен мутационен статус е установен при 55 тумори, които разпределени по пол са 71% при жени (n=39) и 29% при мъже (n=16)(P <0,001) (Фиг. 24).



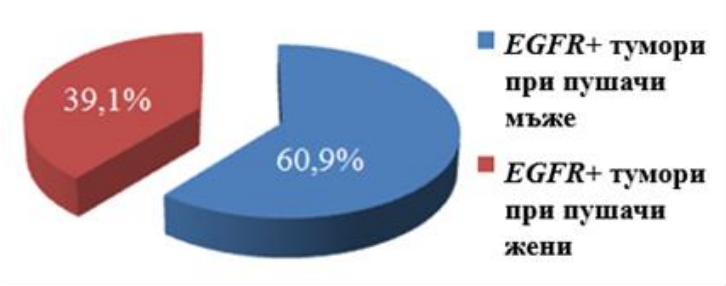
Фигура 24: Разпределението на *EGFR* позитивни тумори според пола на пациентите

В групата на непушачите (n=32) *EGFR* позитивни тумори са установени предимно при жени 93,9% (n=30) (фиг.25).



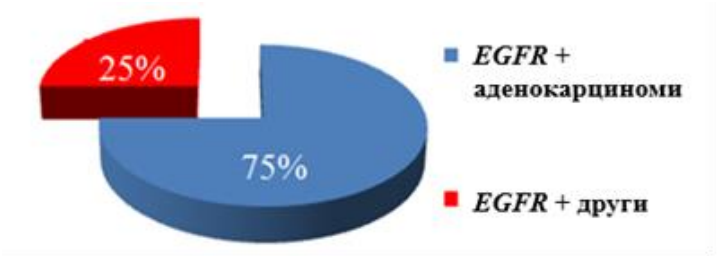
Фигура 25: Разпределение спрямо пола на *EGFR* позитивни тумори при пациенти непушачи

В групата на пушачите (n=23), *EGFR* позитивните тумори предоминират при мъжете - 60,9% (n=14) спрямо жените 39,1% (n=9) (P <0,003) (Фиг. 26).



Фигура 26: Разпределение на *EGFR* позитивни тумори спрямо пола на пациенти пушачи

EGFR позитивни тумори (n=55) са значимо по-чести при аденокарциноми 75% (n=41) спрямо другите хистологични типове 25% (n=14) (P < 0.001) (плоскоклетъчен карцином и едроклетъчен карцином), (фиг. 27).



Фигура 27: Разпределение на *EGFR*-положителните случаи според хистологичния вариант на туморите, при които са установени

На табл.10 са представени обобщените данни за *EGFR* мутациите в изследваната кохорта пациенти с НДКБК по отношение на клинико-патологични данни.

Таблица 10. Обобщени данни за изследваните пациенти за *EGFR* мутационен статус

Клинични и килинопатологични характеристики	**Брой пациенти**	**Брой пациенти с див тип ген или с мутация**		**P-value**
		Mut	**WT**	
Пол				$1x10^{-15}$
Мъжки	407	16	391	
Женки	144	39	105	
Хистология				0.000232
Аденокарцином	276	41	235	
Други хистологични типове	275	14	261	
Тютюнопушене				0.002807
Пушачи	338	23	315	
Непушачи	213	32	181	

3. Данни от NGS секвениране с карциномен панел на ДНК-и от кръвни проби на пациенти с фамилен папиларен тироиден карцином

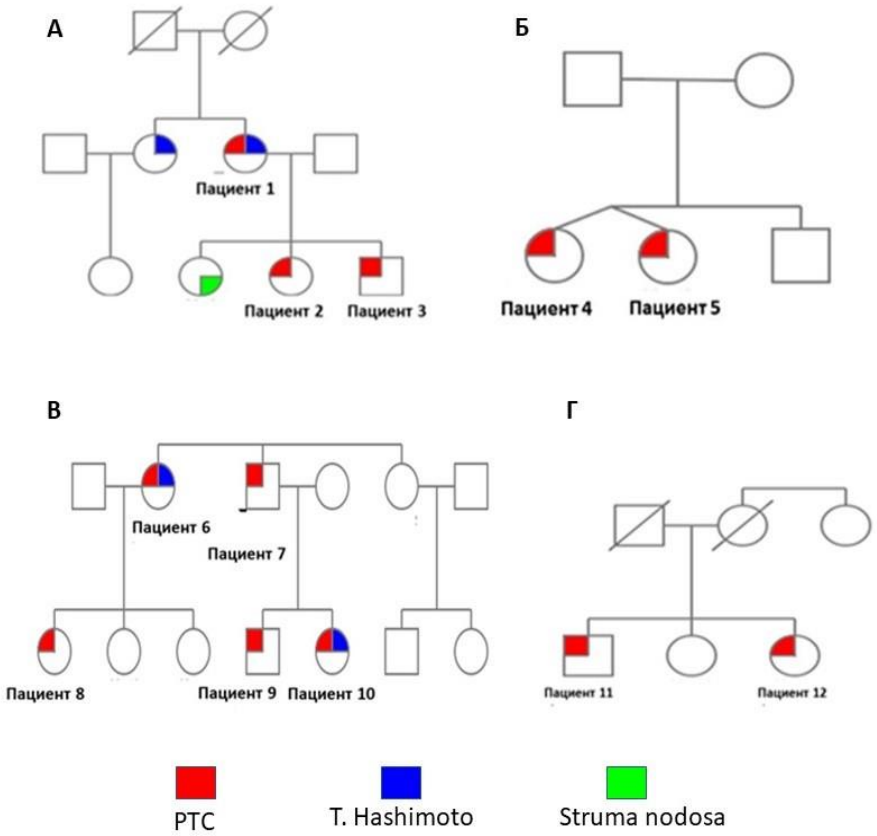
3.1. Родословия на четири фамилии с ФПТК

Извършено е NGS секвениране с карциномен панел на ДНК-и, изолирани от кръвни проби на 12 пациента от четири семейства (I,II,III,IV) с фамилна форма на папиларен карцином на щитовидната жлеза (ФПТК). (Табл. 11)

Таблица 11. Секвенирани пациенти с фамилен папиларен тироиден карцином от четири семейства

Фамилии	Пациенти	Туморен статус
I	№1	pT4mN0M0
	№2.	pT1b N0M0
	№3	pT1a N0M0
II	№4,	pT4m N1M0
	№5.	pT2N1bM0
III	№6,	pT1amN0M0
	№7.	pT1bmN0M0
	№8,	pT2m N1M0
	№9,	pT1bmN1M0
	№10	pT3mN1aM0
IV	№11,	pT2N1aM0
	№12	pT2 N0 M0

На фигури 28 А, Б, В, Г са представени родословните дървета на четирите семейства с различните заболявания на щитовидната жлеза.



Фигура 28: Родословия на четири фамилии с ФПТК;

3.2. Патогенни варианти в гени, свързани с карциногенеза, определени при NGS секвениране на пациенти с ФПТК

Извършено е NGS секвениране с TruSight Cancer Sequencing Panel, Illumina на ДНК-и, изолирани от кръвни проби на 12 пациенти с фамилен папиларен тироиден карцином. Всички открити варианти са анотирани по сенс веригата. Анализирани са варианти в гени, свързани с карциногенезата. Определени са 8 патогенни варианти в 3 гени (*MET*, *RET*, *MLH1*), които се срещат в повече от една фамилия или са редки варианти (Табл. 12 и Табл. 13)

Таблица 12: Патогенни варианти, които се срещат в повече от една фамилия

Гени	Варианти	а	б	в	г	Болни брой	Sift	PolyPhen
MET	rs56391007					3	патогенни	вероятно патогенни
RET	rs181856591					3	патогенни	вероятно патогенни
MLH1	NP_000240.1:p.Glu23Gly					4	патогенни	вероятно патогенни

Таблица 13: Редки патогенни варианти

Гени	Варианти	а	б	в	г	Болни брой	Sift	PolyPhen

MET	NP_001120972.1:p.Leu1080Met					1	патогенни	вероятно патогенни
MET	NP_001120972.1:p.Phe1234Ser					1	патогенни	вероятно патогенни
MET	NP_001120972.1:p.Pro712Gln					1	патогенни	вероятно патогенни
MET	NP_001120972.1:p.Ser823Tyr					1	патогенни	вероятно патогенни
MLH1	NP_000240.1:p.Gly634Val					1	патогенни	вероятно патогенни

От 12 пациенти, 9 са носители на патогенни варианти. От тях 5 пациента имат повече от 1 мутантен ген (Табл. 14).

Таблица 14: Индивидуални патогенни варианти

Пациенти	Гени	Варианти	Генотип	Sift (cutoff=0.05)	PolyPhen
№4	*MET*	NP_001120972.1:p.Leu1080Met	CA	патогенни	вероятно патогенни
	RET	rs181856591	CA	патогенни	вероятно патогенни
№5	*MET*	rs56391007	CT	патогенни	вероятно патогенни
	RET	rs181856591	CA	патогенни	вероятно патогенни
№6	*MET*	NP_001120972.1:p.Pro712Gln	CA	патогенни	вероятно патогенни
	MLH1	NP_000240.1:p.Glu23Gly	AG	патогенни	вероятно патогенни
№7	*MET*	NP_001120972.1:p.Ser823Tyr	CA	патогенни	вероятно патогенни
№8	*MET*	rs56391007	CT	патогенни	вероятно патогенни
	MLH1	NP_000240.1:p.Glu23Gly	AG	патогенни	вероятно патогенни
	RET	rs181856591	CA	патогенни	вероятно патогенни
№9	*MET*	rs56391007	CT	патогенни	вероятно патогенни
№10	*MLH1*	NP_000240.1:p.Glu23Gly	AG	патогенни	вероятно патогенни
№11	*MET*	NP_001120972.1:p.Phe1234Ser	TC	патогенни	вероятно патогенни
	MLH1	NP_000240.1:p.Gly63	GT	патогенни	вероятно

		4Val			патогенни
№12	*MLH1*	NP_000240.1:p.Glu23Gly	AG	патогенни	вероятно патогенни

RET

Вариантът rs181856591 се среща при 25% от пациенти (n=3) в хетерозиготно състояние – CA генотип. Според предикцията по SIFT, този вариант се класифицира като патогенен или вреден (deleterious) и според PolyPhen се определя като вероятно патогенен (probably damaging).

Патогенни мутации в този ген водят до развитието на наследствени ракови синдроми -множествената ендокринна неоплазия тип 1 и тип 2 и медуларен тироиден карцином.

MET

Вариантът rs56391007 се среща при 25% от пациентите (n=3) в хетерозиготно състояние – CT генотип. Според предикцията по SIFT, този вариант се класифицира като патогенен (deleterious) и според PolyPhen се определя като вероятно патогенен (probably damaging).

Варианти в този ген са асоциирани с папиларен карцином на бъбреците и хепатоцелуларен карцином. В 33% (n=4) от пациентите са намерени SNPs варианти в *MET*, които не са анотирани в базите данни и няма докладвана асоциация между тях и патологичен фенотип (NP_001120972.1:p.Leu1080Met, NP_001120972.1:p.Phe1234Ser, NP_001120972.1:p.Pro712Gln, NP_001120972.1:p.Ser823Tyr). Според предикцията по SIFT, тези варианти се класифицират като патогенен (deleterious) и според PolyPhen се определят като вероятно патогенен (probably damaging).

MLH1

В общо 41,7% (n=5) от пациентите се откриха варианти в този ген, като и двата са неанотирани до сега . При 33,3% (n=4) от изследваните пациенти, се открива вариантът NP_000240.1:p.Glu23Gly. Според SIFT стойността той се определя като вреден (deleterious) и според PolyPhen се определя като възможно вреден (possibly damaging). В 8,3% (n=1) е открит също неанотиран вариант - NP_000240.1:p.Gly634Val, който според предикцията по SIFT се класифицира като патогенен (deleterious) и според PolyPhen се определя като вероятно патогенен (probably damaging)..

Патогенни мутации в този ген са докладвани, че са асоциирани с колоректален карцином.

3.3. Фармакогенетични варианти при ФПТК, определени чрез секвениране с карциномен панел

Откритите при пациентите генетични варианти бяха подложени на анализ за ролята им като фармакогенетични варианти свързани с терапията със специфични химиотерапевтици. Всички открити варианти са анотирани по сенс веригата. За референтна база данни, включително и за алелните

честотите в европейска популация, беше използвана PharmGKB. Установи се, че 6 от вариантите имат фармакогенетично значение за определени химиотерапевтици. Тези варианти са локализирани в 6 гена - ERCC2 (rs13181), TP53 (rs1042522), XPC (rs2228001), EGFR (rs2227983), ERCC5 (rs17655) и RET (rs1799939) (Табл. 15). Според двата предиктора SIFT и PolyPhen тези фармакогенетични варианти са съответно толерантни и доброкачествени.

Таблица 15: Варианти асоциирани с терапевтичния отговор към химиотерапевтици при пациенти с фамилен папиларен тироиден карцином

Пациенти	Гени	Варианти	Генотип	SIFT	PolyPhen
№1	*EGFR*	rs2227983	AA	Толерантен	Доброкачествен
	ERCC2	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	GC	Толерантен	Доброкачествен
	XPC	rs2228001	GT	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен
№2	*EGFR*	rs2227983	GA	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
№3	*EGFR*	rs2227983	GA	Толерантен	Доброкачествен
	ERCC2	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	GC	Толерантен	Доброкачествен
№4	*EGFR*	rs2227983	GA	Толерантен	Доброкачествен
	ERCC2	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
	XPC	rs2228001	GT	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен
№5	*ERCC2*	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
	XPC	rs2228001	TT	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен
№6	*ERCC5*	rs17655	GC	Толерантен	Доброкачествен
	TP53	rs1042522	GC	Толерантен	Доброкачествен
	XPC	rs2228001	GT	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен
№7	*ERCC5*	rs17655	GC	Толерантен	Доброкачествен
	ERCC2	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен
№8	*ERCC2*	rs13181	TG	Толерантен	Доброкачествен

	XPC	rs2228001	GT	Толерантен	Доброкачествен
№9	*ERCC5*	rs17655	GC	Толерантен	Доброкачествен
	EGFR	rs2227983	GA	Толерантен	Доброкачествен
	ERCC2	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
	XPC	rs2228001	TT	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен
№10	*ERCC5*	rs17655	GC	Толерантен	Доброкачествен
	ERCC2	rs13181	TG	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
	XPC	rs2228001	GT	Толерантен	Доброкачествен
№11	*EGFR*	rs2227983	GA	Толерантен	Доброкачествен
	TP53	rs1042522	CC	Толерантен	Доброкачествен
	XPC	rs2228001	GT	Толерантен	Доброкачествен
№12	*EGFR*	rs2227983	GA	Толерантен	Доброкачествен
	XPC	rs2228001	TT	Толерантен	Доброкачествен
	RET	rs1799939	GA	Толерантен	Доброкачествен

TP53

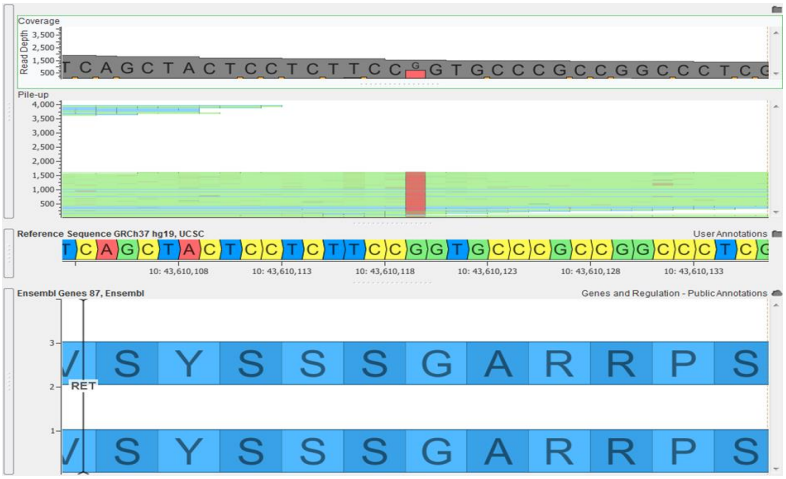
Резултатите показват, че SNP вариантът rs1042522 в TP53 се среща при 83,3% от нашите пациенти (n=10). C-алел (честота в европейска популация 0,71) е представен в хетерозиготно състояние (GC генотип) при 25% (n=3) и в хомозиготно състояние (CC) в 58,3% (n=7) от изследваните пациенти. Алел-C е асоцииран с намален риск за токсичност при лечение с циклофосфамид и флуроурацил в сравнение с алел-G.

ERCC2

SNP вариантът rs13181 в *ERCC2* гена се среща при 66,7% (n=8) от пациентите, като при всички е в хетерозиготно състояние – TG генотип. Честотата на G-алел в европейска популация се оценява на 0,36. При пациентите, лекувани с платинови съединения, се наблюдава по-висока преживяемост при TT генотипв сравнение с тези с GG генотип. В допълнение алел-G се асоциира с повишена токсичност на флуороурацил и леуковорин.

RET

Честотата на SNP варианта rs1799939 е 58,3% (n=7) при изследвани пациенти в хетерозиготно състояние – GA генотип и при хомозиготи GG генотип е 41,7% (n=5) (Фиг. 29). Честотата на A-алела в европейска популация е 0,19. Алел-A не се асоциира с по-висока токсичност при лечение със сунитиниб, но статистическата значимост на това твърдение е недостатъчна.



Фигура 29: SNP вариантът rs1799939 в хетерозиготно състояние GA генотип

EGFR

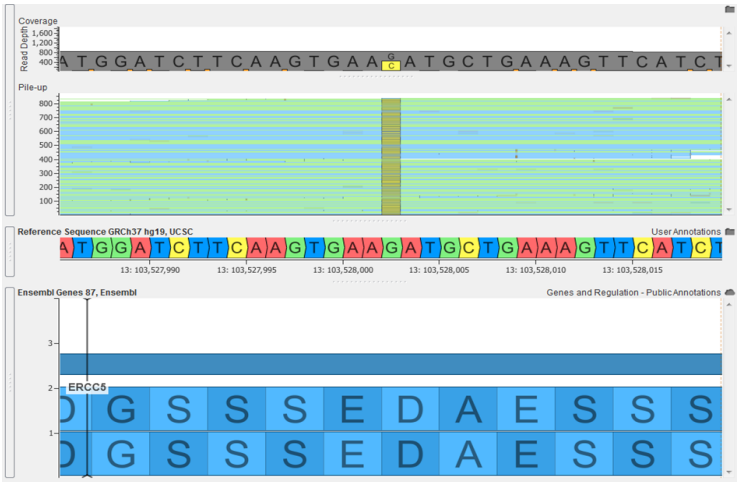
Резултатите показват, че SNP вариантът rs2227983 в *EGFR* гена се среща при 58,3% от нашите пациенти (n=7). Честотата на A-алела в европейска популация е 0,28. От изследваните пациенти 50% (n=6) са хетерозиготи (GA генотип) и 8,3% (n=1) са хомозигот по A-алел (AA генотип),. Генотип AA и генотип GA са асоциирани с намален риск от обрив при лечение с EGFR инхибитори.

XPC

SNP вариантът rs2228001 в гена *XPC* се среща при 75% от нашите пациенти (n=9). Честотата на T-алела в европейска популация е 0,6. От изследваните пациенти 50% (n=6) са хетерозиготи (GT генотип), и 25% (n=3) са хомозиги (TT генотип). Генотип GT и GG се асоциират с повишен риск от токсичност при лечение с цисплатин при сравнение с честотата на токсични реакции при пациенти с TT генотип.

ERCC5

SNP вариантът rs17655 в *ERCC5* гена присъства при 33,3% (n=4) от пациентите в хетерозиготна форма – GC генотип и 66,7% (n=8) от пациентите в хомозиготно състочние GG генотип (Фиг. 30). Честотата на C-алела в европейска популация е 0,25.. Данните за този вариант са противоречиви. Търсената асоциация между него и преживяемостта при лечение с цисплатина е неубедителна според литературата.



Фигура 30: SNP вариантът rs17655 в хетерозиготно състояние GC генотип

IV. Обсъждане

1. Обсъждане на патогенни варианти в гени, асоциирани с карциногенеза при НДКБК

При анализа на данните от NGS секвенирането на 18-те туморни проби се установиха 7 типа патогенни варианти (според PolyPhen и SIFT предикторите), общо 21 на брой, разпределени в четири гена свързани с туморогенезата - BRCA1, PMS2, MLH1 и PALB2.

BRCA1

BRCA1 е тумор-супресорен ген участващ в процесите на ДНК-репарация и регулация на транскрипционната активност в отговор на ДНК-увреждания. BRCA протеините са необходими за поддържане на хромозомна стабилност, като по този начин осъществяват протективна функция спрямо геномни увреждания. Нови данни показват също така, че BRCA регулират транскрипцията на гени, участващи в репарацията на ДНК, клетъчния цикъл и апоптозата. BRCA1 е един от основните гени, мутации, в които водят до развитие на наследствен рак на млечната жлеза и овариума, а варианти с малък самостоятелен ефект в него увеличават риска за развитие на заболяване при спорадичните и фамилни форми на РМЖ. Мутации в този ген също се асоциират още със серозен папиларен карцином на перитонеума и рак на простатата.

Литературните данни подкрепят клиничното значение на вариантите с малък ефект за развитието на фамилната форма на РМЖ. SNP варианти с малък индивидуален рисков потенциал са описани в *BRCA1*, *BRCA2*, *ATM*, *CHEK2* и *TP53*. Те имат кумулативен ефект и едновременното присъствие на всички или някои от тях води до клинично значимо увеличение на риска за РМЖ. Pharoah et al, и Rahman, et al доказват, че варианти в тези гени, при достигане на определен надпрагов брой и в комбинация с определени фактори на средата, стила на живот и репродуктивната история могат да доведат до развитие на заболяването (P = 0,005). Анализът на Johnson et al. също демонстрира ясната тенденция за увеличаване на тежестта и

ролята на вариантите с малък ефект при увеличаване на броя изследвани генетични варианти (P = 0.00004). Всеки потенциално функционален SNP вариант в тези гени има значение, тъй като при акумулиране на достатъчен брой, влияе на проявата на фенотипа.

При анализа на нашите данни се установиха три патогенни SNP варианти в *BRCA1* -rs16941, rs80356964 и rs1799950.

В нашата извадка SNP вариантът rs16941 в *BRCA1* гена се открива в хетерозиготно състояние - генотип TC при 16,7% (n=3), в хомозиготно състояние - CC-генотип в 16,7% (n=3) и в хомозиготно състояние – генотип TT в 66,6% (n=12). Предполага се, че rs16941 влияе на експресията на BRCA1 протеина. В европейската популация този SNP вариант се асоциира с риска за развитие на РМЖ - CC генотип определя слабо повишен риск за РМЖ в сравнение с генотип TT (OR, 1.17, 95% CI, 1.07–1.35, P = 0.033). Няма данни за rs16941 и неговата асоциация с рак на белия дроб.

В нашата извадка SNP вариантът rs1799950 в BRCA1 гена се открива в хетерозиготно състояние - генотип TC при 27,8% (n=5) и в хомозиготно състояние – генотип TT в 72,2% (n=13). Няма установени тумори с генотип CC. Предполага се, че rs1799950 също има влияние върху риска за развитие на РМЖ при европейска популация – CC-генотип се асоциира с леко повишен риск спрямо генотип TT. Няма данни за rs1799950 и неговата асоциация с рак на белия дроб.

В нашата извадка SNP вариантът rs80356964 в *BRCA1* гена се открива в хетерозиготно състояние - генотип GA при 22,2% (n=4) и в хомозиготно състояние – генотип GG в 77,8% (n=14). Няма тумори с други генотипни комбинации. Вариант rs80356964 е с неясно значение. In silico анализът позкава, че този SNP, намиращ се в богатия на серин домен води до промяна на неконсервативен нуклеотид, който в 4/5 програми за in silico анализ се определя като патогенен. Тези предварителни данни трябва да бъдат потвърдени с функционален анализ преди да може да се направи окончателна оценка на ролята му.

При нашия софтуерен анализ двата предкитора дават ралична оценка за патогенността на един от вариантите. Според стойността на PROVEAN коефициента, трите варианта се класифицират като патогенни, докато според SIFT анализа един от вариантите (rs1799950) се определя като „толерантен," а другите два варианта rs16941 и rs80356964 се класифицират като патогенни. Данните от нашата извадка туморни проби са основа за разширяване на изследванията върху ролята и значението на тези варианти в контекста на рака на белия дроб. *BRCA1* е един от тумор-супресорните гени, в който често се установяват соматични мутации при различни неоплазми. Имайки предвид водещата му роля за поддържане на хромозомна стабилност и регулацията на активността на гени, участващи в репарацията на ДНК, клетъчния цикъл и апоптозата не би било изненадващо установяването на ролята на вариантите в *BRCA1* за развитието на туморния процес при редица злокачествени заболявания, включително и рака на белия дроб.

MLH1

MLH1 е тумор-супресорен ген, кодиращ протеин, участващ в отстраняването на погрешно сдвоени бази в ДНК (mismatch repair system) и е един от гените, в които патогенни мутации водят до развитие на наследствен неполипозен колректален карцином като част от спектъра на синдромите на Линч. Носителството на патогенна мутация в MLH1 определя риск от 34% за развитие на КРК до 70 год. възраст при мъжете и 47% при жените. Рискът за развитие на ендометриален карцином при патогенна мутация в MLH1 е 18%.

При нашето изследване се установи един вариант в *MLH1*. Вариантът rs63751049 се среща при 5,6% (n=1) в хетерозиготно състояние - CT генотип и в хомозиготно състояние - TT-генотип в 94,4% (n=17). Хомозиготи по алел C не се наблюдават в нашата извадка. Според двата използвани предиктора (PROVEAN и SIFT) този вариант се определя като патогенен. Литературните данни относно значението му са противоречиви, като класификацията варира от „доброкачествен“ до „вариант с неясно значение“. C-алелът на варианта rs63751049 в MLH1 (c.977T> C (p.Val326Ala)) води до аминокиселинна замяна в еволюционно консервативен регион. 75% от in silico софтуерните предиктори класифицират мутацията като патогенна. Този вариант е докладван при засегнати членове от четири семейства с HNPCC, но за тези семейства не е налице подробна генотипна и фенотипна информация, за да се установи дали вариантът сегрегира с болестта. При популационни проучвания този вариант се открива в 65 от 122538 контролни хромозоми, преобладаващо в южноазиатската субпопулация - честота 0.0033113 (3/906). Тази честота е около 2 пъти над очакваната максимална алелна честота за патогенен *MLH1* вариант (0.0007105), което предполага вероятно доброкачествения характер на този полиморфизъм за южноазиатската популация. В противовес на данните на Spaepen и Hardt, Pastrello et al. докладват семейство, при което вариантът се среща както при засегнати, така и при здрави членове на семейството. В допълнение, туморните тъкани от пациенти с този вариант не показват загуба на хетерозигозност, загуба на експресия на *MLH1* или *MSH2* или микросателитна нестабилност. При двухибриден анализ на дрожди е устновено, че *hMLH1* V326A не влияе на асоциацията на протеина с *hPMS2* и *hEXO1*. Проучванията на Borras et al. показат, че мутацията не засяга и сплайсинга. Всички тези проучвания свидетелстват за липсата на съществен фунционален ефект на този вариант. На базата на публикуваните данни настоящата оценка клони към „доброкачествен / вероятно доброкачествен“. Няма литературни данни дали вариантът спада към групата на вариантите с малък ефект, които могат да определят предразположение към злокачествени заболявания. Необходимо е провеждането на по-мащабни и задълбочени проучвания върху туморогенния му потенциал, както при рак на белия дроб, така и при други неоплазми, за да се даде окончателна оценка на значението му за развитието на неопластичните процеси.

PMS2

PMS2 е тумор-супресорен ген, кодираш протеин, участващ в отстраняването на погрешно сдвоени бази в ДНК (mismatch repair system) и е един от гените, в които патогенни мутации водят до развитие на наследствен неполипозен колректален карцином като част от спектъра на синдромите на Линч. PMS2 е основният партньор на MLH1. Носителството на патогенна мутация в PMS2 определя риск от 19% за развитие на КРК до 70 год. възраст при мъжете и 11% при жените. Рискът за развитие на ендометриален карцином при патогенна мутация в PMS2 е 12%.

При нашето изследване се установи един вариант в *PMS2*. В нашата кохорта вариантът rs146176004 се среща при 11,1% (n=2) от туморните проби в хетерозиготно състояние – CG генотип, в 88,9% (n=16) се открива хомозиготен CC-генотип и не се откриват тумори с генотип GG. И двата софтуера за предикция - PROVEAN и SIFT, определят този вариант като патогенен, влияещ негативно върху функцията на протеина.

Публикуваните данни го определят като „вариант с неясно значение“. Вариантът rs146176004 засяга еволюционно консервативен нуклеотид, което води до промяна на аминокиселината от Gly към Ala. In silico анализът на 4 от 4 софтуерни алгоритми прогнозира патогенния му характер. Тези in silico данни не са потвърждават от експериментални проучвания. Вариантът е открит в 60/118094 контролни хромозоми (1 хомозигот) - честота от 0.0005081. Тази честота може да не е репрезентативна, тъй като използваната технология за определяне на последователността не е в състояние да различи *PMS2* от многобройните му псевдогени. Вариант rs146176004 е докладван при поне един *BRCA1/2*-негативен пациент с ранен рак на млечната жлеза, един пациент с КРК, при който също е установена и делеция в екзон 10 на *PMS2*. Така настоящата оценка за този вариант остава „с неясно значение“ и ролята му общо за злокачествените процеси, както и в по-тесен смисъл за рака на белия дроб, предстои да бъде изяснена.

PALB2

PALB2 е ген кодиращ протеин първоначално идентифициран като партньор на *BRCA2* с решаващо значение за правилното му функциониране. В последствие също така е установено, че си партнира и с *BRCA1*. Герминативните биалелни мутации водят до развитие на анемия на Fanconi, докато моноалелни мутации се асоциират с повишен риск за рак на млечната жлеза и панкреаса. Носителството на патогенна мутация в *PALB2* определя риск от 33% за развитие на РМЖ до 70 год. възраст при жените без фамилна анмнеза за РМЖ и 58% за жените с фамилна анамнеза за РМЖ.

При нашето изследване се установиха два варианта в *PALB2* - rs75023630 и rs4555163.

В нашата извадка вариантът rs75023630 се среща при 5,6% (n=1) от туморните проби в хетерозиготно състояние – CA генотип и в хомозиготни състояние - CC-генотип 94,4% (n=17). Не се открива генотип AA в излседваните тумори. Този вариант се счита за „вероятно доброкачествен“ или „доброкачествен“. Тази оценка се основава на следните критерии:

мутацията води до консервативна аминокиселинна замяна, in silico анализите я определят като вероятно доброкачествена, популационната честота е по-висока от тази характерна за патогенните мутации. Няма литературни данни, доказващи еднозначно патогенния или доброкачествения характер на този вариант.

В нашата извадка вариантът rs45551636 се среща при 5,6% (n=1) в хетерозиготно състояние - CT генотип и в хомозиготно състояние - CC-генотип в 94,4% (n=17). Няма установени тумори с хомозиготен TT генотип. Този вариант се счита за „вероятно доброкачествен“ или „доброкачествен“. Тази класификация се основава на следните критерии: мутацията води до консервативна аминокиселинна замяна, in silico анализите я определят като вероятно доброкачествена, популационната честота е по-висока от характерната за патогенните мутации честота. Ограничените литературни данни показват асоциация на T-алела на *PALB2* c.2993C>T (rs45551636 CT вариант) с повишен риск за РМЖ в семейства с фамилна анамнеза за заболяването - $OR = 3.3$ $p < 0.01$. Няма литературни данни, доказващи еднозначно патогенния или доброкачествения характер на този вариант.

И за двата варианта открити в *PALB2* в нашата туморна извадка литературните данни са оскъдни. Няма достоверни прочувания за ролята им в патогенезата, както на РМЖ така и на рака на белия дроб и другите злокачествени процеси. *PALB2* е участник в един от основните пътища в клетката, чиято регулация е нарушена при неопластичната трансформация при множество злокачествени заболявания и се очаква в бъдеще ролята на вариантите с неясно значение в него, както и в другите тумор-супресорни гени да бъдат подложени на задълбочени анализи при големи кохорти пациенти, които да открият мястото на тези полиморфизми поотделно и заедно в развитието на туморния процес.

2. Обсъждане на фармакогенетични варианти в гени за туморогенеза при НДКБК за избор на химиотерапия

Съгласно препоръките за поведение при НДКБК, изборът на първа линия терапия при напреднал НДКБК се основава на патохистологични изследвания и присъствие на определени генетични маркери – активиращи мутации в *EGFR* гена и *ALK* пренареждания. При пациенти с НДКБК, които имат *EGFR* положителен мутационен статус или *ALK* пренареждане е показано лечение с *EGFR* тирозин-киназни инхибитори като гефтиниб, ерлотиниб или афатиниб или съответно - *ALK* инхибитори като кризотиниб. Прицелната терапия е приложима само при подходящи пациенти, след провеждане на генетично изследване за мутационен статус.

При липса на мутации, които създават чувствителност към тези медикаменти, пациентите се лекуват с платина базирани химиотерапевтични режими.

Като предиктивен маркер за отговора към лечение с платина-базирана химиотерапия се препоръчва изследване на *ERCC1*-експресия. Като предиктивен маркер за отговора към лечение с гемцитабин базирана химиотерапия се препоръчва изследване на *RRM1*-експресия.

Новите NGS технологии разширяват възможностите за диагностициране на генетични варианти и за определяне на SNPs с фармакогенетично значение, които дават възможност за допълнително персонализиране на терапията според мутационния профил на тумора и генетичните характеристики на пациента.

При анализа на данните от NGS-секвенирането на туморите от нашата извадка са установени пет SNP варианта в четири гена - rs1042522 в *TP53*, rs2228001 в *XPC*, rs2227983 в *EGFR*, rs1799793 и rs13181 в *ERCC2*, като четири от вариантите са толерантни или неутрални и само един - rs1799793 в *ERCC2* гена, е патогенен според PROVEAN предиктора (Табл. 16). Тези данни са в съответствие с тезата, че фармакогенетичните варианти не влияят на появата или развитието на онкологичния процес, но могат да имат значение за терапевтичния отговор, преживяемостта и риска от нежелани лекарствени реакции (НЛР).

TP53

TP53 се счита за един от класическите тумор-супресорни гени. P53 е ядрен транскрипционен фактор с про-апоптотична функция. При повече от 50% от човешките туморни заболявания се установяват мутации, които намаляват функционалната му активност. В отговор на разнообразни клетъчни стимули, като увреждане на ДНК, p53 се натрупва в клетъчното ядро, където активира транскрипцията на редица гени. Активираният p53 води до блокиране на клетъчния цикъл, за да се позволи възстановяване на ДНК-увреждането и / или да се стартира процеса на апоптоза и така да се предотврати пролиферацията на клетки с тежки ДНК-увреждания.

В нашата туморна извадка е открит един вариант в *TP53*. SNP вариантът rs1042522 се открива в хетерозиготно състояние - генотип GC при 11,1% (n=2), в хомозиготно състояние - CC-генотип в 77,8% (n=14) и в хомозиготно състояние – генотип GG в 11,1% (n=2). Вариантът rs1042522 в TP53 има слабо клинично значение и несигурен патогенен ефект. Той е един от 25 SNP варианта, за които има данни, че самостоятелно са с малък ефект по отношение на повишаването на риска за развитие на РМЖ, но в комбинация с други варианти с малък ефект могат да доведат до клинично значимо повишен риск за РМЖ. Литературните данни по отношение на фармакогенетичния му ефект показват, че пациентите с РМЖ, които са хомозиготни носители на CC генотип за rs1042522, имат средно 3 години по-висока продължителност на живота от пациенти с хомозиготен GG генотип. Тази корелация е установена в проучване на Bojesen SE et al., включващо 9200 датски граждани. При пациенти с авансирал рак на стомаха хетерозиготният CG генотип на rs1042522 е свързан с по-слаб отговор при комбинираната химиотерапия с паклитаксел и цисплатина в сравнение с хомозиготния генотип (CC) на rs1042522 варианта и допълнително този генотип се асоциира с по-бърза прогресия на злокачествения процес. Честотата на C-алела в европейска популация е 0,71, а в нашата извадка е 0,83 - наблюдава се известно повишение на честотата му. Може да се предположи, че по – високата честота на C-алела би допринасла за потенциално по-добрия ефект на комбинираната химиотерапия с паклитаксел и цисплатина спрямо други терапевтични схеми в тази кохорта.

За G алелът е известно, че в хомозиготно и хетерозиготно състояние повишава риска от токсичност при лечение с циклофосфамид и флуроурацил. G алелът в хомозиготно състояние се установява в 11.1% (n=2) от туморите при пациенти с НДКБК в настоящата извадка. Вероятността за герминативна мутация е висока и ако бъде доказана, съществува повишен риск от токсичност при лечение с циклофосфамид и флуроурацил при тези пациенти. В такива случаи химиотерапията следва да бъде съобразена с генотипа на пациента, за да се избегнат нежелани лекарствени реакции.

Таблица 16: Генетични варианти при НДКБК с информация за фармакогенетичен ефект

Ген	**Хромозома**	**SNP**	**Генотип**	**Брой на пациенти със съответния генотип**	**Фармакогенетичен ефект**	**Тип на полиморфизма**	**Ефект на мутацията в полипептидната верига**	**SIFT cutoff =0.05**	**PROVEAN cutoff = -2.5**
TP53	chr17	rs1042522	CC	14	Намален риск от токсичност при лечение с циклофосфамид и флуроурацил	Missense	72P > R	Толерантен	Неутрален
			GC	2	Повишаване на риска от токсичност при лечение с циклофосфамид и флуроурацил				
			GG	2					
XPC	chr3	rs2228001	TT	4	Намален риск от токсичност при лечение с цисплатин	Missense	939Q > K	Толерантен	Неутрален
			GT	9	Повишен риск от токсичност при лечение с цисплатин				
			GG	5					
EGFR	chr7	rs2227983	AA	2	Намален риск от обрив при лечение с EGFR инхибитори	Missense	521R > K	Толерантен	Неутрален
			AG	7					
			GG	9	Повишен риск от обрив при лечение с EGFR инхибитори				
ERCC2	chr19	rs1799793	TT	2	При терапия с цисплатин се наблюдава удължен	Missense	312D > N	Толерантен	Патогенен

			CT	4	период без прогресия на заболяването.				
			CC	12	При терапия с цисплатин се наблюдава намален период без прогресия на заболяването				
ERCC2	chr19	rs13181	TG	6	По-висока преживяемост, при терапия с платинови съединения	Missense	751K > Q	Толерантен	Неутрален
			TT	9					
			GG	3	намалена преживяемост при лечение с платинови съединения				

XPC

XPC е важен участник в ДНК-репарацията (NER). Експерименталните проучвания показват, че XPC образува заедно с RAD23B комплекс, който участва в разпознаването на ДНК увреждането и дава началото на репаративняи процес чрез NER пътя. Терапевтичните режими на основата на платина са лечение от първа линия при пациенти с НДКБК, като ефикасността му показва значителна вариация - 26% до 60%. Предполага се, че тази разлика в терапевтичния отговор, може да е свързана с индивидуалната вариабилност в способността на туморните клетки да репарират ДНК увреждания, предизвикани от платина базирани химиотерпевтици.

В нашата туморна извадка е открит един варинат в *XPC*. SNP вариантът rs2228001 се открива при 50% (n=9) от туморите в хетерозиготна форма – генотип GT, при 22,2% (n=4) се наблюдава хомозиготен TT генотип и при 27,8% (n=5) от туморите е установен хомозиготен GG-генотип. Генотипи GT и GG се асоциират с повишен риск от токсичност при лечение с цисплатина.

По литературни данни, пациенти с остеосарком и рак на пикочния мехур с GT генотип имат повишен риск от токсичност при лечение с цисплатина, включително загуба на слуха и неутропения, в сравнение с пациентите, които са носители на TT генотип. Резултатите на Sakano S et. al. подкрепят тезата, че SNPs в ДНК-репаративните генни, особено в *XPC*, могат да бъдат прогностични фактори за остра токсичност при CRT (Cathode Ray Tubes) терапия за рак на пикочния мехур и потенциално могат да бъдат използвани за допълнително стратифициране на пациентите при назначаване на лъчетерпия за избягване на токсичност. Въпреки обещаващите първоначални данни са необходими допълнителни проучвания с по-големи размери на извадките, за да могат да бъдат направени окончателни заключения за ползата от генотипирането на пациентите за тези варианти.

При хомозиготни соматични мутации в туморна тъкан не може да се изключи герминативна мутация, която повишава риска от токсичност при лечение с цисплатин. При тези случаи следва да се обсъди алтернативна химиотерапия, за да се избегнат странични лекарствени реакции с токсично действие.

Фармакогенетичната роля на соматични мутации в *XPC* гена следва да се потвърди при едновременни изследвания в кръв и туморна тъкан.

EGFR

Рецепторът за епидермалния растежния фактор (EGFR) е трансмембранен протеин с тирозинкиназа активност, който заема важно място в сигналната трансдукционна каскада, стимулираща клетката към делене. Като се има предвид, че повече от 60% от НДКБК експресират *EGFR* и при 10-20% от тях се открива мутантна форма на протеина, този рецептор се превърна във важна терапевтична мишена при лечението на НДКБК.

В нашата туморна извадка е открит един варинат в *EGFR*. Вариантът rs2227983 се установява в 38,9% (n=7) от пациентите в хетерозиготно

състояние - генотип AG, в 11,1% (n=2) в хомзиготно състояние - AA генотип и в 50% (n=9) - хомозиготен GG-генотип.

A-алелът кодира аминокиселината лизин в позиция 521 вместо аргинин. Braig et al. докладват за резистентност към анти-EGFR моноклонални антитела базирана на този полиморфизъм. При пациенти с плоскоклчетъчен карцином генотип AA и генотип AG се асоциират с намалена преживяемост без прогресия на заболяването при терапия с анти-EGFR моноклонални антитела в сравнение с генотип GG. В разработените ракови модели те демонстрират намаление на ефекта на терапевтичната молекулата, чиято основа е нарушената стабилност на протеина, дължаща се на по-ниски нива на сиализиране на въглехидратните остатъци в молекулата на *EGFR*. Така при терапия с моноклонални антитела наличието на A-алел води до инхибиране в по-ниска степен на EGFR-сигналния път в сравнение с G-алела. Доказаната връзката между високите нива на експресия на EGFR и по-добрата чувствителност към EGFR тирозин-киназни инхибитори подкрепя тезата, че терапевтичният ефект на прицелната терапия зависи от нивата на експресия на мишената и способността на лекарствената молекула да подтисне активността ѝ. Може да се заключи, че варианти, които променят силата на отговора при образуването на комплекса лекарство-мишена влияят на крайния ефект от лечението.

В проучвания върху туморна тъкан от колоректален карцином са доказани намалените нива на фосфорилирането на *EGFR* при наличие на лизин на 521-ва позиция в протеина и по-слаба активирация на c-Myc. В резултат на намалената сила на сигналната тарнсдукция този вариант се асоциира с намалена инвазивност (P = 0,01), по-рядко засягане на лимфните възли (P = 0,02) и по-рядко метастазиране (P <0,01) при пациенти във II/III стадий, при които хирургично е отстранен първичният тумор. При пациентите с метастатичен колоректален карцином този полиморфизъм е свързан с по-добър отговор към лечението с 5-флуороурацил / оксалиплатин (P = 0,02) и по-висока продължителност на живота(P <0,01).

Установена е отлична корелация на генотипа за този вариант в периферна кръв и туморна тъкан. Така при пациенти с генотип AA и генотип AG в туморните проби може да се очаква по-нисък риск от кожни НЛР, тъй като тези генотипи се асоциират с намален риск от кожна токсичност при лечение с EGFR TKIs. Проучването на S Parmar el al. показва, че rs2227983 е свързан с кожен обрив при пациенти с НДКБК. Кожна токсичност се наблюдава в 86% от пациентите, носещи хомозиготен GG генотип за rs2227983, в 61.7% при тези с хетерозиготен GA генотип и в 10% от пациентите с AA генотип (Fisher стойност P = 0,008). В заключение, кожни НЛР са по-чести сред носителите на GG хомозиготен генотип в сравнение с носители на поне един A-алел (OR = 3.24 (1.27-8.31), P = 0.014). Необходимостта от терапевтична намеса при НЛР по типа на кожна токсичност е по-честа сред носителите на GG хомозиготен генотип спрямо носителите на хетерозиготен GA или хомозиготен AA генотип (78,9% срещу 59,4%, Fisher стойност P = 0,037). S Parmar el al. допълнително изследват адитивния ефект на промоторния вариант -216G/T и rs2227983 за

чувствителността към EGFR тирозин-киназни инхибитори. Носителите на -216G,-191C,521G хаплотип имат намален риск да развият кожна токсичност (P = 0.003).

В заключение, значението на алел-А на варианта rs2227983 е различна в зависимост от прилаганата терапия. По отношение на прицелните лекарства - А-алелът се асоциира с намалена преживяемост без прогресия на заболяването при терапия с анти-EGFR моноклонални антитела поради по-слабото подтискане на мишената от лекарството, но при лечение с EGFR TKIs е намален рискът от нежелани лекарствени реакции свързани с кожна токсичност. По отношение на конвенционалните химиотераопевтични протоколи - А-алелът се асоциира с по-рядко засягане на лимфните възли, по-рядко метастазиране, по-добър отговор към лечението с 5-флуороурацил/оксалиплатин и по-висока продължителност на живота.

Честотата на А-алела в европейска популация е 0,28, а в нашата извадка е 0,31. В нашата пациентска извадка се наблюдава слабо повишение на честотата на А-алела, което не се предполага да е статистически значимо и да определя значимо по-висок риск от НЛР при нашите пацинети.

ERCC2

ERCC2 играе ключова роля при възстановяването на ДНК-нарушения индуцирани най-често от химически агенти (вкл. и химиотепаевтици) чрез NER-пътя. Предполага се, че функционални единични нуклеотидни полиморфизми (SNPs) в този ген биха могли да влияят върху клиничните резултати от химиотерапията при онкологични пациенти. В нашата туморна извадка се откриват два варината в *ERCC2*.

Един от най-често изследваните полиморфизми в *ERCC2* с потенциално значение в тази насока е rs1799793. В нашата кохорта се установява: хетерозиготен CT генотип в 22,2% (n=4), хомозиготен TT генотип при 11,1% (n=2) и хомозиготен CC генотип в 66,7% (n=12), като честотата на Т-алела е по-ниска спрямо средната европейска честота. Т-алелът води до замяна на аминокиселината аспартат с аспарагин на 312 позиция в протеина (Asp312Asn). Повечето от проучванията, които анализират ролята на полиморфизма Asp312Asn (rs1799793), показват, че 312Asn-вариантът се асоциира с по-високи нива на ДНК-увреждане под действието на химически агенти в сравнение с тези при пациенти с 312Asp-варианта. При хомозиготите по дивия тип алел – генотип CC на rs1799793 се наблюдава значително по - кратък период без прогресия на заболяването, вероятно поради по-ефективната ДНК-репарационна система, което намалява цитотоксичността на химиотерапевтиците от групата на цисплатината и съответно тяхната ефективност. Така може да се предположи, че наличието на Т-алел в туморната тъкан води до намалена ДНК-репарационна способност на туморните клетки, което повишава тяхната чувствителност към химиотерапевтици, индуциращи ДНК-повреди. В нашата извадка алелът Т се намира с по-ниска честота спрямо средната за европейската популация, което може да се асоциира с по-ниска ефективност на някои терапевтични схеми.

Sullivan et.al. разглеждат генетичните варианти в гени, които участват в механизмите за репарацията на ДНК, тъй като се предполага, че тези гени могат да имат важно значение за ефективността на платиновата химиотерапия и токсичността на лъчетерапията при НДКБК. Sullivan и колеги анализират 17 SNP варианти в осем гена (*ERCC1*, *ERCC2*, *ERCC3*, *ERCC4*, *ERCC5*, *XPA*, *XRCC1* и *XRCC2*), участващи в механизмите за репарация на ДНК и тяхната връзка с терапевтичния отговор при НДКБК и установяват, че при пациентите в стадий III токсичността на лъчетерапията се асоциира със SNP варианти rs1799793 и rs13181 в *ERCC2* гена. При пациенти с генотип CT и TT за rs1799793 се наблюдава по-често остър пневмонит. В нашата кохорта над 80% от пацеинтите са носители на поне един T-алел за rs1799793, което предполага по-добър терапевтичен отговор и по-добра прогноза при химиотерапия с горепосочените медикаменти при тези пациенти. Нашите резултати показват, че 16,7% от НДКБК (n=3) са с генотип GG. При тези пациенти ефективността на платиновата химиотерапия може да е ограничена и потенциално да доведе до намалена преживяемост.

Фармакогенетичен вариант rs13181 се установява в 33,3% (n=6) в хетерозиготно състояние – TG-генотип, в 16,7% (n=3) в хомозиготно състояние - GG-генотип и в 50% (n=9) - хомозиготен TT-генотип. Наличието на T-алел води до замяна на аминокиселината лизин с глутамин на 751-ва позиция в протеина. Предполага се, че 751Gln алелът се асоциира по-високи нива на ДНК-увреда и по-ниска репаративна способност.

Генотипи TG и TT се асоциират с по-висока преживяемост при терапия с платинови съединения спрямо генотип GG. Проучване на Park et. al. установява, че пациентите носители на 751TT генотип (rs13181) имат най-висока степен на ремисия при терапия с 5-флуоцил/оксалиплатин (P = 0,015). Boige et al. установяват, че при пациентите с колоректален карцином и GG генотип за rs13181, лекувани с двойна терапия – флуороурацил и левковорин или с тройна терапия – флуороурацил, левковорин и оксалиплатин, се наблюдават повишен риск от лекарствена токсичност, повишен риск от ранен рецидив и намалена преживяемост без прогресия в сравнение с пациентите, които са носители на TT генотип. Тази тенденция се наблюдава и при хетерозиготно състояние на полиморфизма - GT генотипът също се асоциира с повишена преживяемост без прогресия в сравнение с пациентите, които са носители на GG генотип. Giovannetti et. al. докладват същата асоциация и при рак на панкреаса. Мета-анализът на Wei SZ et. al., показва, че ефективността на платина базираната химиотерапия се определя от комбинация от полиморфизми в *ERCC2* гена, а не индивидуално от генотипа за rs13181.

При пациентите с НДКБК на платина базирана терапия, които са носители на хомозиготния GG генотип в rs13181 се наблюдава повишен риск от пневмонит, спрямо тези, които са носители на GT или TT генотип.

Така крайната оценка за очакваната преживяемост, риск от рецидив и т.н. зависят от сумарния ефект на всички полиморфизми в гените с потенциално фармакогенетично значение, но ролята на *ERCC2*-вариантите се очертава водеща предвид натрупващите се данни в литературата. Корелацията между генотипите на SNP вариантите в *ERCC2* и ефективността на платина

базираните химиотерапевтични протоколи остава тема за обсъждане при пациенти с рак на белия дроб.

3. Обсъждане на мутационния спектър на *EGFR* гена за избор на таргетна терапия при НДКБК

В нашето проучване беше определен мутациионният статус на EGFR гена в 551 туморни проби от НДКБК чрез експресионен PCR анализ в реално време „therascreen EGFR assay“ тест, информативен за 29 мутации. В 10% от изследваните туморни проби бяха установени генни мутации.

Многобройни проучвания показват, че мутациите в EGFR тирозин-киназния домен са най-добрият предиктор за отговора към EGFR тирозин-киназни инхибитори и преживяемостта без прогресия на заболяването при НДКБК. Предишни изследвания доказват високата ефективност на EGFR тирозин-киназните инхибитори при *EGFR* положителни тумори в късните стадии (IIIB и IV). Проучването на Ziegler et al. при пациенти с НДКБК с мутация в *EGFR* гена демонстрира, че употребата на TKIs води до два пъти по – висока преживяемост без прогресия на заболяването спрямо карбоплатин-паклитаксел базираната химиотерапия. При сравнението на ефективността на гефитиниб и карбоплатин-паклитаксел като първа линия лечение за напреднал недребноклетъчен карцином на белия дроб в проучването на Wang et al. също се демонстрира цялостното превъзходство на TKI - честотата на преживяемост без прогресия при пациенти с НДКБК, които са лекувани с гефитиниб е 24,9%, спрямо 6,7% при пациенти лекувани със стандартни химиотерапевтични схеми. Наличието на мутации в *EGFR* гена е най-важният критерий за назначаване на лечение с TKIs при НДКБК.

Като друг предиктивен маркер за ефекта от лечението е изследван и броя копия на *EGFR*, но литературните данни относно ролята на броя *EGFR* копия за ефективността на TKIs са противоречиви. Няколко проучвания издигат хипотезата, че борят *EGFR* копия може да бъде по-добър предиктивен биомаркер за ефикасността на EGFR тирозин-киназните инхибитори. В противовес на тази теза са данните от проучването на Ziegler и колеги, което демонстрира, че по - високата преживяемост без прогресия при пациентите с НДКБК е асоциирана с мутации в *EGFR* гена, които определят чувствителност към тирозин-киназните инхибитори, но не и с високия брой *EGFR* копия. В заключение, настоящите клинични ръководства не препоръчват тест за оценка на броя *EGFR* копия, когато е проведен анализ за *EGFR* мутации.

В клиничната практика се наложи като стандарт определянето на *EGFR* мутационния статус чрез RT-PCR метода, поради неговата бързина, чувствителност, специфичност, възпроизводимост на резултатите и доброто съотношение на цена и информативност на изследването. Друго предимство на RT-PCR базираните протоколи за определяне на *EGFR* мутационен статус е, че те са изцяло фокусирани върху мутации с ясно клинично значение и с утвърден терапевтичен подход при наличието им.

Честотата на *EGFR* мутациите варира сред различните етнически групи – при пациенти с азиатски произход - 38,8% , при пациенти с

европейски произход - 17,4%, при пациенти афро-американци - 17,2% и в смесени популации - 27%.

Нашите резултати показват честота от 10% на *EGFR* мутациите при НДКБК при български пациенти, което е значително по-ниска честота от средната за пациентите от европейска популация. Факторите определящи тези разлики, както между българската и европейската популация, така и между европейската и азиатската популация са неизвестни. Предполага се, че различия във факторите на околната среда и / или генетични варианти, специфични за популациите, могат да са в основата тези разлики в честота на *EGFR* мутациите в световен мащаб.

Литературните данни показват ясна тенденция в характеристиката на *EGFR*-позитивните НДКБК - *EGFR* мутациите са по-чести при аденокарциноми, при жени и при непушачи. Нашите резултати са в съответствие с данните от предишни проучвания. Мнозинството от нашите пациенти с *EGFR* мутации, лекувани с TKI, са жени от европеидната раса и непушачи. В нашата извадка *EGFR* положителен мутационен статус е установен при 71% (n=39) от туморите при жени и 29% (n=16) при мъже, което определя 2,4 пъти по-висока честота на *EGFR* мутациите при пациенти от женски пол. Тези данни свидетелстват за ясния превес на женския пол (P <0,001) докладван и в литературата.

В нашата извадка *EGFR* мутациите са статистически значимо по-чести при аденокарциноми 75% (n=41) спрямо другите хистологични типове (плоскоклетъчен карцином и едроклетъчен карцином) - 25% (n=14), (P < 0.001). Тези данни са в съзвучие с докладваните в литературата честоти на *EGFR* мутации при различните хистологични варианти на НДКБК. Това наблюдение е залегнало в актуализираните препоръки на ESMO (European Society of Medical Oncology), където не се препоръчва определяне на *EGFR*-мутационен статус при плоскоклетъчен карцином, с изключение на случаите при непушачи.

В нашата извадка *EGFR* мутациите са статистически значимо по-чести при непушачи 17,7% (n=32) спрямо пушачи 7,3% (n=23) (P=0.002807). В нашата кохорта *EGFR* мутации се срещат 2,4 пъти по-често при непушачи, което потвърждава литературните данни.

Резултатите от анализа на откритите мутации в нашата извадка показва, че с най-висока честота в нашата кохорта са делециите в екзон 19 - 56.40% (n=31), следвани от мутацията L858R в екзон 21 - 36.40% (n=20). Общо те представляват около 92.7% от всички *EGFR* активиращи мутации. В нашата кохорта тези два вида мутации са установени само в аденокарциноми. Всички делеции в екзон 19 са в проби на пациенти в IV-ти стадий. Мутацията L858R предоминиращо се открива в IV стадий в нашата извадка – 90% (n=18). Останалите мутации са по-редки: инсерции в екзон 20 - 3,6% (n=2), двойна мутация - T790M, комбинирана с делеция в екзон 19 - 1,8% (n=1) и двойна мутация - S768I с делеция в екзон 19 - 1,8% (n=1). Не са доказани останалите мутации, включени в теста: L861Q, S768I, G719A, G719S, G719C. Така в нашата кохорта 92,7% от откритите *EGFR* мутации определят чувствителност към тирозин-киназни инхибитори, а останалите 7,3% от откритите мутации се асоциират с резистентност към първо и второ

поколение тирозин-киназни инхибитори. Разработваните в момента трето поколение тирозин-киназни инхибитори са насочени към T790M положителните пациенти. Те са структурно различни от първо и второ поколение TKIs - свързват се необратимо с *EGFR* и са ефективни, както при чувствителните на първо и второ поколение мутации (делеции в екзон 19 и L858R), така и при резистентната на тях T790M. Вече има FDA и EMA одобрени трето поколение TKIs, които се прилагат в клиничната практика. За сега няма данни в литературата за потенциалното приложение на тези медикаменти при двойните мутанти (комбинация между T790M и друга мутация).

В нашата кохорта най-честата мутация в *EGFR* е делецията в екзон 19 - 56,4% (n=31) от всички *EGFR* мутации и се среща в 5,6% от НДКБК. Честотата на делециите в екзон 19 варира в различните популации. В Южна Азия се наблюдава значително по-висока честота на тази мутация - 20,3% от всички изследвани случаи, в сравнение с всички останали региони. Също така висока честота се наблюдава и в популации от Северна Азия - 12,8%. За Европа е докладвана честота от 9.0%, за Северна Америка - 4,5% и Южна Америка - 4,2%. Така наблюдаваната в нашата кохорта честотата е по – ниска от средната за европейската популация и най-близо до тази рапортувана в северно- и южно-американските популации.

Втората по-честота мутация в нашата кохорта е L858R – 36,4% (n=20) от всички открити *EGFR* мутации и се установява при 3,6% от изследваните пациенти. Най-висока честотата на тази мутация се наблюдава в Южна Азия - 21,6% и Северна Азия -13,9% от всички изследвани пациенти. В останалите региони честотата е какво следва: Европа – 3,3%, за Северна Америка – 3,4% и Южна Америка – 2,6%. Честотата наблюдавана в нашата кохорта е съизмерима с тази рапортувана за Европа и Северна Америка.

Инсерции в екзон 20 са установени при 0,4% от изследваните пациенти и представляват 3,6% от всички установени мутации в нашата кохорта. Средната честота на тази мутация за Европа се оценява на 0,3%, което е съизмеримо с наблюдаваната в нашата извадка. Най-висока честота на тази мутация е докладвана при пациенти от Африка (вкл. и Близкия Изток) – 1,9%. За Северна и Южна Америка честотата е съответно – 0,4% и 0,3%. Най-ниската честота се открива в Северна Азия – 0,07%.

Честотата на двойните мутанти в нашата кохорта е 3,6% от *EGFR* положителните случаи и 0,4% от всички пациенти. Честотата на двойните мутанти в Азия е 2-7% [Wei, 2014, Zhang, 2007, Toshihiko, 2006]. Честотата в Европейска популация е 0,1% от всички изследвани и 0,8% от *EGFR* положителните случаи. Така делът на двойните мутанти в нашата кохорта *EGFR* положителни случаи, както и делът им в общата популация изследвани НДКБК, са значително по-високи в сравнение с този в Европа - (3,6% vs 0,8%) и (0,4% vs 0,1%).

В заключение, общата честота на мутациите в *EGFR* в нашата кохорта е по-ниска от средната за Европа. Делът на отделните мутации е съизмерим с европейския за L858R и инсерциите в екзон 20, докато делециите в екзон 19 са с по-ниска честотата спрямо европейската, а двойните мутанти се срещат по-често от средното за Европа.

4. Обсъждане на патогенни варианти в гени, асоциирани с карциногенеза при ФПТК

При анализа на данните от NGS секвенирането на 12-те кръвни проби от пациенти с ФПТК се установиха 9 патогенни варианти (според PolyPhen и SIFT предикторите) в три гена свързани с туморогенезата - *MET*, *RET* и *MLH1*.

RET

RET генът спада към групата на прото-онкогените. Той кодира трансмембранен рецептор с тирозин-киназна активност. Свързването на лиганди като GDNF (невротрофичен фактор произхождащ от глиална клетъчна линия) и други растежни фактори към рецептора води до активирането на сигнални пътища, които играят роля в клетъчната диференциация, растеж, миграция и оцеляване. Кодираният рецептор е важен за развитието на нервната система и органите и тъканите, произлизащи от невралния гребен. Герминативни мутации повишаващи активността на протеина (gain-of-function) - цитогенетично пренареждане и различни точкови мутации, водят до развитието на наследствени форми на тироиден карцином, докато соматични мутации в него като цяло се асоциират с карциногенезата и се установяват в 50% от случаите на медуларен тироиден карцином. Мутациите, които водят до загуба на функцията (loss-of-function) се асоциират с болестта на Hirschsprung (аганглионарен мегаколон), синдрома на централна хиповентилация, като също така са идентифицирани и при пациенти с бъбречна агенезия.

В нашата извадка SNP вариантът rs181856591 в *RET* гена се открива в хетерозиготно състояние - генотип CA при 25% (n=3) и в хомозиготно състояние - CC-генотип в 75% (n=9). Няма установени пациенти с хомозиготен AA генотип. Според предикцията по SIFT, този вариант се класифицира като патогенен или вреден (deleterious) и според PolyPhen се определя като вероятно патогенен (probably damaging).

A-алелът води до замяна на аминокиселината треонин с лизин в кодон 754 на RET протеина (p.Thr754Lys). Тази замяна не е описана до сега в литературата – не се установява и в популационните бази данни (ExAC, 1000Genomes, TOPMED). В този кодон е описана транзиция C>T, която води до замяна на треонин с метионин. Популационната честота (ExAC) на T-алела за този вариант се оценява на 0,01%, което е съизмеримо с честотата на редица патогенни варианти в *BRCA1* и *BRCA2*. Няма убедителни данни, че T-алелът е патогенен и на базата на наличната информация се класифицирана като вариант на неясно значение. Треониновият остатък се намира в силно консервативен участък и умерената разлика във физикохимичните свойства на треонин и лизин би могла да доведе до промяна в активността на *RET*, но ролята на A-алела (Lys), който се установява за първи път при нашите пациенти, предстои да бъде установена. SNP вариантът rs181856591 - CA генотип, се открива при засегнати инидивиди от две семейства. В фамилия Б и всички засегнати индивиди (n=2) са хетерозиготно носители на CA-генотип. Във фамилия В 20% от засегнатите (n=1) е с хетерозиготен генотип. Така този вариант, при

потвърждаване на патогенния му характер, най-вероятно би бил с умерен ефект по-отношение на риска за развитие на злокачествени заболявания на щитовидната жлеза и би се класифицирал като фактор определящи фамилно предразположение. Фактът,че вариантът до сега не е описан в нито една популация в глобален мащаб, свидетелства за възможния локален характер на този алел.

MET

MET генът спада към групата на прото-онкогените. Той кодира трансмембранен рецептор с тирозин-киназна активност. Свързването на неговия лиганд, хепатоцитен растежен фактор води до активиране на рецептора и стартиране на сигнална трансдукционна каскада, която има роля за клетъчното оцеляване, ембриогенезата, клетъчната миграция и туморната инвазия. Мутациите в този ген са свързани с развитието на папиларен бъбречноклетъчен карцином, хепатоцелуларен карцином и различни злокачествени заболявания на главата и шията. Дерегулация - амплификация и свръхекспресия на този ген, се установяват в множество човешки неоплазми. Свръхактивирането на HGF / MET сигналния път потиска апоптозата и стимулира пролиферацията, инвазията, неоангиогенеза и метастазирането на злокачествения процес.

При нашето изследване се установиха пет варианта в *MET* гена - rs56391007, NP_001120972.1: p.Leu1080Met, NP_001120972.1: p.Phe1234Ser, NP_001120972.1: p.Pro712Gln и NP_001120972.1:p.Ser823Tyr. Само един от тези варианти - rs56391007 е описан в литаратурата.

В нашата извадка вариантът rs56391007 се среща при 25% (n=3) от кръвните проби в хетерозиготно състояние – CT генотип и в хомозиготни състояние - CC-генотип 75% (n=9). Не се открива генотип TT в изследваните тумори. Според предикцията по SIFT, този вариант се класифицира като патогенен (deleterious) и според PolyPhen се определя като вероятно патогенен (probably damaging).

SNP вариантът rs56391007 - CT генотип, се открива при две фамилии – Б и В, като във фамиля Б вариантът се установява в 50% (n=1) от засегнатите, а във фамиля В 40% (n=2) от пациентите с тироиден карцином.

T-алелът на rs56391007 в *MET* гена води до замяна на аминокиселината треонин с изолевцин на 1010 позиция в протеина (Thr1010Ile), която се намира в консервативен нуклеотид и 3/4 in silico софтуерни програми определят патогенния характер на този вариант. Вариантът се намира в c-MET JM домен, който регулира цитоскелетните функции: адхезията, подвижността и миграцията. Има литературни данни, че мутация в този регион в in vitro условия стимулира туморния растеж при рак на белия дроб. Според ExAC популационната база данни общата честота на варианта се оценява на 8/1000, като при 6 от индивидите T-алелът е в хомозиготно състояние. Особено висока честота се наблюдава в африканската субпопулация – 2%. Тези честоти надвишават очакваната максимална алелна честота за патогенен алел в този ген (15/10 000 000). Вариантът е рапортуван, както при здрави индивиди, така и при пациенти с метастатичен рак на гърдата, рак на белия дроб и рак на дебелото черво. Той не се асоциира с повишен риск от развитие на рак на млечната жлеза

при големи асоциативни проучвания и литературните данни за функционалния му ефект са противоречиви. На базата на оценката на множество реномирани бази данни / клинични лаборатории и литературните данни, този вариант се класифицира като „доброкачествен“. Оценката на двете предикторни програми използвани при анализа на нашите NGS-данни е в съответствие с тази на другите предикторни софтуери и класифицира варианта като засягащ функцията на протеина. Тази начална оценка не е окончателно потвърдена при функционалните анализи и са необходими допълнително задълбочени проучвания за еднозначната класификация на този вариант.

В нашата извадка вариантът NP_001120972.1:p.Leu1080Met се среща при 8.3% (n=1) от кръвните проби в хетерозиготно състояние – CA генотип и в хомозиготни състояние - CC-генотип 91.7% (n=11). Не се открива генотип AA в изследваните пациенти. Замяната в аминокиселинната последователност е малко вероятно да доведе до тежки промени във функцията на протеина поради същите физикохимични свойства на двете аминокиселини (неутрални аминокиселини с алифатен радикал).

В нашата извадка вариантът NP_001120972.1:p.Phe1234Ser се среща при 8.3% (n=1) от кръвните проби в хетерозиготно състояние – TC генотип и в хомозиготни състояние - TT-генотип 91.7% (n=11). Не се открива генотип CC в изследваните кръвни проби. Трудно е да се оцени ефекта на замяната в аминокиселинната, тъй като и двете аминокиселина са неутрални, но фенилаланинът е с алифатен радикал за разлика от серина, който има полярен радикал.

В нашата извадка вариантът NP_001120972.1:p.Pro712Gln се среща при 8.3% (n=1) от кръвните проби в хетерозиготно състояние – CA генотип и в хомозиготни състояние - CC-генотип 91.7% (n=11). Не се открива генотип AA в изследваните пациенти. Трудно е да се оцени ефекта на замяната в аминокиселинната, тъй като и двете аминокиселина са неутрални, но пролинът е с алифатен радикал за разлика от глутамина, който има полярен радикал.

В нашата извадка вариантът NP_001120972.1:p.Ser823Tyr се среща при 8.3% (n=1) от кръвните проби в хетерозиготно състояние – CA генотип и в хомозиготни състояние - CC-генотип 91.7% (n=11). Не се открива генотип AA в изследваните проби. Замяната в аминокиселинната последователност е малко вероятно да доведе до тежки промени във функцията на протеина поради същите физикохимични свойства на двете аминокиселини (неутрални аминокиселини с полярен радикал).

Според предикцията по SIFT, тези варианти се класифицират като патогенни (deleterious) и според PolyPhen се определят като вероятно патогенни (probably damaging). Няма експериментални данни за ефекта на вариантите in vitro и in vivo. Тези варианти не са реапортувани до сега в международните бази данни. Тяхното отношение към канцерогенезата не може да бъде изяснено на базата на нашите изследвания. Възможно е да се касае за редки доброкачествени варианти, характерни за изследваната популация или за варианти определящи предразположение към щитовиден карцином с малък до умерен ефект.

MLH1

MLH1 е тумор-супресорен ген, кодиращ протеин, участващ в отстраняването на погрешно сдвоени бази в ДНК (mismatch repair system). Този ген е асоцииран с колоректален и ендометриален карцином.

При нашето изследване се установиха два варианта в *MLH1* - NP_000240.1:p.Glu23Gly и NP_000240.1:p.Gly634Val, които не са анотирани в международните бази данни до сега.

В нашата извадка вариантът NP_000240.1:p.Glu23Gly се среща при 33.3% (n=4) от кръвните проби в хетерозиготно състояние – AG генотип и в хомозиготни състояние - AA-генотип 66.7% (n=8). Не се открива генотип GG в изследваните пациенти. Може да се предположи, че тази промяна в аминокиселинната последователност би имала ефект върху функцията на протеина поради различните физикохимични свойства на референтната аминокиселина (Glu), която е киселинна - с отрицателен електрически заряд и аминокиселината, която се кодира от G-алела (Gly), която е неутрална с алифатен радикал.

В нашата извадка вариантът NP_000240.1:p.Gly634Val се среща при 8.3% (n=1) от изследваните проби в хетерозиготно състояние – GT генотип и в хомозиготни състояние - GG-генотип 91.7% (n=11). Не се открива генотип TT в нашата извадка. Замяната в аминокиселинната последователност е малко вероятно да доведе до тежки промени във функцията на протеина поради същите физикохимични свойства на двете аминокиселини (неутрални аминокиселини с алифатен радикал).

Според предикцията по SIFT и двата варианта се класифицират като патогенни (deleterious) и според PolyPhen се определят като вероятно патогенни (probably damaging). Няма експериментални данни за ефекта на вариантите in vitro и in vivo. Тези варианти не са реапортувани до сега в международните бази данни. Тяхното отношение към канцерогенезата не може да бъде изяснено на базата на нашите изследвания. Възможно е да се касае за редки доброкачествени варианти, характерни за изследваната популация или за варианти определящи предразположение към щитовиден карцином с малък до умерен ефект.

5. Обсъждане на фармакогенетични варианти в гени за туморогенеза при ФПТК за избор на химиотерапия

При анализа на данните от NGS-секвенирането на кръвни проби от нашата извадка са установени шест SNP варианта в шест гена - rs1042522 в *TP53*, rs13181 в *ERCC2*, rs1799939 в *RET*, rs2227983 в *EGFR*, rs2228001 в *XPC* и rs17655 в *ERCC5*, които са толерантни или бенигнени, според SIFT и PolyPhen предикторите (Таблица 17). Тези данни са в съответствие с тезата, че фармакогенетичните варианти не влияят на появата или развитието на онкологичния процес, но могат да имат значение за терапевтичния отговор, преживяемостта и риска от нежелани лекарствени реакции (НЛР).

TP53

В нашата извадка е открит един варинат в *TP53*. SNP вариантът rs1042522 се открива в хетерозиготно състояние - генотип GC при 25%

(n=3), в хомозиготно състояние - CC-генотип в 58.3% (n=7) и в хомозиготно състояние – генотип GG в 16.7% (n=2).

GC и GG генотипите определят повишен риск от токсичност при лечение със стандартни химиотерапевтични режими включващи флуороурацил и циклофосфамид. Проучването на Borchiellini et al. показва, асоциацията на C-алела (Pro) с повишен риск от остеонекроза при пациенти с рак на главата и шията, които са подложени на лъчетерапия.

Интересни са данните за този полиморфизъм и връзката му с рака на щитовидната жлеза. При проучването на Yan et. al. е установено, че разпределението на генотипните честоти при пациенти с рак на щитовидната жлеза за rs1042522 е както следва: Arg/Arg (GG) генотип - 0.286, Arg/Pro (GC) генотип - 0.343 и Pro/Pro (CC) генотип - 0.371. Наблюдаваните генотипни честоти при контролите са Arg/Arg (GG) - 0.45, на Arg/Pro (GC) генотип - 0.37 и Pro/Pro (CC) генотип - 0.18. Честотата на C-алела (Pro) е значително по - висока от честотата на G-алела (Arg) в групата на пациентите (OR = 2,06, 95% C.I = 1,5-2,8). Това изследване показва, че C-алелът на SNP варианта rs1042522 определя по-висок риск за рак на щитовидната жлеза, особено при млади индивиди, жени и непушачи.

Така при 41,7% (n=5) от изследваните пациенти може да се очаква по – висока честота на НЛР, което да наложи понижаване на лекарствената доза или промяна на терапевтичния режим. При 58,3% (n=7), които са хомозиготи по C-алел при прилагане на лъчетерапия може да се очаква повишен риск от остеонекроза. По – слабо изразено повишаване на рисковете свързани с лъчетерапията се очаква при 25% (n=3) от нашите пациенти (хетерозиготи). При пациентите, които са с по-нисък риск за НЛР при химиотерапия (CC-генотип), същевременно се наблюдава по-висок риск за развитие на щитовиден карцином и наличието на този генотип е най-вероятно един от допринасящите фактори за развитие на основното заболяване.

Table 17: Резултати от NGS. Таблицата представя ген; хромозомна локализация; прогнозни параметри (SIFT и PolyPhen); SNP; генотипи; брой пациенти с този генотип; чувствителност към лекарства, свързана с rs (според уебсайта на pharmagkb); вид мутация;

Ген	**Хромозома**	**SIFT/PolyPhen**	**SNP**	**Генотип**	**Брой на пациенти с определения генотип**	**Лекарства**	**Тип на полиморфизъм**
TP53	17	Tolerated/Benign	rs1042522	Хетерозигот GC	3	Повишаване на риска от токсичност към лечение с циклофосфамид и флуроурацил	Missense
				Хомозигот GG	2		
				Хомозигот CC	7	Намален риск от токсичност към лечение с циклофосфамид и флуроурацил	
ERCC2	19	Tolerated/Benign	rs13181	Хетерозигот TG	7	По-висока преживяемост, когато се лекува с платинови съединения	Missense
				Хомозигот TT	1		
				Хомозигот GG	4	Намалява оцеляването при лечение с платинови съединения	
RET	10	Tolerated/Benign	rs1799939	Хетерозигот GA	7	Не са свързани с токсичност към лекарства със сунитиниб	Missense
				Хомозигот AA	0		
				Хомозигот GG	5	Предполагаема токсичност на сунитиниб	
EGFR	7	Tolerated/Benign	rs2227983	Хетерозигот GA	6	Намален риск от обрив при лечение с EGFR	Missense

				Хомозигот AA	1	инхибитори	
				Хомозигот GG	5	Повишен риск от обрив при лечение с EGFR инхибитори	
XPC	3	Tolerated/Benign	rs2228001	Хетерозигот GT	6	Повишен риск от токсичност при лечение с цисплатин	Missense
				Хомозигот GG	3		
				Хомозигот TT	3	Намален риск от токсичност при лечение с цисплатин	
ERCC5	13	Likely benign	rs17655	Хетерозигот GC	4	Повишава оцеляването без прогресия при химиотерапия с платина	Missense
				Хомозигот CC	0		
				Хомозигот GG	8	Намалява оцеляването без прогресия при химиотерапия с платина	

ERCC2

В нашата извадка е открит един вариант в *ERCC2* гена. Фармакогенетичен вариант rs13181 се установява в 58,3% (n=7) в хетерозиготно състояние – TG-генотип, в 33.4% (n=4) в хомозиготно състояние - GG-генотип и в 8.3% (n=1) – в хомозиготен TT-генотип.

Полиморфизмът rs13181 в *ERCC2* гена е известен като фактор, влияещ върху риска за НЛР при химиотерапия. При пациенти с рак на гърдата, носители на GG генотип и GT генотип се установява намален риск от неутропения при лечение с доцетаксел в сравнение с пациенти, които са носители на TT генотип. Според изследванията на Sullivan et. al. върху НДКБК SNP вариантът rs13181 не определя съществени разлики в токсичността на стандартните химиотерапевтични схеми, но се открива значима асоциации между SNPs (rs13181) в *ERCC2* гена и НЛР от типа на остър пневмонит при радиотерапия. Техните данни показват, че 32.3% от пациентите, които са носители на T алел (TT и TG генотип), развиват пневмонит, като при носителите на GG генотип това усложнение се наблюдава в 83.3%. Ролята на варианта rs13181 конкретно при карцином на щитовидната жлеза не е изследвана и предстои да бъде уточнена. При нашите пациенти повишен риск за неутропения при химиотерапия на базата на генотипа на rs13181 може да се очаква при 8,3% (n=1) и повишен риск за пневмонит при лъчетерапия може да се очаква при 33,4% (n=4).

RET

В нашата извадка е открит един вариант в *RET* гена. Фармакогенетичен вариант rs1799939 се установява в 58,3% (n=7) в хетерозиготно състояние – AG-генотип, в 41.7% (n=5) в хомозиготно състояние - GG-генотип. Няма установени пациенти с хомозиготен AA генотип.

A-алелът води до замяна на аминокиселината глицин със серин на 691-ва позиция в протеина. Предполага се, че тази замяна води до появата на две нови места за фосфорилиране на *RET* и може да има слаб ефект върху функцията на протеина. In vitro проучвания показват, че този вариант е свързан с амплифициране на отговора на сигналната каскада и стимулира клетките към репликация / инвазия. Проучването на Quintela-Fandino et al. показва, че наличието на 691Ser 10-кратно повишава нивата на фосфорилиране на *RET* и води до неговата свръхактивация. Авторите предполагат, че тази мутация е подходяща за лечение с тирозин-киназния инхибитор довитиниб, като двамата пациенти с тази соматична мутация подложени на лечението постигат контрол на заболяването и при оттегляне на медикамента неопластичният процес прогресира. Мета анализите показват, че този вариант е с ниска пенетрантност и има модифициращa роля по отношение на риска за развитие на рак.

Проучването на Kim et al. разглежда в дълбочина риска от хематологична токсичност - тромбоцитопения, неутропения и анемия при лечението със сунитиниб. Общо 12 генетични полиморфизми в 8 кандидат-гени (*CYP1A1*, *CYP3A5*, *ABCB1*, *ABCG2*, *PDGFRα*, *VEGFR2*, *RET* и *FLT3*) са анализирани за връзка с терапевтичната токсичност на сунитиниб. При

пациенти с бъбречноклетъчен карцином, SNP (rs1799939) вариант не се асоциира с нежелани странични реакции към сунитиниб.

Няма литературни данни за ефекта на този вариант при тироиден карцином. В нашата извадка 58,3% (n=7) хетерозиготни носители на А-алел. При тези пациенти би било подходящо изследване на туморната проба за RET-активност и включването им в клинични проучвания за довитиниб.

EGFR

В нашата извадка е открит един вариант в *EGFR*. Вариантът rs2227983 се установява в 50% (n=6) от пациентите в хетерозиготно състояние - генотип GA, в 8.3% (n=1) в хомозиготно състояние - AA генотип и в 41.7% (n=5) - хомозиготен GG-генотип. Генотипите GG се асоциират с повишен риск за реакции на кожна токсичност при лечение с TKIs.

Литературните данни за значението на този полиморфизъм са противоречиви. Мета-анализът на Wang et al. на 13 проучвания върху различни видове рак (рак на млечната жлеза, колоректален карцином, белодробен карцином, рак на щитовидната жлеза и др.) показва, че няма значима връзка между полиморфизма rs2227983 в *EGFR* гена и риска за развитие на рак. Тяхното проучване също така не показва еднозначна асоциация между rs2227983 полиморфизма и ефекта от химиотерапия и свидетелства за необходимостта от допълнителни мащабни изследвания, които да потвърдят или отхвърлят такава асоциация.

Противно на данните на Wang, резултатите от проучването на Parmar et al. при пациенти с НДКБК, колоректален карцином и рак на панкреаса показват, че AA и AG генотипите се асоциират с намален риск от кожна токсичност при лечение с тирозин-киназни инхибитори в сравнение с GG генотип (P = 0.008).

В литературата има ограничени данни за ефикасността на лечението с EGFR-TKIs при рак на щитовидната жлеза. Така в нашата извадка при 41,7% (n=5) от пациентите при лечение с тирозин-киназни инхибитори може да се очаква повишен риск от реакции на кожна токсичност. В тези случаи може да се предложи по-активно поведение по отношение на профилактиката на този тип НЛР.

XPC

В нашата извадка е открит един варинат в *XPC*. SNP вариантът rs2228001 се открива при 50% (n=6) от пробите в хетерозиготна форма – генотип GT, при 25% (n=3) се наблюдава хомозиготен TT генотип и при 25% (n=3) от пробите е установен хомозиготен GG-генотип. SNP вариантът се асоциира с риска от токсичност при някои химиотерапевтици. Така при пациенти с генотип GT и GG е отчетен по-висок риск от хематологична токсичност при лечение с цисплатина. Алелът G се асоциира и с по-висок риск от ототоксичност, по – изразен в хомозиготно състояние.

Изследвана е и връзката на този полиморфизъм с ефективността от химиотерапията. Според данните на Caronia et al. наличието на поне един G-алел се асоциира с по - добър терапевтичен отговор в сравнение с хомозиготите за T-алел (OR = 0,34, P = 0,032).

Така вариантите в гените за ДНК репарацията се очертават като добър прогностичен маркер за токсичността на определени терапевтични

режими. При 75% от нашите пациенти на базата на генотипа rs2228001 може да се очаква по-висок риск за потискане на хематопоезата и слухово увреждане. За тези пациенти в бъдеще може да се предвиди, както по-често проследяване на базовите хематологични параметри за навременно третиране на усложненията, така и аудиологично изследване за установяване на отклонения в слуха.

ERCC5

Този ген кодира ДНК ендонуклеаза, която участва в ДНК репарационните механизми при UV-индуцирано ДНК-увреждане. *ERCC5* участва и в други клетъчни процеси – транскрипция на РНК полимераза II. Патогенните мутации в този ген водят до развитието на кожното заболяване xeroderma pigmentosum, характеризиращо се с свръхчувствителност към ултравиолетова светлина и висок риск за развитие на рак на кожата.

В нашата извадка е открит един варинат в *ERCC5*. Вариантът rs17655 се установява в 33.3% (n=4) от пациентите в хетерозиготно състояние - генотип GC, в 66.7% (n=8) в хомзиготно състояние - GG генотип. Няма установени пациенти с хомозиготен CC генотип.

Този SNP се намира в С-края на протеина *ERCC5*, където се намират домените за свързване на протеина с другите участници в NER репарационния комплекс - XPB, XPD, p62 и p44 субединиците на транскрипционния фактор IIH. С-алелът води до промяна на аминокиселината Asp на His на 1104-та позиция в протеина и се предполага, че пречи на изграждането на NER-репарационния комплекс и неговата ефикасност.

В литературата SNP вариантът rs17655 в *ERCC5* гена се асоциира с общата преживяемост и преживяемостта без прогресия при пациенти с рак на яйчника подложени на химиотерапия. Наличието на G-съдържащ генотип е свързано с по-добра прогноза в ранен стадий и с по-лоша прогноза при късната фаза на заболяването. Този феномен може да се дължи на различната физиопатологията на рака на яйчника в ранен и късен стадий. Данните на Caiola et. al. показват, че генотипирането на *ERCC5*, може да помогне за прогнозиране на отговора към платина базирана химиотерапия – при пациентите с CC генотип и в късен стадий се очаква по-лош терапевтичен отговор.

Ролята на този полиморфизъм при шитовиден карцином не е изследвана до сега и предстои да бъде изяснена. В нашата извадка не са открити пациенти хомозиготи с CC генотип, което предполага, при наличие на същата асоциация както при рака на овариума, добър ефект на химиотерапията с платина в тази кохорта.

V. Заключение

Проведеното изследване е в модерна област на медицината, тъй като е свързано с характеризиране на фармакогенетични варианти, асоциирани с химиотерапия при пациенти с НДКБК и ФПТК, използвайки NGS технологията. Изборът на подходящо лечение в зависимост от молекулния туморен фенотип ще доведе до намаляване на нежелани лекарствени реакции и подобряване на ефекта от лечението. Това от своя страна ще

подобри качеството на живот на конкретния пациент и ще намали финансовото бреме на обществото.

Получените резултати имат научно и научно-приложно значение, тъй като определят фармакогенетични варианти в онкогени и тумор-супресорни гени, които имат отношение към ефекта от химиотерапията при пациенти с НДКБК и с фамилна предразположеност към тироиден карцином.

С настоящото проучване е определен мутационния спектъра на EGFR гена при белодробен карцином чрез изследване с real-time PCR. Установен е типа и честотата на чувствителните мутации за прицелна терапия с тирозин киназни инхибитори.

Нашите резултати добавят доказателства в подкрепа на диагнозата, че всички пациенти с аденокарцином на НДКБК трябва да се подложат на генетичен тест за мутации в EGFR гена за лечение с EGFR тирозин-киназни инхибитори.

NGS – технологията има потенциал да подобри откриването на нови генетични варианти свързани с раковите заболявания, което ще доведе до по-добри решения за лечение и различни подходи за мониторинг. Разработването на различни видове NGS панели, ще позволи разкриването на редки фармакогенетични варианти и разработването на стратегии за индивидуализиране на терапията съобразно генните варианти, оптимизиране на ефикасността и минимизиране на токсичността на лекарствените средства.

VI. Изводи

1. В туморни проби на пациенти с НДКБК са установени седем патогенни варианта в четири от 94-те изследвани гена, асоциирани с туморогенеза: rs16941, rs80356964 и rs1799950 в *BRCA1*, rs146176004 в *PMS2*, rs75023630 и rs45551636 в *PALB2* и rs63751049 в *MLH1*.
2. При НДКБК са доказани пет фармакогенетични варианта в четири от 94-те изследвани гена - rs1042522 в *TP53*, rs2228001 в *XPC*, rs2227983 в *EGFR*, rs1799793 и rs13181 в *ERCC2*.
3. Потвърдено е, че активиращите мутации в *EGFR* гена при пациенти с НДКБК са по-чести при женски пол, при аденокарциноми в сравнение плоскоклетъчни или/и едроклетъчни карциноми и при непушачи.
4. Установено е, че общата честота на мутациите в *EGFR* гена в българска популация е по-ниска (10%) в сравнение със средната за Европа (17%). Делът на отделните мутации е съизмерим с европейския по отношение на точкова мутация L858R в екзон 21 (3,6%) и инсерции в екзон 20 (0,4%), докато делециите в екзон 19 са с по-ниска честота (5,6%) спрямо европейската (съответно 3,3%; 0,3%; 9,0%).
5. В кръвни проби на пациенти с ФПТК са установени осем патогенни варианта в три от 94-те изследвани гена, свързани с туморогенезата - два публикувани rs56391007 в *MET* и rs181856591 в *RET* и шест нови варианта, непубликувани в литературата - NP_001120972.1: p.Leu1080Met, NP_001120972.1: p.Phe1234Ser, NP_001120972.1:p.Pro712Gln, NP_001120972.1:p.Ser823Tyr в *MET* и NP_000240.1:p.Glu23Gly, NP_000240.1:p.Gly634Val в *MLH1*.
6. При секвениране с панел TruSight Cancer на пациенти с НДКБК или ФПТК от общо 7800 установени варианти са доказани 7 фармакогенетични варианти, от които 4 се откриват при двата типа карциноми - rs1042522 в *TP53*, rs2228001 в *XPC*, rs2227983 в *EGFR* и rs1318 в *ERCC2*.

VII. Приноси

Приноси с научен характер:

1. Онкогени/туморсупресорни гени освен патогенни варианти могат да съдържат фармакогенетични варианти, асоциирани с неефективно лечение с платинови съединения и намалена преживяемост на пациентите.
2. Доказани са 6 нови патогенни герминативни варианти при фамилен папиларен тиреоиден карцином.
3. Установено е, че българската популация има по-ниска обща честота на мутациите в EGFR гена в сравнение със средната за Европа за сметка на делециите в екзон 19.

Приноси с приложен характер:

1. Секвенирането на ДНК-и от туморни проби с TruSight Cancer Sequence панел може да определи соматични фармакогенетични варианти, свързани с неефективна платина базирана химиотерапия.
2. Секвенирането на ДНК-и от кръвни проби с TruSight Cancer Sequence панел може да определи герминативни фармакогенетични варианти, определящи токсично действие при химиотерапия.

VIII. Научни публикации и съобщения по темата на дисертационния труд

Научни публикации във връзка с дисертационния труд

1. N. Chilingirova · Z. Hammoudeh · L. Balabanski · S. Ivanov · R. Vazharova · D. Nikolova · G. Kurteva · D. Toncheva · P. Chilingirov, **TruSight Cancer Sequencing Panel reveals pharmacogenetic variants associated with sensitivity to chemotherapy in lung cancer**, **Memo – Magazine of European Medical Oncology.** February 2016, Volume 9, Issue 1, pp 30–38. (N. Chilingirova, Z. Hammoudeh, equally contributed to this work).
2. Hammoudeh, Z., Nikolova, D., Balabanski, L., Ivanov, S., Vazharova, R., Weidner, S., Malinov, M., Toncheva, D., **Screening of pharmacogenetic variants associated with drug sensitivity in patients with papillary thyroid carcinoma using next generation sequencing, Biotechnology & Biotechnological Equipment**. Pages 761-765 | Received 18 Feb 2017, Accepted 24 May 2017, Published online: 07 Jun 2017.
3. Hammoudeh Z., Antonova O., Staneva R., Nikolova D., Kyuchukov Y., Penev A., Mintchev T., Koleva V., Hadjidekova S. and Toncheva D., **Detecting EGFR Mutations in Patients with Non-Small Cell Lung Cancer, B J M G - Balkan Journal of Medical Genetics.** June 29, 2018.

Научни съобщения на конгреси във връзка с дисертационния труд

1. Hammoudeh Z. A., Koleva V., Dakova P., Oreshkov K., Minchev C., Penev A., Benova A., Radinov A., Toncheva D. "**EGFR mutation analysis in non-small cell lung cancer". European Human Genetic Conference, June 8 - 11, 2013, Palais des Congrès, Paris, France**, P11.167, 305
2. Hammoudeh Z., Chilingirova N., Nesheva D., Rukova B., Karachanak-Yankova S., Antonova O., Staneva R., Damyanova V., Hadjidekova S., Dimova I., Balabanski L., Ivanov S., Vazharova R., Kurteva G., Toncheva D. "**Pharmacogenetic variants in patients with non-small cell lung cancer". European Human Genetic Conference, June 6 - 9, 2015, Glasgow, Scotland, United Kingdom**, PM12.090, 255
3. Hammoudeh Z., Nikolova D., Balabanski L., Ivanov S., Vazharova R., Weidner S., Malinov M., Toncheva D. "**Pharmacogenetic variants in patients with thyroid cancer". European Human Genetic Conference, May 21 - 24, 2016, Barcelona, Spain**, E-P12.099
4. Hammoudeh Z., Koleva V., Kyuchukov Y., Dakova P., Penev A., Minchev T., Toncheva D.. "**Detecting EGFR Exon 19 Deletions and L858R in Patients with Non-Small Cell Lung Cancer". European Human Genetic Conference, May 27-30, 2017, Copenhagen, Denmark**, P12.136D
5. Hammoudeh Z., Nikolova D., Balabanski L., Antonova O., Vazharova R., Weidner S., Malinov M., Toncheva D "**Pharmacogenetic studies in**

patients with cancer". 12th Balkan Congress of Human Genetics, 8-10 September 2017, Plovdiv, Bulgaria

6. Hammoudeh Z., Nikolova D., Balabanski L., Antonova O., Vazharova R., Weidner S., Malinov M., Toncheva D. "**Screening of pharmacogenetic variants in patients with non-small-cell lung cancer & papillary thyroid cancer using cancer sequencing panel", First Balkan Conference on Personalized Medicine, Best Western Expo Sofia Hotel, Sofia**, 26-27.10.2017

IX. Благодарности

Благодаря сърдечно на научните си ръководители чл.-кор. проф. д-р Драга Тончева и доц. д-р Савина Хаджидекова за възможността да работя с едни от най-добрите екипи и най-съвременните технологии в областта на молекулярно-генетичните изследвания и за помощта и напътствията при разработването на този научен труд. Благодаря на научния си консултант д-р Рада Станева за ценните съвети, помощта при разработване на дисертационния труд и безрезервната подкрепа, надхвърлящи професионалните задължения. Благодаря на проф. Sabine Weidner за предоставените проби от пациенти с рак на щитовидната жлеза и на колегите от СБАЛГАР „Д-р Малинов" за помощта при експерименталните изследвания. Благодаря на гл. асистент Драгомира Николова за съдействието и интерпретацията на резултатите. Благодаря на всички колеги от Катедрата по Медицинска генетика за помощта, подкрепата и приятелската атмосфера. Благодаря на семейството ми за подкрепата, търпението и обичта.

X. Summary

Determination of variants associated with carcinogenesis and pharmacogenetic effects in lung cancer and thyroid cancer

This thesis aims to identify pharmacogenetic variants in two types of cancer patinets (non-small cell lung cancer and papillary thyroid cancer) by using NGS technology. Pharmacogenetic studies target on predicting the response of tumor tissue to standard therapy by genetic profiling and thus choosing the most appropriate drug at optimal dosages for each patient. Aditionaly, pharmacogenetics have the potential to elucidate adverse drug reactions and the positive influence of the drugs based on the genetic profile of the patients. Lung cancer and thyroid cancer are both from the most common cancers worldwide.

Lung cancer is divided into two types: Small cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC). Non-small cell carcinoma is about 85% of lung cancer, while small cell carcinoma is the other 15%. Somatic mutations in NSCLC are detected in *TP53*, *EGFR* and *KRAS* genes. Mutations in the *TP53* gene leds to a modified p53 protein, in result the modifie p53 can't bind to the DNA molecule and can't regulate cell proliferation proparlly, which allows damaged DNA to accumulate in cells. Mutations in the *EGFR* and *KRAS* genes result in the translation of pathologically activated proteins, resulting in continuous cell signal for proliferation.

The cases of papillary thyroid carcinoma (PTC) are increasing, thus makes it the predominant endocrine carcinoma. In papillary thyroid carcinoma, mutations in *RET*, *RAS* and *BRAF* genes are associated with the loss of ability to accept iodine.

The studies gathered in this dissertation explore genetic variants and mutations in NSCLC and PTC using NGS technology and real-time PCR analysis to determine pharmacogenetic variants related to the efficiency of the chemotherapy and target therapy with tyrosine kinase inhibitors. We performed DNA analysis on 581 samples (569 patients with NSCLC and 12 patients with PTC). DNA from NSCLC samples were extracted from formalin–fixed paraffin-embedded tissue and DNA from PTC samples were extracted from blood. Samples of 18 NSCLC and 12 PTC were sequenced by Illumina-MiSeq instrument using TruSight cancer panel of 94 genes and 284 SNPs. The results were analyzed using pharmgkb database (https://www.pharmgkb.org/) for SNP variants associated with sensitivity to certain drugs. We evaluate the presence of active EGFR mutations in 551 of NSCLC samples. To determine the types and the frequency of EGFR mutations and their association with smoking status, sex and histology we used real-time PCR (polymerase chain reaction) method. EGFR mutations were found in 55 NSCLC samples (10%). The frequency of EGFR mutations was greater for females versus males (27,1% versus 3,9%, $P < 0.001$), for never smokers 15% versus smokers 6,8% with $P < 0.003$, and the frequency for adenocarcinomas versus other histological types (squamous cell carcinoma and others) are 14,9% versus 5,1%; $P < 0.001$.