

Integrirana analiza tumora pluća skvamoznih stanica omics metodama

Šutić, Maja

Doctoral thesis / Disertacija

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Rijeka / Sveučilište u Rijeci**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:193:202089>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-04-25**



image not found or type unknown

Repository / Repozitorij:

[Repository of the University of Rijeka, Faculty of Biotechnology and Drug Development - BIOTECHRI Repository](#)



image not found or type unknown

SVEUČILIŠTE U RIJECI
ODJEL ZA BIOTEHNOLOGIJU

Maja Šutić

**INTEGRIRANA ANALIZA TUMORA PLUĆA
SKVAMOZNIH STANICA *OMICS*
METODAMA**

DOKTORSKI RAD

Rijeka, 2022.

SVEUČILIŠTE U RIJECI
ODJEL ZA BIOTEHNOLOGIJU

Maja Šutić

**INTEGRIRANA ANALIZA TUMORA PLUĆA
SKVAMOZNIH STANICA *OMICS*
METODAMA**

DOKTORSKI RAD

Mentor: Doc. dr. sc. Jelena Knežević
Komentor: Prof. dr. sc. Ljiljana Bulat-Kardum

Rijeka, 2022.

UNIVERSITY OF RIJEKA
BIOTECHNOLOGY DEPARTMENT

Maja Šutić

**INTEGRATIVE ANALYSIS OF SQUAMOUS
CELL LUNG CANCER BY USING OMICS
APPROACH**

DOCTORAL THESIS

Rijeka, 2022.

Mentor rada: Doc. dr. sc. Jelena Knežević

Doktorski rad obranjen je dana _____ na Sveučilištu u Rijeci, Odjel za biotehnologiju, pred povjerenstvom u sastavu:

- 1.
- 2.
- 3.
4. Doc. dr. sc. Jelena Knežević- mentor
5. Prof. dr. sc. Ljiljana Bulat-Kardum- komentor

ZAHVALE

Veliko hvala mentorici dr.sc. Jeleni Knežević na iskaznom povjerenju tijekom zapošljavanja na HRZZ projektu “Genetički i epigenetički biomarkeri urođene imunosti u KOPB-u i tumoru pluća” čime mi je dala priliku da pokažem svoju motivaciju i želju za radom. Zahvaljujem i na svim poticajima i pomoći bez koje ne bi bilo moguće uspješno realizirati izradu ove doktorske disertacije.

Zahvaljujem kolegama Jurici Baranašiću, Filipu Rokiću i Robertu Belužiću za svu pomoć, savjete i potporu što je neizmjereno pridonijelo kvaliteti ove disertacije.

Zahvaljujem voditelju Laboratorija za naprednu genomiku dr.sc. Oliveru Vugreku i svim ostalim članovima laboratorija, kolegama sa zavoda, instituta, fakulteta i stranih institucija u kojima sam boravila na poticajnom znanstvenom okruženju za rad kao i za svaki put kada su mi pomogli na bilo koji način.

Zahvaljujem članovima povjerenstva koji su uvelike zaslužni za poboljšanje kvalitete ove disertacije: komentorici Prof. dr. sc. Ljiljani Bulat-Kardum na velikoj pomoći u zadnjoj fazi izrade doktorske disertacije, zatim predsjednici komisije _____ na iscrpnoj recenziji te članovima komisije _____ na svim preporukama tijekom pisanja rada.

Zahvaljujem svim članovima svoje obitelji koji su me podupirali u svim aspektima života pa tako i u mojim akademskim postignućima

Zahvaljujem Mariu na apsolutnoj potpori i razumijevanju. Radujem se našim budućim zajedničkim uspjesima.

Eksperimentalni dio ove doktorske disertacije je izrađen u Laboratoriju za Naprednu genomiku Instituta Ruđer Bošković i financiran je projektom Hrvatske zaklade za znanost, HRZZ („Genetski i epigenetski biomarkeri urođene imunosti u KOPB-u i tumoru pluća“) pod vodstvom dr.sc. Jelene Knežević.

Razmjene u drugim znanstveno-istraživačkim institucijama su financirane na sljedeći način:

- Institut za istraživanje raka (DKFZ, Njemačka) –stipendija Instituta Ruđer Bošković za kratkoročni studijski boravak mladih istraživača u europskim laboratorijima
- Institut za istraživanje raka (DKFZ, Njemačka)- HRZZ projekt „Genetički i epigenetički biomarkeri urođene imunosti u KOPB-u i tumoru pluća“

SAŽETAK

Uvod i cilj

Tumor pluća je među najčešće dijagnosticiranim malignim oboljenjima na svijetu, sa visokom stopom smrtnosti. Tumor pluća skvamoznih stanica je relativno česti podtip tumora pluća koji je dijagnosticiran u 20-30 % slučajeva, a od njega je češće dijagnosticirani samo adenokarcinom pluća. Pacijenti oboljeli od tumora pluća skvamoznih stanica ne pokazuju uniformnost u kliničkom ishodu, i liječe se na jednak način. Obzirom da profil genskih varijanti, kao i specifični profil ekspresije gena, još uvijek nije u potpunosti razjašnjen u ovom tipu tumora, pacijenti se još uvijek uglavnom liječe sistemskom kemoterapijom. Pretpostavka je da bi bolja molekularna karakterizacija tumora pluća skvamoznih stanica prema biološkim karakteristikama, mogla dovesti do personaliziranog pristupa liječenju pacijenata, boljeg predviđanja odgovora na terapiju i u konačnici boljeg ishoda liječenja ove vrste tumora. Stoga su potrebna daljnja istraživanja koja će doprinijeti boljoj molekularnoj karakterizaciji ove vrste tumora.

Cilj ove doktorske disertacije je analiza i razumijevanje profila ekspresije molekula mRNA, proteina i miRNA u tumoru pluća skvamoznih stanica s krajnjom svrhom identifikacije ključnih molekula u tumorigenezi ove bolesti.

Materijali i metode

Za potrebe ovog istraživanja prikupljena su 23 uzorka tkiva tumora pluća skvamoznih stanica i 4 tkiva pluća zdravih osoba koja su fiksirana formalinom i uklopljena u parafin. Rezovi debljine 10 μ m korišteni su za pripremu knjižnica za sekvenciranje. Za analizu na spektrometru masa izolirani su proteini iz 3 tumorska i 3 zdrava uzorka, u trima biološkim replikama. Za identifikaciju gena sa promijenjenom ekspresijom korišten je alat DESeq2. Genima sa promijenjenom razinom ekspresije dodijeljene su anotacije pomoću analize genske ontologije i analize obogaćivanja setovima gena. Metodama ESTIMATE i GSVA procijenjen je stupanj infiltracije imunoloških stanica u tumorsko tkivo i ispitana je povezanost stupnja infiltracije s preživljenjem pacijenata. Pomoću analize proteinskih interakcija identificirani su klasteri u proteinskoj mreži te visoko umreženi geni. Pomoću WGCNA metode, ispitana je povezanost skupine gena objedinjenih u klastere sa kliničkim

parametrima pacijenata. Za validaciju rezultata genske ekspresije, ponovljena je DESeq2 analiza na uzorcima tumora pluća skvamoznih stanica sekvenciranih u sklopu TCGA projekta (223 ispitanika), i zdravih kontrola sekvenciranih u sklopu GTEx projekta (288 ispitanika). Učinak razine genske ekspresije na preživljenje pacijenata provjeren je na validacijskoj kohorti pomoću Kaplan-Meier krivulje preživljenja i log-rank testa. Diferencijalno eksprimirane miRNA identificirane su pomoću DESeq2 alata. Ciljni geni miRNA sa promijenjenom ekspresijom identificirani su pomoću multiMir paketa u R programu. Proteini sa promijenjenom razinom ekspresije u tumorskom tkivu identificirani su pomoću t-testa i programa Perseus. Rezultati ekspresije proteina potom su uspoređeni sa rezultatima ekspresije gena i miRNA.

Rezultati

Rezultati analize genske ontologije na obje ispitivane kohorte pokazuju da geni sa pozitivnom promjenom u razini ekspresije u tumorskom tkivu sudjeluju u regulaciji staničnog ciklusa, regulaciji stanične diobe, DNA metaboličkim procesima i diferencijaciji epitelnih stanica. Geni sa negativnom promjenom u razini ekspresije sudjeluju u regulaciji imunološkog sustava, poput regulacije lučenja citokina ili aktivacije leukocita, te u regulaciji apoptoze i stanične adhezije. Na temelju stupnja infiltracije tumora imunološkim stanicama identificirana su četiri imunosna podtipa, ali pripadnost određenom podtipu ne povezuje se sa preživljenjem pacijenata. Pomoću analize proteinskih interakcija identificirano je 11 klastera u proteinskoj mreži. Dva najveća klastera, Klaster 1 i Klaster 2, čine geni koji uglavnom sudjeluju u regulaciji staničnog ciklusa, regulaciji imunološkog odgovora i KRAS signalnom putu. Pomoću WGCNA metode identificiran je klaster *sienna3* koji se povezuje sa procesom metastaziranja. Unutar *sienna3* klastera identificirano je 12 gena koji mogu utjecati na progresiju tumora pluća skvamoznih stanica. Nadalje, pronađena je povezanost razine ekspresije 89 gena sa sveukupnim preživljenjem pacijenata. Geni HOXC4, GALNT2, PYGB, SMOX, C1orf106, SNR31 i FGL2 su identificirani kao novi potencijalni prognostički biomarkeri tumora pluća skvamoznih stanica. Analiza diferencijalne ekspresije miRNA pokazala je da 75 miRNA ima pozitivno a 47 miRNA negativno promijenjenu razinu ekspresije u tumorskom tkivu, u odnosu na zdrave kontrole. hsa-miR-577 ima najveću pozitivnu promjenu, a hsa-miR-490-3p najveću negativnu promjenu u razini ekspresije u

tumorskom tkivu. Interaktom mRNA-miRNA pokazuje da hsa-miR-182-5p, hsa-miR-590-3p, hsa-miR-107 i hsa-let-7c-5p imaju najveći broj potencijalnih ciljnih gena što ukazuje na njihovu potencijalno bitnu ulogu u tumoru pluća skvamoznih stanica. Rezultati analize diferencijalne ekspresije proteina pokazali su 1404 proteina ima pozitivnu, a 271 protein negativno promijenjenu ekspresiju u tumorskom tkivu, u odnosu na zdrave kontrole. Od navedenog broja, 245 prekomjerno eksprimiranih proteina, prekomjerno je eksprimirano i na transkripcijskoj razini, a za 144 gena i proteina zabilježena je smanjena ekspresija. Od 89 identificiranih prognostičkih gena, na proteinskoj razini potvrđena je prekomjerna ekspresija 9 proteina (GALNT2, PYGB, LAMB3, MCCC1, DCPS, DSG3, PC, FARSA i DTYMK), kao i smanjena ekspresija proteina SLC9A3RN.

Zaključci

Promjene razine mRNA, miRNA i proteina, koje su uočene u tumoru pluća skvamoznih stanica mogu biti odgovorne ili pridonositi tumorskoj patologiji. Prema dobivenim rezultatima, prekomjerna ekspresija gena koji sudjeluju u regulaciji staničnog ciklusa, kao i smanjena ekspresija gena koji sudjeluju u regulaciji imunološkog odgovora obilježja su tumora pluća skvamoznih stanica. Predložena je povezanost razine ekspresije 89 gena sa sveukupnim preživljenjem pacijenata s ciljem njihovog budućeg istraživanja u prognostičke svrhe. Predložena je tumor-supresorska uloga hsa-miR-107, hsa-miR-125 i hsa-let-7c-5p u tumoru pluća skvamoznih stanica.

Ključne riječi

Tumor pluća skvamoznih stanica, sekvenciranje slijedeće generacije, spektrometrija masa, prognostički biomarkeri

Ovaj doktorska disertacija sadrži 228 stranica, 70 slika, 35 tablica i 211 literaturnih navoda.

SUMMARY

Introduction and aims

Lung cancer is among the most commonly diagnosed malignancies in the world, and has a high mortality rate. Squamous cell lung cancer is a relatively common subtype of lung cancer, diagnosed in 20-30% of cases, of which only lung adenocarcinoma is more commonly diagnosed. Patients with squamous cell lung cancer do not show uniformity in clinical outcome, and are treated in the same way, regardless of morphological subtype. Because the profile of gene variants, as well as the specific profile of gene expression, is still not fully elucidated in this type of tumor, patients are still mostly treated with systemic chemotherapy. It is hypothesized that better molecular characterization of squamous cell lung cancer by biological characteristics could lead to a personalized approach to patient treatment, better prediction of response to therapy, and ultimately a better treatment outcome for this type of tumor. Therefore, further research is needed that will contribute to better molecular characterization of this type of tumor. The aim of this doctoral dissertation is to analyze and understand the expression profile of mRNA, protein and miRNA molecules in squamous cell lung cancer with the ultimate goal of identifying key molecules in the tumorigenesis of this disease.

Materials and methods

For the purposes of this study, 23 samples of squamous cell lung cancer tissue, together with 4 samples of lung tissue of healthy individuals were collected, fixed with formalin and embedded in paraffin. 10 µm thick sections were used to prepare the sequencing libraries. For mass spectrometry, proteins from 3 tumor and 3 healthy controls were isolated, in three replicates. The DESeq2 tool was used to identify genes with altered expression. Genes with altered expression levels were assigned annotations using gene ontology analysis and gene set enrichment analysis. ESTIMATE and GSVA methods were used to assess the degree of infiltration of immune cells into tumor tissue and to examine the relationship between the degree of infiltration and patient survival. The analysis of protein interactions identified clusters in the protein network and highly connected genes. Using the WGCNA method, the association of genes grouped in clusters and the clinical parameters of patients

was examined. To validate the gene expression results, the DESeq2 analysis was repeated on squamous cell lung cancer samples sequenced as part of the TCGA project (223 subjects), and healthy controls sequenced as part of the GTEx project (288 subjects). The effect of gene expression levels on patient survival was tested on a validation cohort using the Kaplan-Meier survival curve and log-rank test. Differentially expressed miRNAs were identified using the DESeq2 tool. Targeted miRNA genes with altered expression were identified using multiMir packages in the R program. Proteins with altered expression levels in tumor tissue were identified using the t-test and the Perseus program. Expression profiles of genes, proteins and miRNA were finally compared.

Results

The results of the gene ontology analysis, in both examined cohorts, show that genes with a positive change in the level of expression in tumor tissue participate in cell cycle regulation, regulation of cell division, DNA metabolic processes and epithelial cell differentiation. Genes with a negative change in the level of expression are involved in the regulation of the immune system, such as the regulation of cytokine secretion or leukocyte activation, and in the regulation of apoptosis and cell adhesion. Based on the degree of tumor infiltration by immune cells, four immune subtypes have been identified, but affiliation to a particular subtype is not associated with patient survival. By analyzing protein interactions, 11 clusters in the protein network were identified. The two largest clusters, Cluster 1 and Cluster 2, consist of genes that are mainly involved in cell cycle regulation, immune response regulation, and the KRAS signaling pathway. Using the WGCNA method, a *sienna3* cluster was identified that is associated with the metastasis process. Within the *sienna3* cluster, 12 genes have been identified that may influence squamous cell lung cancer progression. Furthermore, an association was found between the expression levels of 88 genes and the overall survival of patients. Genes HOXC4, GALNT2, PYGB, SMOX, C1orf106, SNR31 and FGL2 are proposed as new prognostic genes of squamous cell lung cancer. Analysis of differential miRNA expression showed that 75 miRNAs had positive and 47 miRNA negatively altered levels of expression in tumor tissue, compared to healthy controls. hsa-miR-577 has the largest positive change, and hsa-miR-490-3p has the largest negative change in the level of expression in tumor tissue. Interaction mRNA-miRNA shows that hsa-miR-

182-5p, hsa-miR-590-3p, hsa-miR-107 and hsa-miR-107 and hsa-let-7c-5p have the largest number of potential target genes indicating their potentially important role in squamous cell lung cancer. Differential expression analysis of proteins showed overexpression of 1404, and downregulation of 271 proteins in lung tissues. Among 1404 overexpressed proteins, 245 were also overexpressed on mRNA level. For 145 downregulated proteins in tumor tissues, 144 were downregulated on mRNA level as well. Out of 89 identified prognostic genes, overexpression on protein level was confirmed for 9 proteins (GALNT2, PYGB, LAMB3, MCCC1, DCPS, DSG3, PC, FARSA i DTYMK), as well as downregulation of SLC9A3RN.

Conclusions

Changes in mRNA, miRNA, and protein levels observed in squamous cell lung cancer may be responsible for or contribute to tumor pathology. According to the results obtained, overexpression of genes involved in cell cycle regulation and decreased expression of genes involved in the regulation of the immune response are characteristic of squamous cell lung cancer. The association of the expression level of 89 genes with the overall survival of patients with the aim of their future research for prognostic purposes has been proposed. hsa-miR-107, hsa-miR-125 and hsa-let-7c-5p are miRNA with proposed tumor suppressive role.

Keywords: Squamous cell lung cancer, next-generation sequencing, mass spectrometry, prognostic biomarkers

This doctoral thesis contains 228 pages, 70 images, 35 tables and 211 literature references.

SADRŽAJ

1. UVOD	1
1.1. Tumor pluća	3
1.2. Podjela tumora pluća	5
1.3. Histološka podjela tumora pluća	6
1.4. Podjela tumora pluća prema opsegu proširenosti	8
1.5. Etiologija i patofiziologija tumora pluća	10
1.6. Dijagnostika tumora pluća	12
1.7. Liječenje tumora pluća	13
1.8. Tumorski mikrookoliš	16
1.8.1. Karakteristike tumorskog mikrookoliša	16
1.8.2. Imunološke stanice u tumorskom mikrookolišu	17
1.9. Profiliranje molekularnih karakteristika tumora pluća skvamoznih stanica	19
2. CILJEVI I HIPOTEZA ISTRAŽIVANJA	24
3. MATERIJALI I METODE	26
3.1. Plan istraživanja	27
3.2. Ispitanici	28
3.3. Izolacija RNA iz parafinskih kocki	30
3.4. Provjera kvalitete izolirane RNA	31
3.5. Priprema knjižnica za sekvenciranje	32
3.6. Bioinformatička obrada podataka RNA-Seq analize	36
3.6.1. Provjera kvalitete neobrađenih sekvenci	37

3.6.2. Poravnavanje neobrađenih sekvenci na referentni genom GRCh38.....	40
3.6.3. Analiza diferencijalne ekspresije gena.....	42
3.6.4. Analiza genske ontologije i GSEA analiza.....	44
3.6.5. Analiza mreže proteinskih interakcija, analiza modula i identifikacija ključnih gena.....	47
3.6.6. Analiza ponderirane korelacijske mreže (engl. <i>Weighted gene co-expression network analysis</i> , WGCNA).....	48
3.6.7. Analiza sveukupnog preživljenja.....	51
3.6.8. Procjena razine infiltriranosti tumora sa imunološkim stanicama i povezanost sa sveukupnim preživljenjem.....	52
3.7. Analiza profila ekspresije malih molekula RNA.....	54
3.7.1. Provjera kvalitete neobrađenih sekvenci.....	54
3.7.2. Poravnavanje neobrađenih sekvenci na referentni genom i identifikacija miRNA.....	55
3.7.3. Analiza promijenjeno eksprimiranih miRNA.....	56
3.7.4. Predikcija ciljnih gena miRNA.....	56
3.7.5. Analiza interakcija miRNA-mRNA.....	57
3.8. Analiza proteoma.....	58
3.8.1. Izolacija proteina iz FFPE rezova tkiva.....	58
3.8.2. Određivanje koncentracije proteina.....	59
3.8.3. Priprema uzoraka za spektrometriju masa.....	59
3.8.4. Identifikacija proteina.....	61

3.8.5. Provjera kvalitete proteomskih podataka i identifikacija diferencijalno eksprimiranih proteina.....	61
3.8.6. Usporedba razine ekspresije mRNA, proteina i miRNA.....	62
4. REZULTATI.....	63
4.1. Rezultati analize transkriptoma	64
4.1.1. Rezultati provjere kvalitete neobrađenih sekvenci.....	64
4.1.2. Rezultati poravnavanje neobrađenih sekvenci na referentni genom GRCh38.....	69
4.1.3. Rezultati analize diferencijalne ekspresije gena.....	71
4.1.4. Rezultati analize genske ontologije i analize GSEA	73
4.1.5. Rezultati analize mreže proteinskih interakcija, analize modula i identifikacije ključnih gena.....	83
4.1.6. Rezultati analize ponderirane korelacijske mreže (engl. <i>Weighted gene co-expression network analysis</i> , WGCNA).....	91
4.1.7. Rezultati procjene razine infiltriranosti tumora sa imunološkim stanicama i povezanost sa sveukupnim preživljenjem.....	100
4.1.8. Rezultati validacije transkriptomskih analiza na TCGA setu podataka.....	105
4.1.9. Rezultati utjecaja razine ekspresije gena na preživljenje.....	114
4.2. Rezultati analize malih RNA molekula.....	118
4.2.1. Rezultati provjere kvalitete neobrađenih sekvenci i uklanjanja sekvenci adaptera.....	118
4.2.2. Rezultati poravnavanja neobrađenih sekvenci na referentni genom GRCh38_p10.....	122
4.2.3. Rezultati analize diferencijalne ekspresije miRNA.....	124

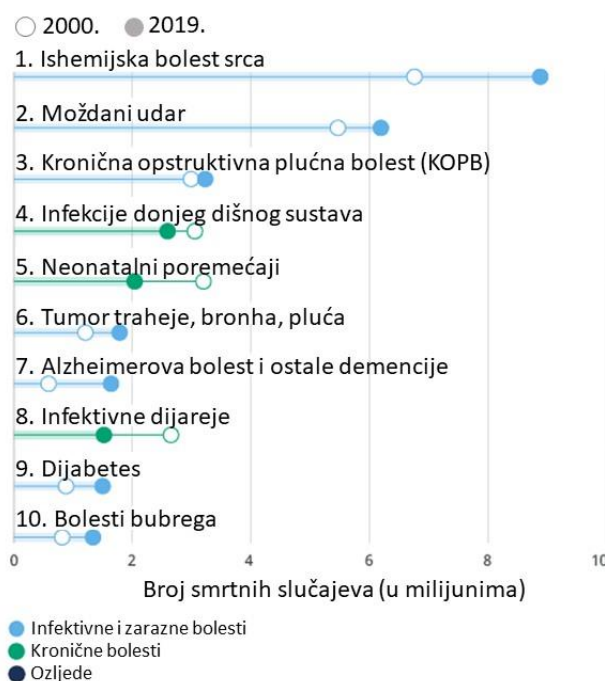
4.2.4. Rezultati predikcije ciljnih gena miRNA sa promijenjenom razinom ekspresije i interakcije mRNA-miRNA.....	126
4.3. Rezultati analize proteoma.....	131
4.3.1. Rezultati identifikacije proteina i provjere kvalitete podataka.....	131
4.3.2. Rezultati analize diferencijalne ekspresije proteina.....	136
4.3.3. Rezultati usporedbe ekspresije proteina sa ekspresijom gena i miRNA.....	138
5. RASPRAVA.....	143
5.1. Transkriptomski profil tumora pluća skvamoznih stanica.....	145
5.2. Klasteri gena i visokoumreženi geni u tumoru pluća skvamoznih stanica.....	149
5.3. Utjecaj razine ekspresije gena na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica.....	153
5.4. Profil tumorskog mikrokoliša i utjecaj na preživljenje pacijenata.....	156
5.5. Profil ekspresije miRNA u tumoru pluća skvamoznih stanica.....	159
5.6. Profil ekspresije proteina u tumoru pluća skvamoznih stanica.....	161
5.7. Interakcija mRNA, proteina i miRNA u tumoru pluća skvamoznih stanica...	163
6. ZAKLJUČCI.....	165
7. LITERATURA.....	168
8. POPIS KRATICA.....	194
9. POPIS ILUSTRACIJA.....	200
9.1. Slike.....	201
9.2. Tablice.....	205
10. ŽIVOTOPIS.....	208

1. UVOD

5 slika, 4 tablice

1. UVOD

Početkom dvadesetog stoljeća tumor pluća bio je rijetko dijagnosticirana bolest. Danas je tumor pluća jedan od najčešće dijagnosticiranih vrsta tumora (Slika 1.) i vodeći uzrok smrtnosti od malignih bolesti [1,2]. Mnogobrojna znanstvena istraživanja usmjerena su na pronalazak uzroka tumora pluća, kao i na ispitivanje novih mogućnosti liječenja ove bolesti, ali unatoč svim dosadašnjim znanstvenim naporima, smrtnost od ove bolesti i dalje je visoka. Visokoj stopi smrtnosti pridonosi i činjenica da se samo petina (20 %) slučajeva tumora pluća dijagnosticira u ranoj fazi bolesti, kada je tumor lokaliziran i nije zahvatio obližnje limfne čvorove, i kada je očekivano petogodišnje preživljenje najviše (>60 %), u odnosu na kasnije faze bolesti [3]. Liječenje pacijenata sa novim vrstama terapije, poput ciljane terapije ili imunoterapije, dovelo je do boljeg preživljenja, ali za relativno malu skupinu pacijenata [4]. Stoga je smrtnost od tumora pluća još uvijek vrlo visoka.



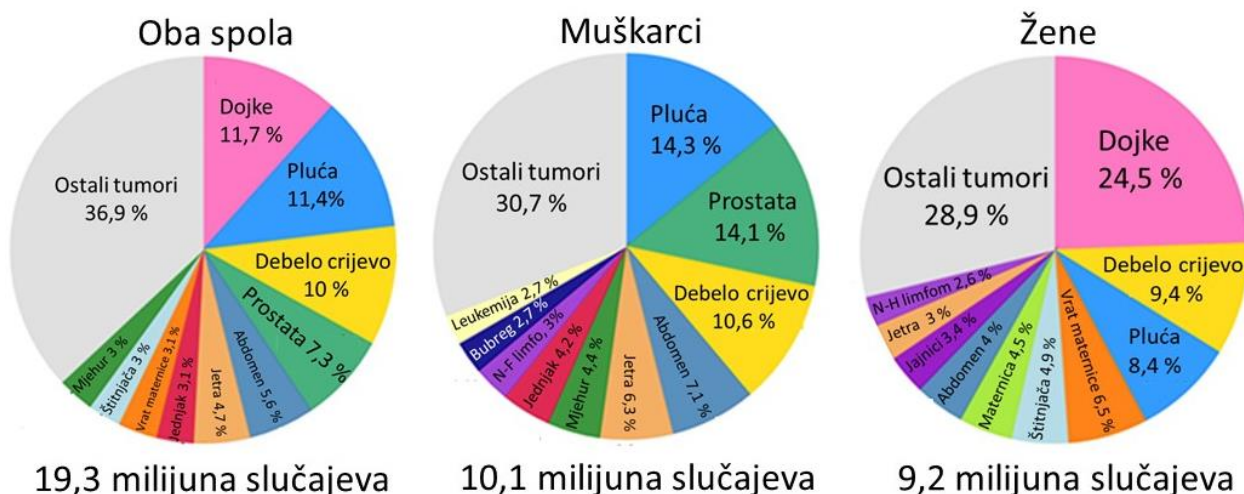
Slika 1. 10 najčešćih uzročnika smrti na svjetskoj razini u razdoblju 2000.-2019. godine. Prazan krug označava broj smrtnih slučajeva 2000. godine, a puni krug broj smrtnih slučajeva 2019. godine. Preuzeto i prilagođeno iz: Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. Geneva, World Health Organization; 2020 [5].

1.1. TUMOR PLUĆA

Tumor pluća je jedan od najčešće dijagnosticiranih malignih oboljenja na svijetu u oba spola. Prema GLOBOCAN bazi podataka, procijenjeno je da će se u svijetu 2020. godine dijagnosticirati 2,2 milijuna novih slučajeva tumora pluća. U Hrvatskoj se na godišnjoj razini dijagnosticira oko 3000 novih slučajeva, što tumor pluća čini drugim najčešće dijagnosticiranim oblikom tumora kod muškaraca (nakon tumora prostate) i trećim kod žena (nakon tumora dojke i debelog crijeva) [6] (Slika 2).

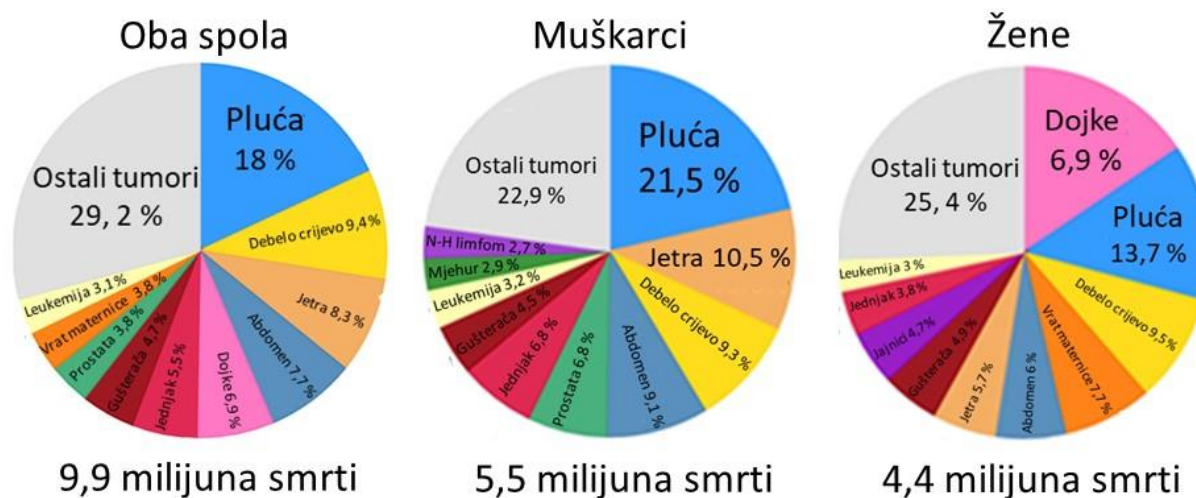
Maligna oboljenja su među vodećim uzrocima smrtnosti u cijelome svijetu. Prema procjenama Svjetske zdravstvene organizacije (SZO), u 112 od 183 zemalja svijeta (61 %), tumori su vodeći ili drugi po redu uzrok smrti prije napunjene sedamdesete godine života [7]. U odnosu na ostale vrste tumora, najveći broj zabilježenih smrti pripisuje se tumoru pluća (Slika 3.), što ga čini vodećim uzrokom smrtnosti od malignih bolesti kod muškaraca i drugim uzrokom smrtnosti kod žena (nakon tumora dojke), zbog čega tumor pluća predstavlja veliki javnozdravstveni problem na globalnoj razini [2].

Stopa incidencije tumora prema spolu



Slika 2. Globalna stopa incidencije 10 najčešćih vrsta tumora u 2020. Na lijevoj strani prikazana je stopa incidencije tumora za oba spola, u sredini za muškarce, a na desnoj strani za žene. N-H limfom- Ne-Hodgkinov limfom. Preuzeto i prilagođeno iz: Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. 2018, 68, 394–424, doi:10.3322/caac.21492 [2].

Stopa smrtnosti od tumora prema spolu



Slika 3. Stopa smrtnosti od 10 najčešćih vrsta tumora u 2020. Na lijevoj strani prikazana je stopa smrtnosti od tumora za oba spola, u sredini za muškarce, a na desnoj strani za žene. N-H limfom-Ne-Hodgkinov limfom. Preuzeto i prilagođeno iz: Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. 2018, 68, 394–424, doi:10.3322/caac.21492.

Unatoč brojnim znanstvenim istraživanjima temeljenima na pronalasku uzroka tumora pluća i pronalasku novih modaliteta liječenja ove bolesti, petogodišnje preživljenje pacijenata u većini zemalja svijeta i dalje je relativno nisko, oko 10-20 %. Razvijenije zemlje svijeta, poput Japana, Sjedinjenih Američkih Država ili Kanade, imaju veći postotak petogodišnjeg preživljenja (20-30 %), u odnosu na primjerice Hrvatsku u kojoj je ono 11 % [8]. Jedan od razloga koji tome pridonosi je i kasna detekcija same bolesti u većini zemalja svijeta. Gotovo 50 % dijagnoza tumora pluća postavi se u trenutku kada se tumor već proširio na udaljene limfne čvorove ili udaljene organe [3], što za posljedicu ima preživljenje pacijenata od svega 18 mjeseci od postavljene dijagnoze [4]. Liječenje tumora pluća novim vrstama terapija, poput ciljane terapije ili imunoterapije, dovelo je do boljeg ishoda liječenja za dio pacijenata koji su kandidati za navedene terapije. Pacijenti koji nisu kandidati za navedene terapije i dalje se liječe sistemskom kemoterapijom [9].

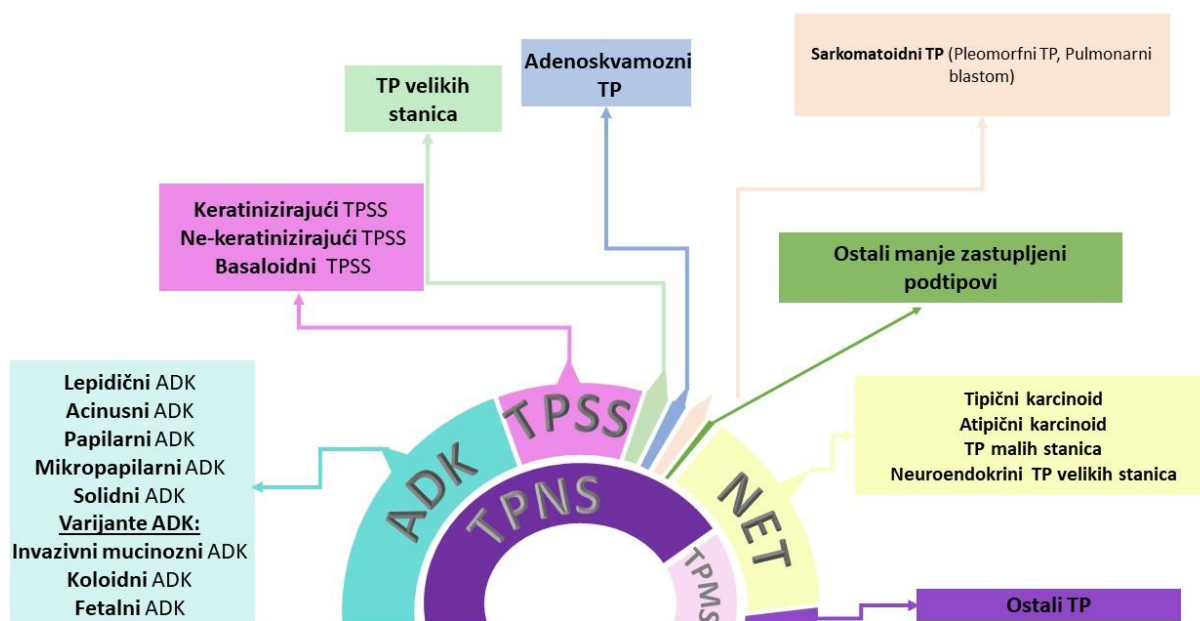
1.2. Podjela tumora pluća

Razvoj i primjena naprednih visokoprotočnih tehnologija u znanstvenim istraživanjima, poput sekvenciranja slijedeće generacije, doprinijeli su novim znanstvenim spoznajama i boljem razumijevanju biologije i patogeneze tumora pluća. Nove spoznaje o pokretačkim mutacijama [10] koje doprinose nastanku tumora pluća, te utječu na rast i diobu tumorskih stanica, omogućile su razvoj ciljane terapije za tumor pluća. Detekcija pokretačkih mutacija u određenim histološkim tipovima tumora pluća pretpostavlja povoljan odgovor bolesnika na ciljanu terapiju. Obzirom da je profil identificiranih pokretačkih mutacija različit u podtipovima tumora pluća [11,12], a najveći broj lijekova je odobren za liječenje adenokarcinoma, postojeći odobreni lijekovi nisu primjenjivi za liječenje svih pacijenata. Upravo iz toga razloga važno je prilikom dijagnoze tumora pluća utvrditi točno koji podtip tumora pluća se razvio u pacijentu.

Tumor pluća se na temelju morfoloških karakteristika tradicionalno dijeli u dvije glavne skupine: tumor pluća nemalih stanica (engl. *non-small cell lung cancer*, *NSCLC*) i tumor pluća malih stanica (*small cell lung cancer*, *SCLC*). Tumor pluća malih stanica je neuroendokrini tumor pluća visokog stupnja diferencijacije, zbog čega ima karakteristike agresivnog širenja i najlošiju prognozu u odnosu na ostale vrste tumora pluća [13]. Tumor pluća malih stanica čini 15-20 % dijagnoza tumora pluća, dok je tumor pluća nemalih stanica puno češći sa 80-85 % od ukupnih dijagnoza tumora pluća. Tumor pluća nemalih stanica se nadalje klasificira u nekoliko podskupina. Najčešća podskupina NSCLC-a u oba spola je adenokarcinom pluća, sa incidencijom oko 40 %. Tumor pluća skvamoznih stanica, koji je do ranih 1990-ih godina bio najčešće dijagnosticirani tip tumora pluća, danas čini 20-30 % od ukupnih dijagnoza tumora pluća, a po udjelu dijagnoza slijedi ga tumor pluća velikih stanica sa 5-10 % dijagnoza [14]. Obzirom da su sve do zadnjih desetak godina opcije za liječenje tumora pluća bile limitirane, daljnja podjela tumora pluća nije bila potrebna. Prikupljanjem novih znanstvenih saznanja o tumoru pluća, mijenjala se i klasifikacija tumora pluća kroz povijest. Tradicionalnoj podjeli je tako dodana nova kategorija tumora pluća nazvana neuroendokrini tumori pluća. Novije klasifikacije uzimaju u obzir i poznate pokretačke mutacije za koje postoje odobreni lijekovi.

1.3. Histološka podjela tumora pluća

Histološka klasifikacija tumora pluća je kompleksna, obzirom da tumor pluća predstavlja heterogenu skupinu tumora koja broji više od pedeset histomorfoloških podtipova. Prema posljednjoj klasifikaciji Svjetske zdravstvene organizacije iz 2021. godine [15] (Slika 4.) tumori pluća dijele se na adenokarcinom pluća, tumor pluća skvamoznih stanica, tumor pluća velikih stanica, adenoskvamozni tumor pluća, tumor pluća vretenastih stanica (engl. *spindle cell*), pleomorfni, tumor pluća orijaških stanica (engl. *giant cell*), manje zastupljeni tumori pluća i neuroendokrini tumori pluća među koje se ubraja i ranije navedeni tumor pluća malih stanica. Navedene skupine tumora nadalje se dijele u podskupine. Primjerice tumor pluća skvamoznih stanica čine keratinizirajuća, nekeratinizirajuća i bazaloidna podskupina tumora pluća.

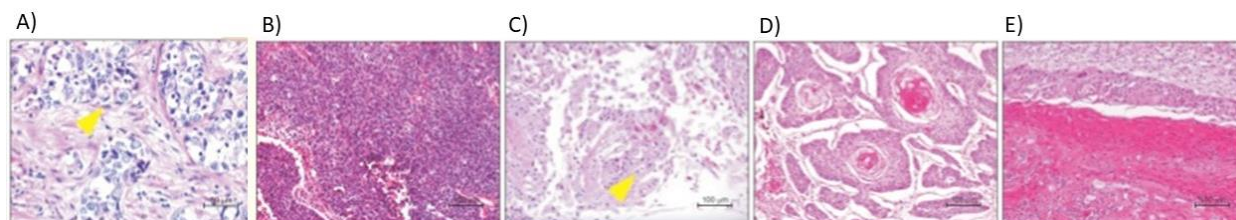


Slika 4. Podjela tumora pluća prema klasifikaciji Svjetske zdravstvene organizacije iz 2021. ADK- adenokarcinom, TPSS- tumor pluća skvamoznih stanica, NET- neuroendokrini tumor pluća, TP- tumor pluća. Preuzeto iz: Šutić, M.; Vukić, A.; Baranašić, J.; Försti, A.; Džubur, F.; Samaržija, M.; Jakopović, M.; Brčić, L.; Knežević, J. Diagnostic, Predictive, and Prognostic Biomarkers in Non-Small Cell Lung Cancer (NSCLC) Management. *J. Pers. Med.* **2021**, *11*, 1102, doi:10.3390/JPM11111102 [16]

Adenokarcinom je najčešći oblik tumora pluća, dijagnosticiran u 40-50 % slučajeva tumora pluća. Adenokarcinom pluća je maligna epitelna neoplazma sa žljezdanom diferencijacijom ili

sekrecijom mucina [17]. Nastaje većinom u perifernim dijelovima pluća, ali može biti i centralno smješten u plućima. Adenokarcinom pluća uglavnom eksprimira pneumocitne markere, poput *Thyroid transcription factor* (TTF-1) i *Napsin A* čija se ekspresija može provjeriti imunohistokemijski [14]. Prema morfološkim osobinama adenokarcinom pluća dijeli su u nekoliko podtipova (Slika 4.) poput solidnog adenokarcinoma, lepidičnog adenokarcinoma, acinarnog adenokarcinoma, mikropapilarnog i papilarnog adenokarcinoma. Subklasifikacija adenokarcinoma na temelju morfoloških obilježja tumora utječe na očekivanu prognozu. Primjerice, najbolja prognoza očekuje se za pacijente oboljele od lepidičnog adenokarcinoma, zatim acinarnog i papilarnog, dok se najlošija prognoza očekuje za solidne i mikropapilarne adenokarcinome [18].

Tumor pluća skvamoznih stanica do 1990.-ih godina bio je najzastupljeniji podtip tumora pluća, nakon čega se njegova prevalencija smanjuje. Danas je tumor pluća skvamoznih stanica drugi po učestalosti tumora pluća, sa 20-30 % od svih postavljenih dijagnoza [14]. Obično se javlja u središnjem dijelu pluća, a tumorigeneza započinje metaplazijom bronhalnog epitela u pločasti epitel. Potom slijedi stvaranje displastičnog pločastog epitela koji napreduje u preinvazivni, a zatim i u invazivni tumor. Tumor pluća skvamoznih stanica morfološki je često okarakteriziran keratinizacijom. Može doći do keratinizacije pojedinačnih tumorskih stanica ili stvaranja nakupina keratinizirajućih pločastih stanica. Također, tumorske stanice obično imaju hiperkromatske jezgre i međustanične mostove (Slika 5.), bez stvaranja strukture žljezdanih stanica ili lučenja mucina. Ukoliko morfološke karakteristike nisu vidljive na mikroskopskom preparatu nakon bojanja bojama hematoksilin-eozin, pločastu diferencijaciju potrebno je dokazati imunohistokemijski pomoću protutijela za citokeratin-5, citokeratin-6, p40 ili p63 [17].



Slika 5. Histološke karakteristike tumora pluća skvamoznih stanica. A) Prisutnost hiperkromatskih jezgri, B) Tumor pluća skvamoznih stanica bez prisutnosti keratinizacije, C) Citoplazmatska keratinizacija, D) Keratinske nakupine, E) Keratinizacija u slojevima

Preuzeto iz: Park, H.J.; Cha, Y.J.; Kim, S.H.; Kim, A.; Kim, E.Y.; Chang, Y.S. Keratinization of Lung Squamous Cell Carcinoma Is Associated with Poor Clinical Outcome. *Tuberc. Respir. Dis. (Seoul)*. **2017**, 80, 179, doi:10.4046/TRD.2017.80.2.179 [19]

Tumor pluća velikih stanica dijagnosticiran je u 5-10 % slučajeva. Nastaje iz epitelnih stanica, u perifernim dijelovima pluća i slabo je diferenciran. Karakteriziraju ga velike, poligonalne stanice [17].

Neuroendokrini tumori pluća su relativno česta skupina tumora pluća, sa 20-25 % udjelom od ukupnih dijagnoza. Morfološke karakteristike ih jasno odvajaju od ostalih vrsta tumora pluća. Karakteriziraju ih organoidni uzorak rasta, granulirani kromatinski uzorak i ekspresija neuroendokrinih markera poput kromogranin A, sinatofizin i CD56. Kao što je prikazano na Slici 4. prema klasifikaciji SZO-a, neuroendokrini tumori nadalje se dijele u četiri glavne kategorije, uključujući tipični i atipični karcinoid pluća, tumor pluća malih stanica i neuroendokrini tumor velikih stanica [17].

1.4. Podjela tumora pluća prema opsegu proširenosti

Tumori pluća se, osim prema morfološkim karakteristikama, često dijele i prema opsegu proširenosti. U klinici se najčešće upotrebljava TNM klasifikacija tumora pluća koja se bazira na karakteristikama primarnog tumora, poput veličine tumora (T, iz engl. *tumor*), zahvaćenosti limfnih čvorova tumorom (N, iz engl. *nodes*) i prisutnosti metastaza (M, iz engl. *metastasis*) [20]. Svaki od TNM parametara podijeljen je u nekoliko skupina koji opisuju opseg proširenosti tumora (Tablica 1.).

Na temelju informacija o sva tri navedena parametra, određuje se stadij bolesti na način kako je prikazano u Tablici 2. Određivanje stadija tumora pluća važno je za procjenu prognoze bolesti, kao i za odabir vrste liječenja pacijenta. Najveći stadij tumora pluća, IV B, ukazuje na proširenu bolest sa višestrukim metastazama i ima najlošiju prognozu u odnosu na ranije stadije bolesti.

Tablica 1. Opis TNM parametara na temelju osmog izdanja TNM klasifikacije iz 2017. godine. Preuzeto i prilagođeno iz: Detterbeck, F.C.; Boffa, D.J.; Kim, A.W.; Tanoue, L.T. The Eighth Edition Lung Cancer Stage Classification. *Chest* 2017, 151, 193–203 [20].

T (primarni tumor)	
T0	Bez prisutnosti primarnog tumora
Tis	Tumor <i>in situ</i>
T1	
T1mi	Minimalno invazivni adenokarcinom
T1a	Tumor ≤ 1 cm
T1b	Tumor > 1 cm, ali ≤ 2 cm
T1c	Tumor > 2 cm, ali ≤ 3 cm
T2	
T2a	Tumor > 3 cm, ali ≤ 4 cm
T2b	Tumor > 4 cm, ali ≤ 5 cm
T3	Tumor > 5 cm, ali ≤ 7 cm ili tumor koji se proširio u područje perikarda, ili limfne čvorove istok plućnog reznja
N (regionalni limfni čvorovi)	
N0	Tumor ne zahvaća regionalne limfne čvorove
N1	Tumor zahvaća ipsilateralne ili hilarne limfne čvorove
N2	Tumor zahvaća ipsilateralne ili subkarinalne limfne čvorove
N3	Tumor zahvaća kontralateralne, hilarne ili supraklavikularne limfne čvorove
M (udaljene metastaze)	
Mo	Bez metastaza
M1a	Maligna pleuralna ili perikardijalna efuzija; pleuralna ili perikardijalni ili odvojeni tumorski noduli u kontralateralni režanj
M1b	Jedna metastaza izvan torakalnog područja
M1c	Više metastaza izvan torakalnog područja

Tablica 2. Određivanje stadija tumora pluća prema TNM parametrima. Preuzeto i prilagođeno iz Detterbeck, F.C.; Boffa, D.J.; Kim, A.W.; Tanoue, L.T. The Eighth Edition Lung Cancer Stage Classification. *Chest* 2017, 151, 193–203: [20]

T/M	Podkategorija	N0	N1	N2	N3
T1	T1a	IA1	IIB	IIIA	IIIB
	T1b	IA2	IIB	IIIA	IIIB
	T1c	IA3	IIB	IIIA	IIIB
T2	T2a	IB	IIB	IIIA	IIIB
	T2b	IIA	IIB	IIIA	IIIB
T3	T3	IIB	IIIA	IIIB	IIIC
T4	T4	IIIA	IIIA	IIIB	IIIC
M1	M1a	IVA	IVA	IVA	IVA
	M1b	IVA	IVA	IVA	IVA
	M1c	IVB	IVB	IVB	IVB

1.5. Etiologija i patofiziologija tumora pluća

Tumor pluća je kompleksna bolest i na njegov nastanak i progresiju utječu brojni faktori, od genetske podloge pojedinca do utjecaja okoliša u kojem pojedinac živi. Tumor pluća razvija se u nekoliko koraka koji podrazumijevaju patološke promjene na epitelnim stanicama pluća uzrokovane nakupljanjem genetskih i epigenetskih promjena [20].

Studije su pokazale da epitelne stanice koje oblažu respiratorni trakt, pod utjecajem rizičnih čimbenika, poput pušenja, nakupljaju brojne molekularne promjene koje uzrokuju nastajanje tumora. Plućni epitel pritom prolazi kroz nekoliko faza morfoloških promjena nazvanih hiperplazija, metaplazija, displazija i karcinom *in situ*. Daljnjom proliferacijom transformiranog epitela, premaligne lezije mogu prijeći u invazivni karcinom [21].

Tumor pluća dijeli glavna obilježja sa drugim solidnim tumorima. Karakteriziraju ga povećana proliferacija tumorskih stanica, izbjegavanje supresora rasta, induciranje procesa angiogeneze, izbjegavanje apoptoze, deregulacija stanične energije, što je često praćeno upalom koja potiče daljnji rast tumora, kao i mnogobrojnim akumuliranim mutacijama u tumorskim stanicama [22]. Biološke karakteristike tumora pluća sažete su u Tablici 3.

Početkom dvadesetog stoljeća tumor pluća je bio rijetka bolest. Primjerice, 1912. godine Adler I. opisao je 374 slučajeva primarnog tumora pluća prilikom autopsije pacijenata iz SAD-a i zapadne Europe, što odgovara 0,5 % udjelu tumora pluća u odnosu na ostale maligne bolesti [1]. Porastom broja aktivnih pušača, kao i porastom broja popušanih cigareta po osobi, primijećena je i veća incidencija tumora pluća. 1964. Zavod za javno zdravstvo SAD-a izvijestio je o štetnim utjecajima pušenja na zdravlje i zaključio da je pušenje povezano sa povećanjem broja smrti i do 70 % kod muškaraca pušača i da pušenje predstavlja rizični faktor za nastanak tumora pluća [23]. Mehanizmi tumorigeneze uzrokovane pušenjem uključuju formaciju DNA adukata i njihovih metabolita, kao i oštećenja nastala djelovanjem slobodnih radikala [24], stoga pušači imaju 10-30 puta veće šanse za razvoj tumora pluća u odnosu na nepušače [25]. Osim pušenja, rizični čimbenici koji se povezuju sa nastankom tumora pluća su i dugotrajna izloženost jako zagađenom zraku, izloženost radioaktivnom zračenju, kao i profesionalna izloženost karcinogenima azbestu, kromu ili niklu [26].

Tablica 3. Biološke karakteristike tumora pluća. Preuzeto i prilagođeno iz: Chalela, R.; Curull, V.; Enríquez, C.; Pijuan, L.; Bellosillo, B.; Gea, J. Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy. *J. Thorac. Dis.* **2017**, 9, 2142–2158, doi:10.21037/JTD.2017.06.20. [22]

Karakteristika	Zdrave stanice	Tumorske stanice	Liječenje
Proliferacija stanica	Dioba stanice počinje lučenjem proliferacijskih signala (prema potrebi)	Proliferacijski signali konstantno aktivni	EGFR inhibitori
Izbjegavanje supresora rasta	Supresori rasta inhibiraju neželjenu proliferaciju stanica	Inhibicija supresora rasta što omogućava daljnji rast tumora	Inhibitori ciklin-ovisnih kinaza
Poticanje angiogeneze	Lučenje VEGF-a po potrebi	Pojačano lučenje VEGF-a zbog veće potrebe za nutrijentima	Inhibitori VEGF signalnog puta
Replikacijska besmrtnost	Limitirana replikacija zbog progresivnog gubitka telomeraza svakom staničnom diobom	Neograničena replikacija omogućena lučenjem telomeraze koja omogućava replikaciju telomera	Inhibitori telomera
Izbjegavanje stanične smrti	Programirana stanična smrt (apoptoza) i nekroza uklanjaju stanice sa oštećenjem DNA	Izbjegavanje apoptoze, pojačana proliferacija, progresija tumora i rezistencija na terapiju	Lijekovi koji imitiraju pro-apoptotski BH3
Invazija i metastaziranje	Organizirani rast sa granicama diferencijacije	Tumor je prošao tkivne barijere i može se širiti na druge organe ili vaskularno i limfatičko tkivo	Inhibitori HGF/ c-MET
Izbjegavanje imunološkog odgovora	T limfociti detektiraju abnormalne stanice pomoću površinskih markera i uništavaju ih	Imunološki odgovor izbjegava se na različite načine, uglavnom smanjenom ekspresijom površinskih markera	Imunoterapija
Deregulacija stanične energije	Kisik iz krvotoka koristi se za iskorištavanje energije iz glukoze	Zbog neograničenog rasta troši se puno energije. Javlja se anaerobni metabolizam glukoze	Inhibitori aerobne glikolize

Iako se više od 80 % slučajeva tumora pluća dijagnosticira kod pušača [25], oko 20 % pušača zaista i razvije tumor pluća. Ova varijabilnost u razvoju tumora pluća kod pušača ukazuje da na razvoj tumora pluća utječu i ostali rizični čimbenici kojima su pojedinci izloženi.

Genetska predispozicija također ima ulogu u patogenezi tumora pluća. U zadnjih dvadesetak godina asocijacijske analize povezale su oko tisuću gena sa genetičkom predispozicijom za razvoj tumora pluća. Primjerice, cjelogenomska analiza povezanosti (engl. *Genome-wide association study*, GWAS) pronašla je povezanost rijetkih nasljednih varijanti BRCA2 i CHEK2 gena sa nastankom tumora pluća skvamoznih stanica [27]. Druge GWAS i genotipizacijske studije pronašle su povezanost gena RNASET2, SECISBP2L, NRG1, CHRNA2, OFBC1, RTEL1 [28], ARID1B i REL [29] sa razvojem tumora pluća, sugerirajući da je tumorigeneza pluća povezana sa modifikacijama kromatina, upalom i restrukturiranjem tumorskog mikrokoliša.

Osim akutnih upalnih bolesti pluća, poput upale pluća ili tuberkuloze [30], i kronične bolesti na plućima se povezuju sa povećanim rizikom od razvoja tumora. Primjerice, kronična opstruktivna bolest pluća (KOPB) [31] i intersticijska plućna fibroza (IPF) [32] se najčešće povezuju sa rizikom od nastanka tumora pluća, u usporedbi sa ostalim plućnim bolestima.

Navedene spoznaje ukazuju na važnu ulogu molekularnih mehanizama u nastanku tumora pluća. Dosadašnja istraživanja također ukazuju na potencijalnu važnost gena koji sudjeluju u regulaciji staničnog ciklusa, proliferaciji, popravku oštećene DNA u stanici, kao i regulaciji upalnog odgovora i apoptozi u tumorigenezi pluća [16]. Međutim, potrebna su dodatna istraživanja koja će identificirati ključne gene za nastanak tumora pluća.

Obzirom da je tumorigeneza pluća, barem djelomično, određena genetikom pojedinca, osobe sa obiteljskom anamnezom tumora pluća ili određenog genetskog profila, pripadaju rizičnoj skupini za razvoj tumora pluća. Uključivanje pojedinaca iz rizične skupine, kao i pušača, u *screening* programe za ranu detekciju tumora pluća, može dovesti do smanjenja smrtnosti uzrokovane ovom bolesti.

1.6. Dijagnostika tumora pluća

Inicijalna evaluacija pacijenata kod kojih postoji sumnja na tumor pluća, najčešće se radi pomoću rendgena pluća, računalne tomografije (PET) ili pozitronske emisijske tomografije (PET-CT). Rendgenska snimka može detektirati sumnjive tvorevine na plućima, ali nije dovoljna za sigurnu dijagnozu tumora pluća. Također, ima značajno visoku stopu lažno negativnih nalaza, sa osjetljivošću od približno 75 %. CT i PET-CT su osjetljiviji od rendgena, a ujedno mogu generirati snimke više organa čime se dobiva šira slika o lokalizaciji i opsegu proširenosti tumora [33]. Niskodozni CT (LDCT) također može imati ulogu u ranom otkrivanju tumora pluća u rizičnim skupinama.

Obzirom da pristup liječenju ovisi o tipu tumora pluća, potrebna je točna identifikacija histološkog podtipa tumora. Generalna diferencijacija između podtipova bazira se na prepoznavanju morfoloških obilježja tumorskog tkiva, zbog čega je potrebno prikupiti dio tumorskog tkiva iz pacijenta. U svrhu dijagnostičkih postupaka tkivo pluća može se prikupiti bronhoskopijom, citološkom punkcijom, torakoskopijom ili biopsijom. Za diferencijaciju na temelju morfoloških obilježja tumora, prikupljeni rezovi tumorskog tkiva boje se kombinacijom hematoksilina i eozina. Međutim, morfološka obilježja ponekad nisu dovoljna za pouzdanu

klasifikaciju tumora pluća kod pacijenta. To je najviše izraženo na primjeru slabo diferenciranih tumora ili u slučajevima kada tumor nema specifična morfološka ili fenotipska obilježja [34]. Stoga se preporučuje upotreba imunohistokemijske i imunocitokemijske metode klasifikacije koje koriste protutijela koja se specifično vežu na epitope stanica u uzorku. Primjerice TTF-1 i Napsin A markeri su za identifikaciju adenokarcinoma, dok se za tumor pluća skvamoznih stanica preporučuje korištenje protutijela za p40 [14].

Ciljana terapija pripisuje se ukoliko je u tumorskom tkivu detektirana neka od poznatih pokretačkih mutacija pomoću metoda molekularne biologije. Neke od metoda koje se koriste u dijagnostičke svrhe su PCR (engl. *polymerase chain reaction*), NGS (engl. *next generation sequencing*), FISH (engl. *fluorescence in situ hybridization*) i RT-PCR (engl. *reverse transcription-PCR*). Odabir metode ovisi o mutacijama koje se žele ispitati [35].

1.7. Liječenje tumora pluća

Uspješno liječenje tumora pluća ovisi o nekoliko važnih čimbenika: stadiju bolesti u trenutku dijagnoze, histološkom tipu tumora pluća i molekularnim karakteristikama tumora. Najbolje šanse za oporavak imaju pacijenti kojima je tumor dijagnosticiran u ranom stadiju, kada je još lokaliziran. U takvim slučajevima liječenje podrazumijeva uklanjanje tumorske mase operacijom, koja može biti praćena radioterapijom. Liječenje lokalno uznapredovalog tumora pluća koji zahvaća obližnje limfne čvorove, također uključuje ili samostalno uklanjanje tumora operacijom ili kombinaciju operacije i neoadjuvantne kemoterapije. Za liječenje ne-operabilnih tumora pluća stadija III, moguća je i primjena imunoterapije. Za liječenje metastatske bolesti, moguća je primjena sistemske kemoterapije bazirane na platini, kao i primjena ciljane terapije ili imunoterapije [36].

U prvoj polovici dvadesetog stoljeća dostupni oblici liječenja tumora pluća nisu bili efikasni i uglavnom se pacijentima nudila samo palijativna njega. Kako su operacije postajale sigurnije za pacijenta, postale su jedan od glavnih vrsta liječenja tumora pluća, zajedno sa radioterapijom [37]. Sedamdesetih godina dvadesetog stoljeća u liječenju tumora pluća počela je primjena prvih kemoterapeutika, poput metotreksata i doksorubicina, ali je klinički utjecaj na tijek bolesti bio skroman [38]. Pojava kemoterapije zasnovane na platini u osamdesetim i devedesetim godinama dvadesetog stoljeća, poput taksana, vinorelbina i gemcitabina, dovela je do značajnog povećanja stope preživljenja pacijenata u odnosu na palijativnu njegu [39] i na ranije

kemoterapeutike [40]. Posljednjih desetak godina ciljana terapija ili imunoterapija postaje sve češće primarni način liječenja, ukoliko su pacijenti kandidati za navedene terapije.

Imunoterapija se bazira na upotrebi protutijela koji blokiraju vezanje imunosupresivnih receptora ili liganada tumorskih stanica na T limfocite, čime se povećava citotoksična aktivnost T limfocita. U kliničkoj primjeni za liječenje tumora pluća koriste se inhibitori kontrolnih točaka PD-1 (engl. *Programmed cell Death protein*) i PD-L1 (engl. *Programmed Death-Ligand 1*), dok se inhibitori CTLA-4 (engl. *Cytotoxic T-Lymphocyte-Associated protein 4*) rijetko koriste [16]. PD-1 je receptor dominantno eksprimiran na aktiviranim citotoksičnim T limfocitima, dok se PD-L1 konstitutivno eksprimira na površini makrofaga, antigen-prezentirajućih stanicama, B i T limfocitima. Mehanizam supresije imunološkog odgovora do kojeg dolazi vezanjem PD-L1 liganda na PD-1 receptor bitan je u prevenciji autoimune reakcije organizma tijekom upale [41]. Međutim, ovaj mehanizam koriste i tumorske stanice koje ekspresijom PD-L1 liganda na svojoj površini i njegovim vezanjem na PD-1 receptor na citotoksičnim T limfocitima izbjegavaju imunološki odgovor organizma [42]. Iako je primjena imunoterapije dovela do značajnog produljenja vremena bez progresije bolesti i preživljenja pacijenata u odnosu na kemoterapiju [43], još uvijek nije u potpunosti razjašnjeno koji pacijenti su najbolji kandidati za ovu vrstu liječenja.

Ciljana terapija bazirana je na upotrebi lijekova koji inhibiraju djelovanje prekomjerno aktiviranih staničnih puteva u tumorskim stanicama uzrokovano raznim mutacijama. Obzirom da mutacije u tumorskim stanicama najčešće nastaju u genima koji direktno ili indirektno kontroliraju staničnu proliferaciju, primjenom ciljane terapije nekontrolirana stanična proliferacija se inhibira te se na taj način sprječava daljnji rast tumora. Do sada je poznato mnogo pokretačkih mutacija u tumoru pluća te postoji sve veći spektar lijekova koji su dostupni pacijentima koji imaju neku od mutacija [16]. U Tablici 4. sažete su najčešće mutacije za koje su dostupni lijekovi. Prema smjernicama nekoliko organizacija za istraživanje tumora pluća preporuča se rutinsko testiranje svih uznapredovalih tumora pluća koji imaju komponentu adenokarcinoma na mutacije u EGFR, ALK ili ROS1 genu. Ukoliko je moguće testiranje više gena simultano, uz navedene preporuča se i testiranje mutacija u RET, MET, Her2, KRAS i BRAF genu. Također, preporučuje se molekularno testiranje za pacijente mlađe od pedeset godina oboljele od tumora pluća skvamoznih stanica, kao i testiranje nepušača [44–46].

Tablica 4. Najčešće mutacije tumora pluća za koje je dostupna ciljana terapija. Preuzeto i prilagođeno iz: Šutić, M.; Vukić, A.; Baranašić, J.; Försti, A.; Džubur, F.; Samaržija, M.; Jakopović, M.; Brčić, L.; Knežević, J. Diagnostic, Predictive, and Prognostic Biomarkers in Non-Small Cell Lung Cancer (NSCLC) Management. *J. Pers. Med.* 2021, *11*, 1102, doi:10.3390/JPM11111102 [16]

Biomarker	Promjena	Ciljana terapija	Godina FDA odobrenja	Godina EMA odobrenja
EGFR	Delecija eksona 19	Erlotinib	2013	2011
		Gefitinib	2015	2009
	Ekson 21 (L858R) supstitucija	Afatinib	2013	2013
		Osimertinib	2018	2018
		Dacomitinib	2018	2019
		Erlotinib + ramucirumab	2020	
EGFR	T790M	Osimertinib	2015	2016
EGFR	L861Q, G719X, S768I	Afatinib	2018	
ALK	Promjene u ALK genu	Crizotinib Ceritinib Alectinib Brigatinib Lorlatinib	2011 2014 / 2017 2017 2017 2018	2012 / 2015 2015 / 2017 2017 2018 / 2020 2019
ROS1	Promjene u ROS1 genu	Crizotinib Entrectinib Ceritinib	2016 2019 2019	2016
BRAF	V600E mutacija	Dabrafenib + Trametinib	2017	2017
NTRK1/2/3	Fuzije gena	Larotrectinib Entrectinib	2018 2019	2019 2020
MET	Preskakanje eksona 14	Capmatinib Tepotinib	2020 2021	
RET	Promjene u RET genu	Selpercatinib Pralsetinib	2020 2020	

Iako je primjena imunoterapije i ciljane terapije dovela do boljeg preživljenja za dio pacijenata, za većinu pacijenata nije primjenjiva. Razlog tome je što je većina pacijenata negativna na mutacije u EGFR i ALK genima [47], dok su samo trećina pacijenata kandidati za primjenu imunoterapije. Pacijenti koji nisu kandidati za navedene terapije i dalje se liječe kemoterapijom baziranoj na platini u prvoj liniji liječenja [16].

1.8. Tumorski mikrookoliš

Dišni sustav je stalno izložen patogenima i česticama iz okoliša koji potencijalno mogu narušiti njegov integritet. U dišnom sustavu postoji nekoliko mehanizama zaštite koji pokušavaju spriječiti štetno djelovanje patogena i čestica te ih eliminirati iz dišnog sustava. Prvi mehanizam zaštite pluća od infekcije je sluz koja oblaže gornje dišne puteve a sadrži mnoge antimikrobne spojeve. Sluz luče vrčaste stanice dišnog sustava (engl. *goblet cells*), a izbacuju je epitelne stanice [48,49]. Epitelne stanice dišnog sustava također eksprimiraju receptore za prepoznavanje molekularnih uzoraka povezanih sa patogenima (engl. *Pathogen-associated molecular patherns*, PAMPs) ili česticama iz okoliša (engl. *Danger-associated molecular patherns*, DAMPs), pomoću kojih mogu modulirati i usmjeravati imunološki odgovor [48]. Navedene receptore također eksprimiraju i imunološke stanice u plućima, poput dendritičnih stanica, granulocita i alveolarnih makrofaga [48,50]. Aktivacija navedenih receptora i nizvodnih signalnih puteva omogućava brzo regrutiranje drugih imunoloških stanica u pluća, u svrhu efikasnog uklanjanja patogena. Limfociti, poput CD8+ i CD4+ T limfocita, neutrofil i prirodnoubilačke NK stanice također su prisutni u plućima [51]. Najrasprostranjenija podskupina CD4+ T limfocita u plućima su T pomoćničke stanice tipa I (Th1) koje luče interferon-gama (IFN- γ). T pomoćničke stanice tipa II (Th2), koje luče interleukin 4, i regulatorne T stanice (Treg), koje luče IL-10, otkrivene su u niskim razinama [48].

Održavanje homeostaze u plućnom tkivu je kompleksan posao u koji je uključen veliki broj stanica, citokina i kemokina. Međutim, kada je homeostaza u plućima narušena, kao kod upalnih stanja poput primjerice kronične opstruktivne bolesti pluća (KOPB) [52], promijenjeni mikrookoliš pluća pokazuje značajke koje mogu podržati tumorigenezu.

1.8.1. Karakteristike tumorskog mikrookoliša

Tumorske stanice nisu izolirane u tijelu, već su u stalnoj interakciji sa stanicama u okruženju. Stanice koje se nalaze u okruženju tumora nazivamo mikrookolišem tumora (engl. *tumor microenvironment*, TME)[53],[54]. TME čine ekstracelularni matriks (engl. *Extracellular matrix*, ECM), stromalne, vaskularne, endotelne stanice i stanice imunološkog sustava. U TME pluća, tumorske stanice mogu reprogramirati tumor-infiltrirajuće stromalne stanice da doprinose tumorigenezi [55],[56],[57]. U tumoru pluća nemalih stanica, stupanj infiltracije imunoloških

stanica u tumor ovisi o stadiju tumora [53],[54], što ukazuje da TME pridonosi tumorigenezi pluća i da se može koristiti kao indikator prognoze bolesti u pacijentima [57].

Upala je uključena u sve faze tumorigeneze, od maligne transformacije i inicijacije tumora do invazije tumora i metastaza. 2011. godine znanstvenici Hanahan i Weinberg, smatrajući da upala može promicati tumorigenezu, predložili su promicanje upale tumora kao jednu od 10 karakteristika tumora [58], [57]. Iako mehanizam kojim upala potiče tumorigenezu nije u potpunosti razjašnjen, pojavile su se dvije međusobno povezane hipoteze: jedna je unutarnji put vođen genetskim promjenama koje dovode do tumora i upale; drugi je vanjski put vođen upalnim stanjima koja povećavaju rizik od tumora [57]. Istraživanja upalne komponente TME pokazala su da intervencija u upalni TME može smanjiti razvoj tumora pluća kod životinjskih modela i pacijenata s tumorom pluća [59,60],[57].

Tumorske stanice nastoje izbjeći imunološki nadzor i uništenje kroz niz mehanizama. Na temelju mehanizama imunološke evazije, tijekom posljednjih desetljeća razvijen je i klinički primijenjen niz pristupa imunoterapiji. Za razliku od tradicionalne kemoterapije, imunoterapija djeluje uglavnom kroz iskorištavanje imunoloških stanica unutar ili izvan TME za posebno prepoznavanje i napad na stanice tumora, što teoretski čini da imunoterapijski pristup posjeduje veću specifičnost i manje nuspojave za pacijenta [61].

1.8.2. Imunološke stanice u tumorskom mikrookolišu

Dendritične stanice imaju sposobnost prezentiranja tumorskih antigena T limfocitima u limfnim čvorovima, zbog čega imaju važnu ulogu u borbi protiv tumora [62]. Zrele dendritične stanice eksprimiraju razne molekule koje stimuliraju aktivaciju T limfocita, dok rezidentne neaktivirane dendritične stanice ne mogu stimulirati T limfocite. Dendritične stanice mogu pod utjecajem tumora i tumorskog mikrookoliša promijeniti protutumorski fenotip u tumor tolerantni ili imunosupresivni fenotip. Primjerice, u tumoru pluća nemalih stanica dendritične stanice mogu pojačano eksprimirati molekulu B7-H3 koja inhibira aktivaciju T limfocita [63],[57]. Imunosupresivna uloga dendritičnih stanica u tumorskom mikrookolišu može se očitovati i lučenjem transformirajućeg faktora rasta (TGF- β) koji potiče diferencijaciju CD4⁺ limfocita u regulatorne T limfocite (Treg) koji onda potiskuju proliferaciju T limfocita [64],[57]. Manipulacijom fenotipa dendritičnih stanica može se utjecati na sastav i ulogu stanica u tumorskom okruženju što može dovesti do lokalnog i sistemskog protutumorskog odgovora. Primjerice, prva

faza kliničkih ispitivanja na pacijentima oboljelim od tumora pluća pokazala je da intratumoralna administracija adenoviralno transduciranih dendritičnih stanica sa pojačanom ekspresijom CCL21 (engl. *CC-chemokine ligand 21*), dovodi do indukcije sistemskog imunološkog odgovora na tumorspecifične antigene, pojačanu infiltraciju tumora sa CD8⁺ T limfocitima i pojačane ekspresije PD-L1 na tumorskim stanicama [65].

Tumoru pridruženi makrofazi (engl. tumor-associated macrophages, TAM) česte su komponente imunološkog infiltrata u tumoru pluća nemalih stanica [66]. Postoje dva fenotipa TAM-ova, M1 fenotip koji je aktiviran klasičnim putem, podupire upalu i ima protutumorsko djelovanje, dok je M2 fenotip aktiviran alternativnim putevima, podupire angiogenezu i imunosupresivan mikrookoliš [66],[67],[57]. Tumorske stanice mogu lučiti neke signalne molekule i interleukine kako bi regrutirale makrofage u tumorski mikrookoliš. Primjerice, otkriveno je da je ekspresija faktora stimulacije kolonije 1 (CSF-1) i interleukina-6 (IL-6) u tumoru pluća nemalih stanica usko povezana s infiltracijom TAM-a u tumorsku stromu i napredovanjem tumora pluća [68]. Regrutacija makrofaga u tumorski mikrookoliš također može biti posredovano i lučenjem citokina i kemokina IL-34, CCL2, CCL5 i CXCL4 iz tumorskih stanica [69].

Neutrofili su najzastupljenija vrsta imunoloških stanica u tumoru pluća nemalih stanica [54]. Neutrofili mogu izražavati jedan od dva fenotipa: N1 ili N2. Neutrofili sa N1 fenotipom pojačano izlučuju proupalne citokine i imaju protutumorsku ulogu. U ranim fazama razvoja tumora N1 neutrofili prevladavaju, međutim kako tumor napreduje, pokazalo se da se N2 neutrofili akumuliraju u korist N1 fenotipa. Mehanizam koji dovodi do ove promjene fenotipa nije u potpunosti razjašnjen. N2 neutrofili pokazuju pojačanu ekspresiju MMP-9 i imaju pro-tumorigenu ulogu [70]. Nakupljanjem N2 neutrofila u tumorski mikrookoliš povećava se proliferacija tumorskih stanica, angiogeneza i supresija protutumorskog imunološkog odgovora [70].

Prirodnoubilačke stanice (engl. *Natural killer cells, NK*) imaju bitnu ulogu u protutumorskom odgovoru jer mogu izravno prepoznati tumorske antigene [71]. Odgovor NK stanica u mikrookolišu tumora posredovan je transformirajućim faktorom rasta (TGFβ) koji posreduje polarizaciju NK stanica prema proangiogenom fenotipu. NK stanice koje su izgubile svoju protutumorsku sposobnost djelovanja karakterizira smanjena regulacija NK receptora, smanjena degranulacija i gubitak ekspresije IFNγ [72], [57].

Limfociti T koji infiltriraju tumore, posebice CD4⁺ i CD8⁺ T stanice, kao i njihovi imunoregulatorni citokini, izvršavaju ključne efektorske citotoksične funkcije u TME-u i

posreduju u odgovoru na inhibiciju imunoloških kontrolnih točki [57]. Protutumorsko djelovanje CD8⁺ citotoksičnih T stanica posredovano je lučenjem IFN γ , faktorom nekroze tumora (TNF) i granzimom B [73]. Stanice Treg mogu izravno potisnuti protutumorsko djelovanje stanica CD8⁺ T [74]. Nadalje, trajna izloženost antigenu može dovesti do disfunkcije stanica T ili iscrpljenosti koju karakterizira gubitak efektorskih i memorijskih funkcija [57]. Imunosupresivne stanice Treg inducirane su regulacijom prostaglandina E2 (PGE2) u tumorskim stanicama posredovanom ciklooksigenazom 2 (COX2). Smanjenje broja stanica Treg povezano je s povećanom ekspresijom granzima A, granzima B, perforina i IFN γ u limfocitima CD8⁺ T [74]. Stanice $\gamma\delta$ T igraju važnu ulogu u tumorigenezi i trenutno se istražuju kao meta imunoterapije tumora. Stanice $\gamma\delta$ T imaju dvojaku ulogu u tumorskom mikrookolišu. Protumorigeno djelovanje očituje se induciranjem angiogeneze, poticanjem proizvodnje vaskularnog endotelnog faktora rasta, regrutiranjem mijeloidnih supresorskih stanica i lučenjem citokina IL-4, IL-10 i TGF- β [75,76]. Njihovo protutumorsko djelovanje očituje se kroz sposobnost izravne citotoksičnosti (perforin/granzim, TRAIL, FasL i ADCC putevi), promicanjem sazrijevanja dendritičkih stanica, induciranjem Th1 i Th17 diferencijacije i aktiviranjem stanica CD8⁺ T i NK [75,77,78].

Limfociti B također mogu infiltrirati tumor. Limfociti B koji infiltriraju tumor, zajedno sa stanicama CD4⁺ T, nalaze se unutar tercijarnih limfoidnih struktura (TLS), a te strukture koreliraju s boljom prognozom [78]. Pokazalo se da limfociti B mogu predstaviti endogene tumorske antigene limfocitima CD4⁺ T i promijeniti njihov fenotip *in vitro*. Aktivirane stanice B koje infiltriraju tumor povezane su sa aktiviranim stanicama IFN γ +CD4⁺ T, a iscrpljene stanice B koje infiltriraju tumor povezane su sa imunosupresivnim stanicama Treg [79]. Ove studije sugeriraju da u TME, sposobnost stanica B koje infiltriraju tumor da utječu na CD4⁺ TIL-e predstavlja potencijalni novi terapijski cilj u imunoterapiji NSCLC [57].

1.9. Profiliranje molekularnih karakteristika tumora pluća skvamoznih stanica

Razvoj i primjena naprednih visokoprotočnih tehnologija u znanstvenim istraživanjima, poput sekvenciranja slijedeće generacije, doprinijeli su novim spoznajama o tumoru pluća. Primjerice, identificirane su molekularne karakteristike adenokarcinoma pluća koje su dovele do razvoja i primjene ciljane terapije u kliničkom okruženju [16].

Dosadašnja istraživanja molekularnih karakteristika tumora pluća skvamoznih stanica pokazala su da je ova vrsta tumora heterogenija nego na što ukazuje podjela na temelju morfoloških

karakteristika. Također, novija istraživanja ukazuju na potencijalno bitnu ulogu određenih gena i staničnih puteva u tumorigenezi ove vrste tumora pluća. Primjerice, komparativna istraživanja pokazala su da je 77 % promijenjeno eksprimiranih gena zajedničko preinvazivnom i invazivnom tumoru pluća skvamoznih stanica. Zajednički geni povezani su sa odgovorom na oštećenje DNA, metabolizmom DNA/RNA i upalom kao glavnim biološkim karakteristikama. Navedeno ukazuje da su promjene u regulaciji kontrolnih točki staničnog ciklusa, promjene u odgovoru na oštećenje DNA, kao i promjene u brojnim signalnim putevima, poput PI3K/Akt signalnog puta, događaji koji se događaju u ranim fazama skvamozne tumorigeneze. Stoga bi promijenjena razina ekspresije gena koji sudjeluju u regulaciji staničnog ciklusa, poput PTTG1 gena, mogla biti indikator skvamozne tumorigeneze [79]. Nadalje, povišena razina ekspresije ciklina, ciklin-ovisnih kinaza, gena iz E2F obitelji transkripcijskih faktora i MCM (engl. *Minichromosome maintenance proteins*) gena ukazuje da su promjene u regulaciji staničnog ciklusa ovisne o Notch, Wnt i ErbB signalnim putevima. Također, prekomjerna ekspresija poznatih onkogeni, poput PIK3CA, MYC i RAD51, može ukazivati na pojačanu agresivnost i metastatski potencijal tumora pluća skvamoznih stanica, u odnosu na adenokarcinom pluća [80]. Nadalje, pojačana ekspresija keratina, poput KRT5, KRT6, KRT14 i KRT17, kao i gena PERP, TRIM29, GPC1, CELSR2, S100A2 i TUBA1C je karakteristična za ovu vrstu tumora [81]. Smanjena ekspresija gena koji sudjeluju u regulaciji imunološkog odgovora i u odgovoru na transformirajući faktor rasta (TGF- β) također je uočena u tumoru pluća skvamoznih stanica [82]. Rezultati sekvenciranja eksoma ukazuju na veliku opterećenost tumora pluća skvamoznih stanica mutacijama, sa srednjom stopom od 8,1 mutacije po megabazi. Ovako visoka stopa mutacija nije zabilježena u drugim tumorima iz TCGA (engl. *The Cancer Genome Atlas*) projekta. Primjerice, mutacije u TP53 genu uočene su u 81 % testiranih uzoraka, a česte su i mutacije u NOTCH1, NOTCH 2, HLA-A, SOX2, TP63, ASCL4 i FOXP1 genima. CDKN2A gen, poznati tumor-supresor u tumoru pluća skvamoznih stanica, inaktiviran je u 72 % uzoraka [11].

Profil genske ekspresije korišten je u nekim istraživanjima tumora pluća skvamoznih stanica u svrhu identifikacije podtipova koji bi mogli služiti kao faktor prognoze preživljenja pacijenata ili kao prediktivni faktor odgovora na terapiju. Primjerice, pomoću analize diferencijalne ekspresije i analize obogaćivanja setovima gena, identificirana su četiri podtipa skvamoznog tumora pluća. Promijenjeno eksprimirani geni u identificiranom bazalnom podtipu imaju ulogu u adheziji stanice, geni u klasičnom podtipu povezuju se sa metabolizmom ksenobiotika, dok se geni

u sekretornom podtipu povezuju sa prekomjernom ekspresijom imunskih gena. Pacijenti čiji tumori pokazuju karakteristike primitivnog podtipa, kojeg karakteriziraju promijenjeno eksprimirani geni koji sudjeluju u regulaciji proliferacije stanica, imaju očekivanu najgoru prognozu preživljenja u odnosu na ostale podtipove [83]. Druga studija pokazala je da pripadnost klasičnom, primitivnom, sekretornom ili bazalnom podtipu tumora pluća skvamoznih stanica može biti prediktivni faktor odgovora na terapiju [84].

Iz navedenih istraživanja može se zaključiti da su tumori pluća skvamoznih stanica heterogena skupina tumora prema profilu genske ekspresije i opterećenju mutacijama. Bolje razumijevanje njihovih profila, moglo bi dovesti do personaliziranog pristupa liječenju, preciznijeg predviđanja odgovora na terapiju i u konačnici boljeg ishoda liječenja ove vrste tumora.

Osim potrebne bolje karakterizacije tumora pluća skvamoznih stanica na genskoj razini, proučavanjem genskih promjena ili promjena ekspresije mRNA, potrebna je i bolja karakterizacija na epigenetskoj razini kako bi bolje razumjeli tumorigenezu ovog tipa tumora pluća. miRNA su stabilne jednolančane RNA molekule veličine 20—25 nukleotida koje vezanjem na molekule mRNA mogu mijenjati ekspresiju ciljne molekule na post-transkripcijskoj razini. Biološka aktivnost miRNA povezana je sa regulacijom mnogih vitalnih staničnih procesa kao što su stanični rast, apoptoza i regulacija upale [85]. Profil miRNA u tkivima tumora pluća, primjenom naprednih tehnologija, poput sekvenciranja slijedeće generacije, slabo je istražen. Dosadašnja istraživanja ukazuju da je miR-338 značajno utišana, dok je njezin ciljni gen, NKX2 gen, povezan sa progresijom tumora pluća skvamoznih stanica [86]. Meta-analiza profila ekspresije miRNA u tumorima pluća pokazala je da su najčešće prekomjerno eksprimirane miR-210, miR-203 i miR-205, dok miR-126, miR-30a, miR-30d najčešće imaju negativno promijenjenu ekspresiju u odnosu na zdravo tkivo [86]. Neke studije ukazuju i na povezanost profila ekspresije miRNA sa stadijima i prognozom bolesti. Primjerice, Gallach S i suradnici pokazali su da je povećana ekspresija miR-21-5p i miR-188-5p u tumoru pluća ne-malih stanica povezana s lošijom prognozom [87]. Ma J i suradnici pronašli su povezanost povišene ekspresije miRNA-577, miRNA-183-3p i miRNA-944 sa uznapređovalim stadijima tumora pluća skvamoznih stanica. Također, pokazali su povezanost ekspresije miR-944 sa povećanom zahvaćenošću limfnih čvorova ukazujući da bi miRNA-944 mogla biti prediktor procesa metastaziranja [88]. Drugim metodama, poput RT-qPCR ili mikroeseja, prepoznala se potencijalna mogućnost korištenja određenih miRNA u dijagnozi tumora pluća iz tkivnih preparata [89] ili pomoću neinvazivne tekuće biopsije iz plazme i seruma [90].

Rezultati objavljenih radova koji su se bavili analizom miRNA, primjenom standardnih metoda molekularne medicine (RT-qPCR, mikroeseji), izolirane iz tkiva ili seruma/plazme upućuju na zaključak da miRNA posjeduje veliki potencijal u dijagnozi, predikciji odgovora na terapiju i prognozu bolesti.

Iako je bolja karakterizacija genskih i epigenetskih promjena bitna za bolje razumijevanje procesa tumorigeneze, karakterizacija profila ekspresije proteina je također izuzetno važna, obzirom da su to ključne izvršne molekule za funkciju stanice. Profiliranje tkiva spektrometrijom masa, u odnosu na druge manje osjetljive metode, omogućuje osjetljivu identifikaciju velikog broja proteina u uzorku. Primjerice, profiliranje proteoma primarnog tumorskog ksenografta ukazuje na karakterističnu ekspresiju 30 proteina u tumoru pluća skvamoznih stanica, poput povišene ekspresije fosfoglicerat-dehidrogenaze, citokeratina i ostalih komponenti epitelne barijere [91]. Također, prekomjerna ekspresija NF κ BIE, EDF1, HMGA1 i LPXN proteina smatra se specifičnom za tumor pluća skvamoznih stanica [92]. Proteomske analize također ukazuju da su modifikatori kromatina, poput LSD1 i EZH2, potencijalno podložni lijekovima u tumorima u kojima je SOX2 prekomjerno eksprimiran [93]. Nadalje, amplifikacija kromosomske regije 8p11-12 je česta u etiologiji tumora pluća skvamoznih stanica. U ovoj regiji nalazi se i gen FGFR1 kojeg se smatralo jednim od pokretača tumorigeneze ove vrste tumora [94]. Međutim, novija proteomska istraživanja [93,95] ukazuju na onkogenu ulogu NSD3 proteina (engl. *Histone-lysine N-methyltransferase NSD3*) u ovoj kromosomskoj regiji. Na mišjem modelu tumora pluća skvamoznih stanica pokazano je da ukidanje ekspresije NSD3, ali ne i FGFR1 proteina, utječe na smanjenje tumorske mase i dulje preživljenje miševa [95]. Stewart PA i suradnici napravili su integriranu analizu skvamoznog tumora pluća na razini genskih promjena, ekspresije mRNA i proteinskog profila. Rezultati tih analiza ukazali su na postojanje tri različita podtipa tumora pluća skvamoznih stanica koji su se međusobno razlikovali u stupnju infiltracije imunološkim stanicama, promjenama u oksidacijsko-redukcijskim putevima te aktiviranim signalnim putevima. Prema rezultatima ove studije može se zaključiti da bi pacijenti sa višim stupnjem infiltracije tumora imunološkim stanicama mogli imati koristi od imunoterapije [91].

Veliki izazov u liječenju malignih bolesti predstavlja činjenica da tumori vremenom razvijaju otpornost na sve oblike liječenja. Stoga su potrebna daljnja istraživanja koja će doprinijeti boljoj molekularnoj karakterizaciji svih tumora, pa i tumora pluća. Pretpostavka je da bi sveobuhvatna karakterizacija tumorskog mikrookoliša potencijalno dovela do identifikacije novih

prediktivnih ili prognostičkih biomarkera tumora pluća skvamoznih stanica ili identifikacije novih meta za razvoj lijekova. Također, karakterizacija novih tkiva tumora i sistematizacija novih podtipova prema molekularnim karakteristikama je od iznimnog značaja za bolje razumijevanje procesa tumorigeneze ove vrste tumora.

2. CILJEVI I HIPOTEZA ISTRAŽIVANJA

2. CILJEVI I HIPOTEZA ISTRAŽIVANJA

Opći cilj doktorske disertacije je priprema knjižnica za sekvenciranje mRNA i miRNA te priprema uzoraka za spektrometriju masa u svrhu istraživanja profila ekspresije mRNA/miRNA i proteina u tumoru pluća skvamoznih stanica. Rezultati genske ekspresije validirati će se usporedbom s podacima dobivenim analizom pacijenata iz TCGA kohorte. Razumijevanje molekularnih procesa u pozadini oboljenja omogućiti će probir novih specifičnih ekspresijskih biljega.

Specifični ciljevi ovog istraživanja su:

1. Analizom transkripcijskog profila odrediti gene koji su promijenjeno eksprimirani u tumorskom i zdravom tkivu, dodijeliti im funkcionalne anotacije i odrediti signalne puteve u kojima sudjeluju.
2. Metodom GSVA procijeniti stupanj infiltracije imunoloških stanica u tumorsko tkivo i odrediti povezanost stupnja infiltracije s preživljenjem pacijenata.
3. Primjenom metode spektrometrije masa, na razini pilot studije, odrediti ekspresijski profil proteina u tumorskom i zdravom tkivu i usporediti dobivene rezultate s rezultatima ekspresije na razini mRNA.
4. Analizom miRNA profila tumorskog i zdravog tkiva odrediti miRNA koje su promijenjeno eksprimirane, identificirati ciljne gene miRNA sa promijenjenom ekspresijom i usporediti dobivene rezultate s rezultatima ekspresije mRNA i proteina.

Hipoteza istraživanja: Integrirana analiza diferencijalne ekspresije na transkripcijskoj, post-transkripcijskoj i translacijskoj razini može identificirati ključne gene, genske produkte i molekule miRNA kao potencijalne terapeutske mete tumora pluća skvamoznih stanica

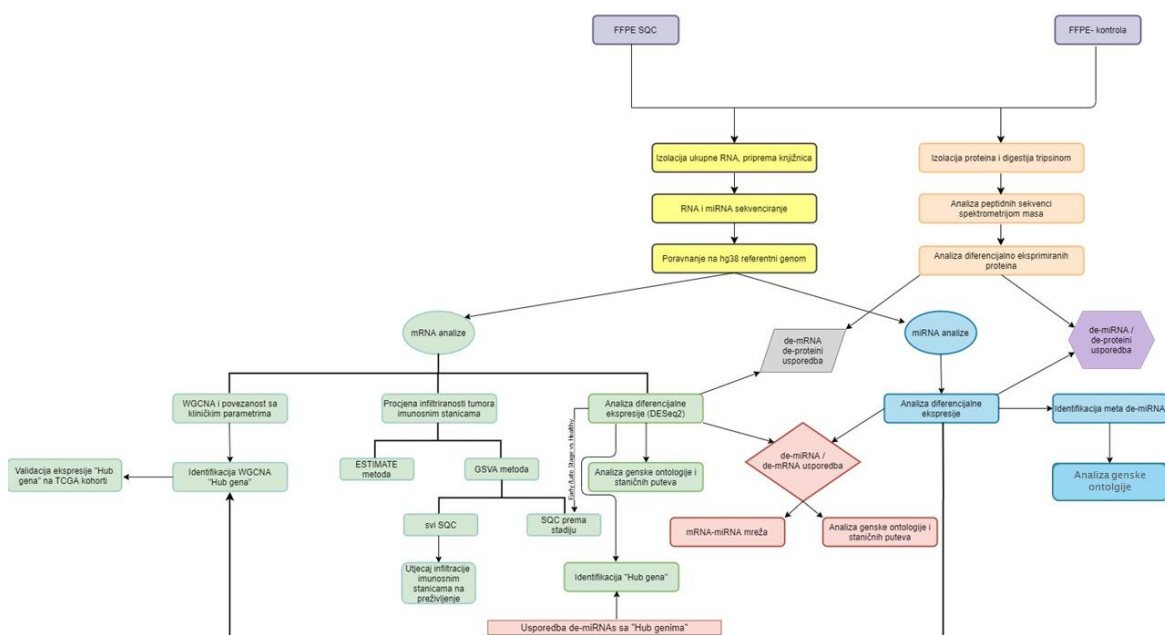
3. MATERIJALI I METODE

5 slika, 4 tablice

3. MATERIJALI I METODE

3.1. Plan istraživanja

U ovom istraživanju za potrebe profiliranja tumora pluća skvamoznih stanica na transkripcijskoj, post-transkripcijskoj i translacijskoj razini, koristili smo uzorke prikupljane u sklopu projekta IP-06-2016/1441 „Genetički i epigenetički biomarkeri urođene imunosti u KOPB-u i karcinomu pluća“. Prikupljanje uzoraka započelo je 2017. godine nakon odobrenja Etičkog povjerenstva KBC-a Zagreb. Istraživanje je provedeno u Laboratoriju za naprednu genomiku Instituta Ruđer Bošković, u suradnji sa Klinikom za plućne bolesti Jordanovac KBC-a Zagreb, Kliničkim zavodom za patologiju i citologiju KBC-a Zagreb i Zavodom za anatomiju, histologiju i patologiju Fakulteta za dentalnu medicinu i zdravstvo Sveučilišta u Osijeku. Iz prikupljenih uzoraka izolirana je ukupna RNA bez/sa miRNA frakcijom koja je korištena za pripremu knjižnica za sekvenciranje. Pripremljene knjižnice sekvencirane su na Illumina platformi u Macrogen-u, Koreja. Podaci dobiveni sekvenciranjem obrađeni su bioinformatičkim metodama. Priprema i analiza uzoraka spektrometrijom masa provedeno je u suradnji sa Laboratorijem za sistemsku biologiju i prijenos signala iz Njemačkog centra za istraživanje raka (DKFZ) u Heidelbergu, Njemačka. Rezultati spektrometrije masa također su obrađeni bioinformatičkim alatima. Konačno, preklapili smo rezultate transkriptomike i proteomike sa rezultatima ekspresije miRNA. Shema plana istraživanja prikazana je na Slici 6.



Slika 6. Shema plana istraživanja

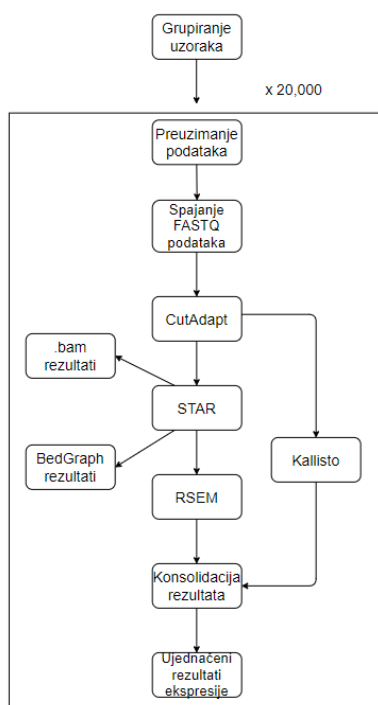
3.2. Ispitanici

Za potrebe ovog istraživanja prikupljeno je dvadeset i tri uzoraka tkiva ispitanika s dijagnozom tumora pluća skvamoznih stanica i četiri uzorka tkiva zdravih ispitanika. Ispitanici uključeni u ovo istraživanje potpisali su informirani pristanak i evidentirani su pod šifrom poznatom samo glavnom istraživaču čime je zajamčena anonimnost ispitanika. Uzorci ispitanika s tumorom pluća skvamoznih stanica prikupljeni su u Kliničkom bolničkom centru Zagreb, u Klinici za plućne bolesti Jordanovac. Tumorska tkiva prikupljana su tijekom dijagnostičke obrade ispitanika kirurškom resekcijom, iz razdoblja 2013.-2019. godine. Rezovi tumorskih tkiva obojani su hematoksilin-eozin kombinacijom boja te ih je pregledao patolog iz Zavoda za patologiju i citologiju KBC-a Zagreb. Tkiva tumora su uključena u istraživanje ukoliko su zadovoljila slijedeće kriterije: u uzorku je potvrđena prisutnost minimalno 60% tumorskih stanica, potvrđena je dijagnoza tumora pluća skvamoznih stanica i maksimalni postotak nekroze u tkivu ne prelazi 30%. Uzorci zdravih ispitanika prikupljeni su u Zavodu za anatomiju, histologiju i patologiju Fakulteta za dentalnu medicinu i zdravstvo, Sveučilište u Osijeku. Tkiva zdravih ispitanika prikupljena su tijekom patološke resekcije zdravih pojedinaca koji su umrli iznenadnom smrću. Zdrava tkiva pregledana su od strane kliničkog patologa i ustanovljeno je da ispitanici u trenutku smrti nisu obolijevali od kroničnih respiratornih bolesti, upalnih bolesti pluća niti su pronađene maligne promjene na plućnom tkivu ispitanika. Tumorska i zdrava tkiva su nakon uzimanja resekcijom fiksirana formalinom i uklopljena u parafin (engl. *Formalin fixed paraffin embedded, FFPE*) te pohranjena na +4°C do analize. Također su prikupljeni klinički podaci ispitanika sa dijagnozom tumora pluća skvamoznih stanica, uključujući informacije o pušenju, TNM stadiju, veličini tumora izraženoj u centimetrima, kao i podaci o preživljenju ispitanika.

Obzirom na relativno malenu eksperimentalnu kohortu ispitanika, u istraživanje smo uključili validacijsku kohortu sa većim brojem ispitanika prikupljenih iz mrežno dostupnih podataka. Mrežno dostupni rezultati sekvenciranja na mRNA razini porijeklom su iz različitih laboratorija i projekata te nije moguće direktno uspoređivanje razina ekspresije gena između projekata. Stoga smo skinuli set podataka genske ekspresije korištenih u sklopu *UCSC Toil RNA-seq recompute compendium* projekta [96] koristeći mrežno dostupnu *UCSC Xena* platformu. Cilj navedenog projekta bio je re-procesuiranje rezultata RNA sekvenciranja (*RNA-Seq*) za oko 20 000 uzoraka analiziranih u sklopu tri različita projekta, kako bi se kreirali konzistentni rezultati ekspresije gena kojima su računalno korigirani učinci pripadnosti različitim skupinama (engl.

computational batch effects). U sklopu navedenog projekta korišten je Toil program koji omogućava provedbu kompleksnih radnih zadataka u mrežnom oblaku ili računalnom okruženju koji zahtjeva sofisticirane računalne performanse za obradu petabitnih količina podataka (PB; 1 PB=1024 TB). Shema protokola korištenog za obradu očitanih sekvenci prikazana je na Slici 1. Tijekom projekta procesuirani su uzorci iz projekata TCGA (engl. *The Cancer Genom Atlas*), TARGET (engl. *Therapeutically Applicable Research to Generate Effective Treatments*) i GTEx (engl. *Genotype–Tissue Expression*).

TCGA projekt, započet 2006. godine suradnjom Nacionalnog instituta za rak (engl. *National Cancer Institute, NCI*) i Nacionalnog instituta za istraživanje humanog genoma (engl. *National Human Genome Research Institute, NHGRI*) jedan je od najvećih i najpoznatijih projekata genomske karakterizacije tumorskih tkiva. U sklopu projekta molekularno je okarakterizirano više od 20 000 uzoraka 33 vrste primarnih tumora, uključujući i tumor pluća skvamoznih stanica, i pripadnog ne-tumorskog tkiva. U sklopu GTEx projekta molekularno su okarakterizirana zdrava tkiva oko 1000 pojedinaca sa 54 različitih dijelova tijela, što ovaj projekt čini jednim od najvećih izvora podataka za ne-maligna tkiva.



Slika 7. Shema protokola koji je korišten u sklopu *UCSC Toil RNA-seq recompute compendium projekta*. Preuzeto i prilagođeno od Vivian J i suradnika [96].

Obzirom da su podaci genske ekspresije dobiveni iz TARGET projekta povezani sa pedijatrijskim malignim oboljenjima, ovaj set podataka uklonjen je iz validacijske kohorte. Iz TCGA seta podataka skinuti su rezultati genske ekspresije za tumor pluća skvamoznih stanica. Kako bi validacijska kohorta bila što sličnija eksperimentalnoj kohorti, navedeni set podataka filtriran je prema slijedećim kriterijima: pacijenti trebaju biti pripadnici bijele rase; ne biti pripadnici latinoameričkog etniciteta; u trenutku uzimanja uzorka tkiva pacijenti bi trebali biti bez propisane terapije i mjesto biopsije ili resekcije trebaju biti pluća. Pacijenti sa dijagnozom više primarnih tumora nisu uključeni u validacijsku kohortu. Ispitanici iz GTEx projekta filtrirani su prema mjestu uzimanja uzorka tkiva i u validacijsku kohortu su uključeni oni ispitanici kojima su kao mjesto uzimanja istaknuta pluća. Sukladno navedenim kriterijima u validacijsku kohortu je u konačnici uključeno ukupno 511 ispitanika, 223 ispitanika sa dijagnozom tumora pluća skvamoznog porijekla, porijeklom iz TCGA projekta, i 288 zdravih ispitanika, porijeklom iz GTEx projekta. Set podataka korišten za validaciju u ovom istraživanju dostupan je na mrežnoj stranici <https://xenabrowser.net/datapages/?cohort=TCGA%20TARGET%20GTEx&removeHub=https%3A%2F%2Fxcna.treehouse.gi.ucsc.edu%3A443>.

3.3. Izolacija RNA iz parafinskih kocki

U svrhu ispitivanja profila genske ekspresije i profila ekspresije miRNA u tkivima tumora pluća skvamoznih stanica, izolirana je ukupna RNA iz rezova tkiva uklopljenih u parafin (FFPE). Ukupna RNA bez miRNA frakcije izolirana je pomoću komercijalno dostupnog kompleta *RNAEasy* FFPE kit (Qiagen, SAD). Ukupna RNA sa miRNA frakcijom izolirana je pomoću komercijalno dostupnog kita *miRNeasy* FFPE kit (Qiagen,SAD), prema uputi proizvođača. Protokol za izolaciju ukupne RNA je isti za oba kita, sa razlikom da je veličina pora na kolonicama za izolaciju ukupne RNA sa miRNA frakcijom manja u odnosu na izolaciju ukupne RNA bez miRNA frakcije. Ukratko, parafinska kocka s uklopljenim tkivom tumora narezana je na ploške debljine 10 µm. Dva reza tkiva su potom prebačena u 1,5 ml Eppendorf tubice i dodano im je 1 ml ksilena u svrhu deparafinizacije tkiva. Rezovi tkiva su potom kratkotrajno (10 sekundi) vorteksirani i centrifugirani 2 minute. Nakon uklanjanja supernatanta, rezovi su dodatno isprani od ksilena sa 1 ml 96 postotnog etanola i centrifugirani u trajanju 2 minute. Supernatant je u potpunosti uklonjen te su uzorci inkubirani 10 minuta sa otvorenim tubicama na sobnoj temperaturi, kako bi eventualno zaostali etanol u potpunosti ispario. Potom su uzorci inkubirani 15 minuta na 56°C u 150 µl pufera

za lizu PKD i 10 μ l proteinaze K, nakon čega je odmah slijedila inkubacija na 80°C u trajanju 15 min. Ovaj korak inkubacije na 80°C u PKD puferu je ključan jer djelomično uklanja potencijalne modifikacije nukleinskih kiselina prouzrokovanih fiksacijom tkiva formaldehidom. Uzorci su potom ohlađeni na ledu 3 minute i centrifugirani 15 minuta. Supernatant je potom prebačen u drugu 1,5 ml tubicu, i tretiran sa 26 μ l DNaze u pripadajućem puferu tijekom 15 minuta na sobnoj temperaturi. Kako bi se RNA vezala za kolonicu, u tubicu sa tretiranim rezovima tkiva dodano je 320 μ l RBC pufera i 720 μ l 96 postotnog etanola. Potom se uzorak prebacio na kolonicu i višekratno isprao sa RPE puferom. Nakon posljednjeg ispiranja, RNA sa kolonice se isprala u 30 μ l vode bez RNA-za i spustila sa eluatom u tubicu. Koraci centrifugiranja bez upotrebe kolonica izvođeni su pri brzini rotacije od 20 000x g, dok su koraci centrifugiranja koji uključuju kolonice izvođeni pri rotaciji od 10 000x g.

3.4. Provjera kvalitete izolirane RNA

Za provjeru koncentracije izolirane ukupne RNA iz rezova tkiva, sa ili bez miRNA frakcije, korišten je Qubit 2.0 fluorometar i RNA BR (engl. *broad range*) kit (Thermo Fisher, SAD). Kitovi za mjerenje koncentracije na Qubit uređaju u sebi sadrže fluorescentnu boju koja se specifično veže za RNA i fluorescentni signal se emitira samo ako dođe do vezanja boje na RNA. Koncentracija se izračunava usporedbom relativne fluorescencije RNA u uzorku i krivulje dobivene iz relativne fluorescencije standarda poznate koncentracije. Uzorci su za potrebe mjerenja koncentracije razrijeđeni 200 puta u puferu sa bojom, inkubirani 2 minute na sobnoj temperaturi te im je izmjerena koncentracija. Čistoću izolirane RNA provjerili smo pomoću NanoDrop uređaja koji UV apsorbancijom određuje koncentraciju nukleinskih kiselina, pomoću omjera A260/280 nm, i čistoću uzorka, omjerom A260/230 nm. RNA čiji je omjer apsorbancije A260/230 iznosio oko 2 smatrana je čistom RNA, bez kontaminacije solima ili alkoholom.

Provjera integriteta RNA u uzorku ključan je korak u pripremi knjižnica za sekvenciranje. Kvalitetu izolirane ukupne RNA u uzorcima provjerili smo pomoću RNA Nano 6000 kita na BioAnalyzer 2100 uređaju (Agilent, SAD). Princip rada BioAnalyzer uređaja temelji se na razdvajanju i analizi RNA fragmenata kapilarnom elektroforezom. Često se prilikom pripreme knjižnica za sekvenciranje kvaliteta RNA u uzorku procjenjuje pomoću RIN vrijednosti (engl. *RNA Integrity Number*) izračunate pomoću omjera 28S i 18S rRNA fluorescentnog signala. Međutim, u FFPE uzorcima, gdje se očekuje veći postotak fragmentirane RNA, u odnosu na RNA dobivenu iz svježih tkiva ili stanica, RIN vrijednost nije najprimjereniji odabir za izražavanje kvalitete RNA.

Za ovaj tip uzoraka primjerenije je kvalitetu izolirane RNA u uzorcima procijeniti na temelju distribucije veličine fragmenata. DV200 vrijednost predstavlja postotak RNA fragmenata duljih od 200 nukleotida i pokazuje visoki stupanj korelacije sa prinosom pripremljene knjižnice. Stoga smo se odlučili kvalitetu RNA u uzorcima provjeriti pomoću DV200 vrijednosti. Ukratko, u gel za elektroforezu je dodana fluorescentna boja te se gel pročistio na kolonicama centrifugiranjem kroz 10 minuta, kako bi se uklonili mogući zaostaci većih komadića gela. Gel (9 μ l) je potom nanijet na BioAnalyzer čipove zajedno sa uzorcima (1 μ l) i markerom (5 μ l). Uzorak poznatih veličina fragmenata RNA (engl. *ladder*) (1 μ l) dodan je u posebnu jažicu. Uzorci su potom vorteksirani 1 minutu na 2400 rpm. Čipovi sa uzorcima odmah nakon pripreme su stavljeni na BioAnalyzer 2100 uređaj. Odabran je esej DV200 Nano RNA. Esaj je uspješan ukoliko se na elektroferogramu *laddera* mogu vidjeti 6 RNA pika i 1 pik markera, dobro razdvojenih. Sukladno preporukama kita za pripremu knjižnice za transkriptomsku analizu, samo uzorci sa DV200 vrijednostima većim od 30% su uključeni u analizu.

Ukupnu RNA sa izoliranom miRNA frakcijom, nakon RNA Nano 6000 eseja sa DV200 postavkom, dodatno smo provjerili na BioAnalyzer 2100 uređaju pomoću *Small RNA* kita (Agilent, SAD). Obzirom da je izolirana ukupna RNA porijeklom iz FFPE rezova, kako se povećava degradacija ukupne RNA, udio manjih fragmenata se povećava a duljih RNA molekula u uzorku smanjuje. Ako je RNA u uzorku jako degradirana, može doći do pogrešne procjene količine miRNA frakcije u uzorku. Zbog toga smo RNA sa miRNA frakcijom provjerili koristeći oba navedena eseja. *Small RNA* esej, u odnosu na *Nano RNA 6000* esej, omogućava bolje razdvajanje malih RNA fragmenata duljine 6 do 150 nukleotida. Također, ovaj esej može kvantificirati apsolutne količine miRNA u uzorku, izraženo u pg/ μ L, ili kao relativni udio (%) malih RNA u cijelom uzorku. Princip protokola za pripremu *Small RNA* eseja istovjetan je pripremi gore opisanog *RNA 6000 Nano* eseja.

3.5. Priprema knjižnica za sekvenciranje

Uzorci u kojima je pokazano da izolirana ukupna RNA nije previše fragmentirana, sukladno dobivenim DV200 vrijednostima, korišteni su za pripremu knjižnica za sekvenciranje. RNA bez miRNA frakcije korištena je za izradu knjižnice za sekvenciranje transkriptoma (engl. *RNA-Seq*) sa *TruSeq RNA Exome* kitom (Illumina, SAD). Ovaj kit bazira se na principu obogaćivanja kodirajućih regija pomoću PCR reakcije (engl. *Polymerase chain reaction*), i prikladniji je za

pripremu knjižnica iz FFPE uzoraka. Knjižnice koje koriste oligo-dT magnete za obogaćivanje mRNA frakcije manje su prikladne za FFPE uzorke jer oligo-dT magneti hvataju polyA repove mRNA a oni mogu biti fragmentirani u ovoj vrsti uzoraka, čime može doći do utjecanja na rezultate sekvencioniranja.

Knjižnice za sekvenciranje transkriptoma su pripremljene za 23 tumorska uzorka i 3 zdrave kontrole. Za pripremu knjižnice korišteno je 20-100 ng ukupne RNA, ovisno o DV200 vrijednostima (Tablica 5). Manje fragmentirani uzorci imaju veću DV200 vrijednost i za njih je korištena manja količina startne RNA (20 ng), dok je za uzorke sa jače fragmentiranom RNA i nižim DV200 vrijednostima (30-40%) korištena najveća količina startne RNA (100 ng), sukladno preporukama proizvođača.

Tablica 5. Prikaz količine ukupne RNA korištene za pripremu knjižnice *TruSeq RNA Exome kitom*

Kvaliteta RNA	DV200 vrijednost	Količina RNA korištene za pripremu knjižnice
Očuvana RNA	> 70 %	20 ng
Blago degradirana RNA	50-70 %	40 ng
Jače degradirana RNA	30-50 %	60-100 ng
Previše degradirana RNA	<30 %	Uzorci isključeni iz studije

Prema uputama proizvođača, prvi korak protokola (fragmentacija RNA) je preskočen, obzirom da je RNA izolirana iz FFPE uzoraka i pokazuje znakove degradacije. Izolirana ukupna RNA prvo je prevedena u jednolančanu cDNA pomoću nasumičnih heksamera i SuperScript II reverzne transkriptaze. Zatim slijedi uklanjanje RNA predloška (engl. *template*) i sinteza zamjenskog lanca. U ovom koraku se umjesto dTTP-a (deoksitimidin-trifosfat) za sintezu drugog lanca cDNA koristi uridin. Inkorporacija uridina sprječava sintezu drugog lanca tijekom amplifikacije knjižnice. Dvolančana cDNA se iz reakcijske smjese uklanja pomoću magnetskih perli. Obzirom da dvolančana cDNA ima tupe krajeve, dodaje joj se jedan adenzin na 3' kraju u svrhu sprječavanja ligacije fragmenata cDNA tijekom ligacije adaptera, a komplementarni timidin na 3' kraju adaptera omogućava ligaciju fragmenta cDNA sa adenozinom na adapter. Slijedeći korak je obogaćivanje DNA fragmenata pomoću PCR reakcije. U ovom koraku umnažaju se samo oni DNA fragmenti koji imaju ligirane adaptere, uz pomoć početnica koje prijanjaju na adaptere.

Nakon PCR reakcije, knjižnice su provjerene na BioAnalyzer 2100 uređaju uz pomoć *DNA High Sensitivity* eseja, prema ranije opisanom principu. Uspješno pripremljenim knjižnicama su potom dodane probe za hibridizaciju na kodirajuće regije. *TruSeq RNA Exome enrichment* kit uključuje više od 425,000 biotiniliranih proba, a svaka proba je konstruirana prema NCBI37/hg19 referentnom genomu, pokrivajući 98,3% referentne kodirajuće regije (Tablica 6).

Tablica 6. Specifikacije *TruSeq RNA Exome* eseja korištenog za pripremu mRNA knjižnica

Specifikacije TruSeq RNA Exome eseja	
Broj ciljnih gena	21 415
Broj ciljnih kodirajućih regija	214 126
Broj proba uključenih u esej	425 437
Postotak pokrivenosti referentnog transkriptoma	98.3 %

Magnetske perle sa vezanim streptavidinom su potom korištene za prikupljanje hibridiziranih probi u uzorku. Nakon elucije obogaćenih regija sa magnetskih perli, ponavlja se korak hibridizacije proba na kodirajuće regije. Nakon drugog koraka hibridizacije, pripremljene obogaćene knjižnice se pročišćavaju pomoću AMPure XP magnetskih perli, kako bi se uklonili zaostaci reagensa korištenih u prethodnom koraku. Nakon pročišćavanja, pripremljene knjižnice se amplificiraju pomoću PCR reakcije u 15 ciklusa te se ponovno pročišćavaju pomoću AMPure XP magnetskih perli. Veličina fragmenata pripremljenih knjižnica provjerena je pomoću BioAnalyzer 2100 uređaja i *DNA High Sensitivity* kita (Agilent, SAD). Knjižnice su potom razrijeđene i sekvencirane na HiSeq4000 platformi (Illumina, SAD).

Shema protokola za pripremu transkriptomске knjižnice prikazana je na Slici 8.



Slika 8. Shema protokola za pripremu transkriptomске knjižnice. Preuzeto i prilagođeno iz: Document # 1000000039582 v01, TruSeq RNA Exome Reference Guide, Illumina

RNA sa miRNA frakcijom korištena je za izradu knjižnice pomoću *TruSeq Small RNA* kita (Illumina, SAD). Na ukupno izoliranu RNA prvo se ligiraju adapteri, potom se RNA sa adapterima prevodi u cDNA. Fragmenti u uzorku se potom razdvajaju na temelju veličine na agaroznom gelu elektroforezom. Fragmenti koji veličinom odgovaraju miRNA frakciji se izrezuju iz gela i pročišćavaju na kolonicama. Knjižnice su pripremljene za 16 tumorskih uzoraka (za koje su pripremljene i mRNA knjižnice) i 4 zdrave kontrole. Za pripremu knjižnice korišten je 1 µg ukupne RNA sa miRNA frakcijom i *TruSeq Small RNA* kit (Illumina, SAD), prema uputi proizvođača. Ukratko, prvo su adapteri ligirani na 3' i 5' krajeve fragmenata ukupne RNA u uzorku. Nakon ligacije adaptera, ukupna RNA u uzorku prevedena je u cDNA pomoću enzima SuperScript II i umnožena PCR-om (15 ciklusa) pomoću početnica koje se vežu na krajeve adaptera. Pripremljene knjižnice nanese su potom na 6% poliakrilamidni gel, zajedno sa fragmentima poznate veličine u posebnim jažicama. Fragmenti cDNA u uzorcima migriraju na temelju veličine u polju struje od 145V tijekom 60 minuta. Potom je gel obojan etidij-bromidom (10µg/ml) u trajanju 2-3 minute i vizualiziran na UV transluminatoru. Fragment duljine 147 nukleotida sadrži miRNA frakciju, a

nastao je iz RNA fragmenata duljine 22 nukleotida kojem su dodani adapteri. Fragment duljine 157 nukleotida sadrži *piwi* RNA, neke miRNA vrste i ostale regulatorne male molekule, a nastao je iz RNA fragmenata duljine 30 nukleotida kojem su dodani adapteri. Pomoću skalpela iz poliakrilamidnog gela izrezan je dio ukupne RNA koji sadrži miRNA frakciju, odnosno izrezani su bendovi koji odgovaraju duljini fragmenta 140-160 nukleotida. DNA iz gela je potom eluirana u 300 μ l mqH₂O tijekom dva sata uz laganu rotaciju (250 rpm). Potom je eluat pročišćen na kolonicama veličine pora 5 μ m centrifugiranjem (1 minuta, 600 x g), kako bi se uklonili zaostaci gela. U svrhu povećavanja koncentracije gotovim knjižnicama, DNA je precipitirana dodavanjem natrijevog acetata (3M NaOAc) uz ispiranje peleta 100 % alkoholom etanolom. Gotove knjižnice eluirane su u 10 μ l 10 mM Tris-HCl (pH 8.5). Veličine pripremljenih knjižnica provjerene su pomoću BioAnalyzer uređaja i *DNA High Sensitivity* kita na ranije opisani način. Knjižnice su sekvencirane na HiSeq2500 platformi (Illumina, SAD).

Shema protokola za pripremu miRNA knjižnice prikazana je na Slici 9.



Slika 9. Shema protokola za pripremu miRNA knjižnice. Preuzeto i prilagođeno iz: Document # 15004197. *TruSeq® Small RNA Library Prep Reference Guide*, Illumina

3.6. Bioinformatička obrada podataka RNA-Seq analize

Rezultati sekvenciranja transkriptoma su neobrađene sekvence duljine 101 nukleotida. Prije daljnje analize potrebno je provjeriti kvalitetu neobrađenih sekvenci, kako bi se utvrdila kvaliteta samog sekvenciranja. Nakon provjera kvalitete sekvenci, slijedi poravnavanje sekvenci na referentni ljudski genom verzije GRCh38 i kvantifikacija genske ekspresije u uzorcima. U svrhu

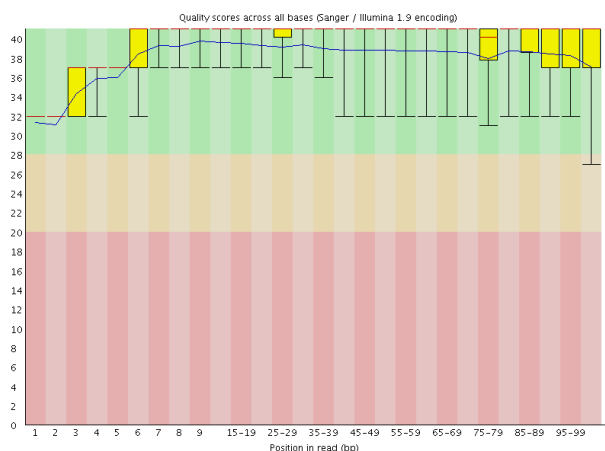
identifikacije gena čija je razina promijenjena u tumorskim uzorcima, u odnosu na zdrave kontrole, napravljena je analiza diferencijalne ekspresije gena. Uloga gena sa promijenjenom razinom ekspresije je potom provjerena pomoću analize genske ontologije i analize obogaćivanja setovima gena. U svrhu boljeg razumijevanja međusobnih odnosa gena sa promijenjenom razinom ekspresije u tumoru napravljena je analiza mreže proteinskih interakcija. Također, identificirani su visokoumreženi geni unutar konstruirane mreže te klasteri gena. Analiza ponderirane korelacijske mreže korištena je u svrhu identifikacije klastera i ključnih gena čija se ekspresija povezuje sa kliničkim parametrima pacijenata. Također, ispitan je utjecaj gena sa promijenjenom razinom ekspresije na sveukupno preživljenje pacijenata. Pomoću metoda ESTIMATE i GSVA procijenjen je stupanj infiltracije tumorskog tkiva sa imunološkim stanicama te je ispitan njegov utjecaj na preživljenje pacijenata. Rezultati diferencijalne ekspresije validirani su na većoj kohorti pacijenata iz TCGA projekta.

3.6.1. Provjera kvalitete neobrađenih sekvenci

U svrhu karakterizacije transkriptomskog profila tumora pluća skvamoznih stanica, sekvencirali smo transkriptom tumorskih i zdravih uzoraka. Pripremljene knjižnice sekvencirane su na željenoj dubini (100x) pomoću HiSeq4000 platforme (Illumina, SAD). Rezultati sekvenciranja dobiveni su u .fastq formatu koji se koristi za daljnju obradu podataka. Obzirom da očitane sekvence nisu savršene, za daljnje analize potrebno je procijeniti točnost očitavanja svake baze u sekvenci, odnosno kvalitetu očitanih baza. Postoji nekoliko softverskih opcija za analizu kvalitete očitavanja sekvenci, uključujući FastQC i FaQC. Nenormalnosti koje se uoče ovom analizom mogu se ukloniti obrezivanjem dijela sekvenci ili označiti za poseban tretman tijekom kasnijih postupaka.

U našim analizama kvaliteta očitanih sekvenci provjerena je pomoću mrežnog oblaka *BaseSpace Sequence Hub* i aplikacije *FASTQC* (Illumina, SAD). Aplikacija *FASTQC* sastoji se od nekoliko modula koji procjenjuju kvalitetu podataka i ukazuju na potencijalne probleme koje mogu utjecati na daljnje analize. Kvaliteta očitanih sekvenci se ispituje kako bi se osiguralo da su dodijeljene baze u sekvenci očitane sa visokom pouzdanošću, da udio GC baza u sekvencama odgovara očekivanoj distribuciji, da motivi kratkog niza (engl. *k-mers*) nisu previše zastupljeni i da je stopa dupliciranih sekvenci prihvatljiva. Svakom modulu aplikacija dodjeljuje prolaznu ocjenu ili upozorava na manja ili velika odstupanja od očekivanih rezultata.

Sažetak kvalitete sekvence po bazama prikazuje se pomoću *BoxWhisker* grafa za svaku bazu u sekvenci. Kvaliteta očitane sekvence izražava se u Phred vrijednostima kvalitete, koji označavaju vjerojatnost da je očitana baza u sekvenci zaista dobro očitana. Primjerice, Phred vrijednost kvalitete od 10 označava 90 % vjerojatnost da je očitana baza dobro očitana, odnosno predstavlja 10 % vjerojatnost da je uređaj za sekvenciranje bazu očitao krivo. Veća Phred vrijednost kvalitete označava manju vjerojatnost da je baza krivo očitana. Primjer sažetka kvalitete sekvence po bazama prikazan je na Slici 10. Na y-osi grafa označene su Phred vrijednosti kvalitete očitane sekvence. Crvena linija unutar žutog kvadrata označava medijan vrijednosti kvalitete očitane sekvence, žuti kvadrati predstavljaju među-kvartilni raspon (25-75 %) kvalitete, donje i gornje horizontalne linije (engl. *whiskers*) predstavljaju 10 i 90 % očitanih vrijednosti kvalitete, a plava linija predstavlja srednju vrijednost kvalitete očitane sekvence. Pozadina grafa dijeli y-os na veoma dobre vrijednosti kvalitete (označeno zelenom bojom), prihvatljive vrijednosti kvalitete (označeno narančastom bojom) i niske vrijednosti kvalitete očitane sekvence (označeno rozom bojom). Kvaliteta dodjeljivanja baza na većini platformi za sekvenciranje, uključujući HiSeq4000 platformu, se smanjuje kako proces sekvenciranja traje i nije neobično da vrijednost očitavanja kvalitete opada na kraju očitane sekvence.



Slika 10. Primjer rezultata modula Sažetak kvalitete sekvence po bazama

Modul Sažetak vrijednosti kvalitete po očitanoj sekvenci provjerava postoje li sekvence sa niskim vrijednostima kvalitete i koliko je njihov udio. Visoki udio sekvenci sa niskim vrijednostima kvalitete mogu ukazivati na sistemski problem sa očitavanjem sekvenci, a najčešće se pojavljuje radi problema sa očitavanjem sekvenci sa rubnog dijela protočne čelije u uređaju za sekvenciranje. Ukoliko je većini očitanih sekvenci srednja vrijednost kvalitete ispod 27, što odgovara pogrešci očitavanja od 0,2%, modul će izdati upozorenje. Ukoliko je srednja vrijednost

kvalitete za većinu očitanih sekvenci ispod 20, što odgovara pogrešci očitavanja od 1%, FASTQC će smatrati da ovaj modul nije prošao ispit kvalitete.

Modul Provjera očitanih sekvenci po bazama omogućava uvid u zastupljenost svake od četiri baze u sekvenci. Očekivano je da su u knjižnici sve baze u sekvenci zastupljene podjednako, što bi na grafu bilo prikazano kao paralelne linije za svaku bazu. Knjižnice koje su pripremljene pomoću nasumičnih heksamera mogu na početku očitane sekvence imati nasumične sljedove, najčešće u prvih desetak nukleotida. Ukoliko se na grafu primijeti prekomjerna zastupljenost neke od baza, to može ukazivati i na prekomjerno zastupljene sekvence u knjižnici koje su potencijalni kontaminanti knjižnice. Ukoliko je razlika u zastupljenosti A (adenozin), T (timidin), G (gvanin) ili C (citozin) baza veća od 20%, FASTQC će smatrati da ovaj modul nije prošao ispit kvalitete.

Modul Udio GC baza u očitanim sekvencama prikazuje zastupljenost GC baza u sekvenci i uspoređuje njihov udio sa normalnom distribucijom GC baza. U knjižnicama se očekuje relativno normalna distribucija GC baza, a centralni pik odgovara ukupnom udjelu GC baza u referentnom genomu. Neuobičajeni pik može ukazivati na kontaminaciju knjižnice. Ukoliko je zbroj devijacija od normalne distribucije veća u 30% očitanih sekvenci, FASTQC će smatrati da ovaj modul nije prošao ispit kvalitete.

Modul Udio N očitavanja u sekvenci ukazuje na postotak baza koje uređaj za sekvenciranje nije mogao pouzdano odrediti. Svakoj bazi u sekvenci koju uređaj ne može pouzdano odrediti, označiti će sa slovom N. Graf prikazuje postotak N očitavanja na svakoj poziciji. Nije neuobičajeno vidjeti mali postotak N očitavanja, pogotovo na krajevima sekvence. Ukoliko je udio N očitavanja veći od 20% na bilo kojoj poziciji u sekvenci, FASTQC će smatrati da ovaj modul nije prošao ispit kvalitete.

Modul Duplicirane sekvence: U knjižnicama će se većina sekvenci očitati samo jednom. Ukoliko je primijećena niska razina duplikacije to može ukazivati na visoku zastupljenost ciljne sekvence, ali visoka razina duplikacije je vjerojatnije posljedica obogaćivanja knjižnice primjerice PCR-om. Ovaj modul broji razinu duplikacije za sekvence u knjižnici i kreira graf koji pokazuje broj sekvenci sa različitim razinama duplikacije. U ovom modulu FASTQC analizira prvih 200 000 sekvenci kako bi se smanjila potrošnja memorije u analizi, ali to je dovoljno za procjenu razine dupliciranih sekvenci u cijeloj knjižnici. Na grafu se prikazuje do 10 duplikata, ako za sekvence postoji više duplikata, biti će prikazani kao da ih ima 10. Ako duplicirane sekvence predstavljaju

više od 50% ukupnih sekvenci u knjižnici, FASTQC će smatrati da ovaj modul nije prošao ispit kvalitete.

Modul Prekomjerno zastupljene sekvence: očekivano je da su sekvence u knjižnici raznolike i da nema sekvenci koje su prekomjerno zastupljene u odnosu na ostale sekvence u knjižnici. Prekomjerna zastupljenost neke sekvence može značiti da je sekvenca biološki značajna ili pak ukazivati na kontaminaciju knjižnice. Prekomjerno zastupljenim sekvencama se smatraju one sekvence čiji je udio veći od 0,1% ukupnog broja sekvenci. Modul će za svaku prekomjerno zastupljenu sekvencu tražiti poklapanje sa sekvencama najčešćih kontaminanta. Poklapanja moraju biti duljine minimalno 20 parova baza sa maksimalno jednim nepoklapanjem. Ukoliko je jedna sekvenca zastupljena više od 1% u odnosu na ostale sekvence, FASTQC će smatrati da ovaj modul nije prošao ispit kvalitete.

3.6.2. Poravnavanje neobrađenih sekvenci na referentni genom GRCh38

Nakon provjere kvalitete, neobrađene sekvence su poravnate na referentni genom čovjeka. Poravnavanje na referentni genom je napravljeno pomoću mrežnog oblaka *BaseSpace* i aplikacije *RNA-Seq Alignment* (Illumina, SAD). *RNA-Seq Alignment* aplikacija, osim poravnavanja očitanih sekvenci na referentni genom čovjeka, omogućava detekciju jako zastupljenih sekvenci (poput mitohondrijske ili ribosomalne RNA), omogućava detekciju genskih varijanti i detekciju fuzija gena. Za poravnavanje na referentni genom aplikacija koristi alat *Spliced Transcripts Alignment to a Reference* (STAR); za detekciju mitohondrijske (mtRNA) i ribosomalne RNA (rRNA) koristi alat *Bowtie*, za detekciju genskih varijanti aplikacija koristi alat *Strelka*; za detekciju fuzija gena koristi alat *Manta* dok za procjenu ekspresije gena i transkripata koristi alat *Salmon*. Prije samog poravnavanja očitanih sekvenci na referentni genom, aplikacija koristi alat *Bowtie* za poravnavanje očitanih sekvenci na poznate sekvence mtRNA i rRNA. Sekvence koje su identificirane kao mtRNA ili rRNA se filtriraju i ne poravnavaju na referentni genom. Kod analize *paired-end* knjižnica *Bowtie* filtrira sve sekvence koje su pronađene na barem jednom lancu očitane sekvence. Alat *STAR* poravnava preostale očitane sekvence na referentni genom u dva koraka: traženjem *seed* sekvence i njihovim grupiranjem. U prvom koraku *STAR* traga za maksimalnim prefiksom koji se može poravnati (eng. *Maximal Mappable Prefix*- MMP). S obzirom na očitane sekvence R, mjesto čitanja *i* i referentnu sekvencu genoma G, MMP (R, i, G) je definiran kao najduži podniz ($R_i, R_{i+1}, \dots, R_{i+MML-1}$) koji odgovara točno jednom ili više podnizova G, gdje je MML maksimalna

duljina koja se može poravnati. Primjerice, u očitanoj sekvenci koja sadrži jedno mjesto prekrajanja i u kojoj nema nepodudarnosti sa referentnim genomom, STAR algoritam prvo pronalazi MMP počevši od prve baze sekvence. Budući da se očitavanje u ovom primjeru sastoji od mjesta prekrajanja, očitana sekvenca se ne može poravnati uzastopno na genom, pa će tako prvo *seed* sekvenca biti poravnata na mjesto prekrajanja donatora. Zatim se MMP pretraga ponavlja za neporavnati dio očitavanja, koji će u ovom slučaju biti poravnati na mjesto spajanja akceptora. Mjesta prekrajanja detektiraju se u jednom prolazu poravnavanja. U drugoj fazi algoritma, STAR poravnava cijele sekvence čitanjem i spajanjem svih *seed* sekvenci koje su bile poravnate na genom u prvoj fazi [97].

RNA-Seq Alignment aplikacija osim mapiranja sekvenci na referentni genom, kvantificira gene i transkripte pomoću alata *Salmon*. Alat prilikom kvantifikacije transkripata uzima u obzir parametre koji su specifični za svaki uzorak, kao i potencijalne sistematske pogreške (engl. *bias*) koje su specifične za *RNA-Seq* metodu, poput razlika u dubini čitanja sekvenci ovisno o poziciji, *bias* na 5' i 3' krajevima očitane sekvence koji je specifičan za sekvencu, *bias* udjela GC baza u fragmentima, *bias* koji nastaje zbog korištenja specifičnih protokola za pripremu knjižnice i *bias* u distribuciji duljine fragmenata. Alat *Salmon* kvantificira transkripte i gene koristeći metodu kvazi-mapiranja [98].

U našoj analizi, za poravnavanje očitanih sekvenci na referentni genom koristili smo slijedeće postavke:

1. Odabrana je verzija genoma NCBI GRCh38 *Decoy*
2. Odabrana je opcija prepoznavanja sa kojeg lanca očitana sekvenca potječe (engl. *stranded*)
3. Odabrana je opcija prepoznavanja fuzija
4. Odabrana je opcija označavanja dupliciranih sekvenci
5. Odabrana je opcija uklanjanja sekvenci adaptera

Nakon poravnavanja očitanih sekvenci na referentni genom i kvantifikacije transkripata i gena, za analizu diferencijalne ekspresije gena koristili smo rezultate *Salmon* kvantifikacije gena dobivenih u obliku *quant.genes.sf* datoteka.

3.6.3. Analiza diferencijalne ekspresije gena

Nakon poravnavanja sekvenci na referentni genom i kvantifikacije gena, napravljena je analiza diferencijalne ekspresije gena. Svrha ove analize je identificirati one gene čija je razina u ekspresiji značajno promijenjena u uzorcima tumora pluća skvamoznih stanica u odnosu na zdrave kontrole. Pretpostavka je da geni čija je razina promijenjena u tumorskom tkivu utječu na specifični profil genske ekspresije koji doprinosi daljnjoj tumorigenezi ove bolesti.

Prilikom analize transkriptomskih podataka, najčešće se testira nulta hipoteza da logaritamska promjena ekspresije gena (engl. *Logarithmic fold change*, *log2FC*) između tretmana i kontrole iznosi točno nula, odnosno da do promjene ekspresije gena dolazi neovisno o tretmanu. Također, često je cilj analize diferencijalne ekspresije gena napraviti listu gena sa promijenjenom ekspresijom poredanih po prilagođenoj statističkoj značajnosti (engl. *adjusted p value*). Ipak, male promjene u ekspresiji gena, iako statistički jako značajne, možda nisu najzanimljivije za daljnje analize. Također, rangiranje gena prema promjeni ekspresije može biti komplicirano za gene sa malo očitanih sekvenci. Noviji alati za analizu diferencijalne ekspresije gena, poput *DESeq2* alata, u svojim algoritmima kompenziraju navedene potencijalne probleme što dobivene rezultate čini pouzdanijima. *DESeq2* je alat koji omogućava kvantitativnu analizu *RNA-Seq* podataka koristeći procjenitelje skupljanja za disperziju (engl. *Shrinkage estimators for dispersion*) i promjenu ekspresije (*log2FC*), uključujući rangiranje gena prema promjeni ekspresije i vizualizaciju, testiranje nulte hipoteze ispod i iznad granice (engl. *threshold*), kao i procjenu kvalitete i klasteriranje prekomjerno raspršenih očitanih sekvenci pomoću regulirane logaritamske transformacije (engl. *Regularized logarithm transformation*).

Za analizu je potrebno napraviti matricu (K) u kojoj su geni (i) poredani u redovima, a uzorci (j) u kolonama. Vrijednosti u matrici označavaju kvantificirane očitane sekvence (engl. *reads*) za svaki gen u uzorku. Za svaki gen *DESeq2* radi generalizirani linearni model (engl. *Generalized linear model*, GLM) na slijedeći način. Prvo se radi model po kojem matrica K_{ij} prati negativnu binomijalnu distribuciju (zvanu i *gamma-Poisson* distribucija) sa srednjom vrijednošću μ_{ij} i disperzijom α_i . Srednja vrijednost q_{ij} proporcionalna je koncentraciji cDNA fragmenata gena u uzorku, pomnoženo sa normalizirajućim faktorom s_{ij} , $\mu_{ij}=s_{ij}q_{ij}$. *DESeq2* algoritam računa GLM na logaritamskoj skali pomoću formule $\log_2 q_{ij} = \sum r_{jr} \beta_{ir}$, sa elementima matrice x_{jr} i koeficijentom β_{jr} . Prilikom usporedbe dvije skupine uzoraka, poput tumorskih i zdravih uzoraka, elementi dizajna matrice označuju kojoj skupini uzorak j pripada, a GLM model računa koeficijent koji ukazuje na

sveukupnu snagu ekspresije gena i $\log_2FC$ vrijednost između skupina. Varijabilnost između replikata je modelirana disperzijskim parametrom α_i koji opisuje varijance u broju očitanih sekvenci gena pomoću formule $\text{Var}K_{ij} = \mu_{ij} + \alpha_i \mu_{2ij}$. Točna procjena disperzijskog parametra je kritična za statističku snagu analize diferencijalne ekspresije gena, posebice u analizama koje uključuju mali broj uzoraka. *DESeq2* pretpostavlja da geni slične ekspresije imaju sličnu disperziju. Algoritam prvo procjenjuje disperziju za sve gene u uzorku pomoću metode maksimalne vjerodostojnosti (engl. *Maximum likelihood*), koja se bazira na podacima za svaki pojedinačni gen u analizi. Zatim se određuje prostorni parametar distribucije procijenjene disperzije i dodaje prilagođena zakrivljena krivulja. To omogućava točniju procjenu očekivane vrijednosti disperzije za gene određene ekspresije ali ne predstavlja devijaciju pojedinačnih gena sveukupnog trenda. Potom se procjene disperzije svih gena skupljaju (engl. *shrink*) prema vrijednostima predviđenim krivuljom kako bi se dobile konačne vrijednosti disperzije. *DESeq2* koristi empirički Bayes-ov pristup koji omogućava da snaga skupljanja ovisi o procjeni kako blizu krivulji su stvarne vrijednosti disperzije i stupnjevima slobode [99].

Najčešći problem u analizi rezultata sekvenciranja je velika varijacija procijenjene promjene ekspresije gena na logaritamskoj skali ($\log_2FC$) za one gene koji imaju mali broj očitanih sekvenci. *DESeq2* pristupa ovom problemu na način da skuplja vrijednosti promijenjene ekspresije gena prema nuli na način da je skupljanje jače kada je primjerice malo očitanih sekvenci, kada je velika disperzija ili postoji nekoliko stupnjeva slobode. Prvo se ponavlja GLM za dobivanje procjena maksimalne vjerodostojnosti za $\log_2FC$ i potom normalna distribucija centrirana oko nule prilagođava uočenoj distribuciji procijenjenih vrijednosti maksimalne vjerodostojnosti na razini svih gena. Potom se prilagođena distribucija koristi u ponovljenoj GLM analizi u kojoj se računaju konačne procijenjene vrijednosti $\log_2FC$ -a, zajedno sa standardnom pogreškom. Statistička značajnost računa se pomoću Wald testa, koji koristi z-vrijednosti, dobivene iz omjera $\log_2FC$ -a i standardne pogreške, i koja se uspoređuje sa normalnom distribucijom. Za prilagodbu p vrijednosti upotrebljava se postupak Benjamin-Hochberg [99].

U našim analizama, nakon poravnavanja sekvenci na referentni genom i kvantifikacije transkripata i gena, napravili smo analizu diferencijalne ekspresije gena, u svrhu identifikacije različite razine ekspresije gena u tumorskom tkivu u odnosu na zdrave kontrole. U ovu analizu uključena su 23 uzorka tumora pluća skvamoznih stanica i tri uzorka zdravih kontroli. Analiza je napravljena pomoću *BaseSapce* aplikacije *RNA-Seq Differential Expression* (Illumina, SAD) koja

za analizu diferencijalne ekspresije gena koristi alat *DESeq2*. Za *DESeq2* analizu, prvo su pojedinačne datoteke sa procijenjenom ekspresijom gena pomoću alata *Salmon* spojene u jednu ekspresijsku matricu gena, u kojoj su geni poredani u redove, a uzorci u kolone. Uzorci tumora pluća skvamoznih stanica označeni su kao testna skupina, a zdrave kontrole kao kontrolna skupina. Geni su okarakterizirani kao diferencijalno eksprimirani na temelju postavljene granične vrijednosti $|\log_2FC| \geq 2$ i prilagođene statističke značajnosti $p < 0,05$, prema metodi Benjamin-Hochberg [100]. Lista diferencijalno eksprimiranih gena koji zadovoljavaju navedene kriterije korištena je za daljnje analize, poput analize genske ontologije i analize obogaćivanja setovima gena. Također, iz aplikacije je preuzeta matrica sa normaliziranim ekspresijskim vrijednostima u obliku *rlog* transformacije, a koja je korištena u daljnjim analizama, poput analize ponderirane korelacijske mreže ili procjene infiltracije tumorskog tkiva sa imunološkim stanicama.

U svrhu validacije rezultata genske ekspresije na većoj kohorti, ponovili smo analizu diferencijalne ekspresije gena koristeći TCGA kohortu ($n=223$) i GTEx zdrave kontrole ($n=288$). Validacija genske ekspresije napravljena je u R programu, pomoću *DESeq2* paketa. Geni su okarakterizirani kao diferencijalno eksprimirani na temelju vrijednosti $|\log_2FC| \geq 1$ i prilagođene statističke značajnosti $p < 0,05$, prema metodi Benjamin-Hochberg. Dobivenu listu diferencijalno eksprimiranih gena u validacijskoj kohorti usporedili smo sa listom gena sa promijenjenom ekspresijom u eksperimentalnoj kohorti. Za usporedbu ekspresije gena između kohorti, uzeli smo u obzir sve gene čija je promjena u ekspresiji $|\log_2FC| \geq 1$. Sličnosti i razlike u ekspresiji gena između dvije kohorte prikazane su pomoću Venn dijagrama, koji je napravljen u R programu pomoću paketa *VennDiagram*.

3.6.4. Analiza genske ontologije i GSEA

Analiza genske ontologije postala je jedna od najčešće korištenih analiza u *RNA-Seq* protokolima kojima se pokušava objasniti funkcija gena u dobivenoj listi diferencijalno eksprimiranih gena. Sama riječ ontologija označava nešto što je poznato na temelju proučavanja, a pojam genska ontologija podrazumijeva znanje o genima i njihovoj funkciji. Projekt Genske ontologije (engl. *The Gene Ontology Resource*, GO; <http://geneontology.org>) pokrenut je 1998. godine sa ciljem stvaranja mrežne baze podataka o funkciji gena. GO baza podataka strukturirana je korištenjem formalne ontologije, kao usmjereni aciklički graf, u kojem svaki pojam ima definiranu vezu sa jednim ili više pojmova u istoj kategoriji. U GO bazi dostupno je oko 45 000

pojmov, zajedno sa 134 000 opisanih međudnosa pojmova. GO pokriva tri različita aspekta funkcije gena poredanih u tri kategorije: kategorija Molekularne funkcije opisuje aktivnost genskog produkta na molekularnoj razini; kategorija Stanične komponente opisuje lokaciju u odnosu na biološke strukture u kojoj genski produkt vrši svoju funkciju, i kategorija Biološki procesi koja opisuje procese koji koriste molekularnu funkciju gena. Svaki GO pojam unutar ontologije ima svoj naziv, jedinstveni alfanumerički identifikator, definiciju s navedenim citiranim izvorima, i ontologiju koja označava domenu kojoj pripada. GO baza podataka podržava računalnu pretragu i često se koristi u analizama obogaćenih pojmova, često zvanih i analiza staničnih puteva. U sklopu tih analiza potrebno je na mrežnoj stranici projekta unijeti listu gena sa promijenjenom ekspresijom. Uneseni geni se potom uspoređuju sa 20 000 ljudskih protein-kodirajućih gena. Računalnom obradom se sortiraju funkcije gena i stvaraju grupe prema funkciji. Potom se određuje statistička značajnost pripadnosti unesene liste gena određenoj grupi [101].

U sklopu naših analiza, zastupljenost (engl. *enrichment*) pojmova genske ontologije napravljena je pomoću mrežnog alata *Gorilla*, dostupnog na mrežnoj stranici <http://cbl-gorilla.cs.technion.ac.il/>. Obzirom da navedeni alat koristi statistički model koji traži obogaćene GO pojmove na vrhu rangirane liste, ovu analizu smo napravili zasebno za gene sa pozitivno promijenjenom ekspresijom i zasebno za gene sa negativno promijenjenom ekspresijom. Prvo su diferencijalno eksprimirani geni rangirani prema $\log_2FC$ vrijednostima, sa najviše pozitivno ili negativno promijenjenim genima na vrhu liste. Odabran je način rada rangirane liste sa graničnom statističkom vrijednošću (engl. *p-value threshold*) $p < 10^{-4}$. Analiza je napravljena za sve tri GO kategorije: kategoriju Biološki procesi, kategoriju Stanične komponente i kategoriju Molekularne funkcije. Prekomjerno zastupljeni GO pojmovi sa statističkom značajnošću korigirane p vrijednosti manje od 0,05, prema metodi Benjamini- Hochberg, smatrani su značajnima. Lista prekomjerno zastupljenih, statistički značajnih GO pojmova, korištena je za vizualizaciju rezultata u obliku stupičastih grafova u R programu (verzija 3.6.2.) pomoću paketa *ggplot2*.

Osim analize genske ontologije, napravljena je i analiza obogaćivanja setovima gena (engl. *Gene Set Enrichment Analysis*, GSEA). Cilj GSEA analize je utvrditi dali su pripadnici *a priori* definiranih setova gena S nasumično distribuirani u listi gena L, ili se većinski nalaze pri vrhu ili dnu liste gena. Ukoliko geni u listi gena nisu nasumično distribuirani, smatra se da su obogaćeni oni setovi gena u kojima se oni nalaze. Korištenje ove metode stoga omogućava lakše razumijevanje biološke uloge gena sa promijenjenom razinom ekspresije, pomoću definiranih

setova gena. Listu gena L za analizu priređuje ispitivač, najčešće na temelju rangiranih vrijednosti, poput *log2FC* vrijednosti.

GSEA analiza sastoji se od tri glavna koraka:

1. Računanje vrijednosti obogaćivanja (engl. *Enrichment score*, ES): ES predstavlja razinu prekomjerne zastupljenosti gena iz *a priori* definiranih setova gena S na dnu ili vrhu rangirane liste gena L. ES vrijednost se računa od vrha liste L nizvodno, prema dnu liste, na način da se tekući zbroj (engl. *running-sum*) povećava kada se gen nalazi u setu S, a smanjuje se kada gen iz liste L nije u setu S. Magnituda povećanja ovisi o korelaciji gena sa fenotipom. ES je maksimalno odstupanje od nule uočeno tijekom slučajnog gibanja (engl. *random walk*) a računa se ponderiranim (engl. *weighted*) Kolmogorov-Smirnovljevim testom.
2. Procjena razine značajnosti ES-a: Statistička značajnost ES-a procjenjuje se empirijskim permutacijskim testom, temeljenom na fenotipu, koji održava kompleksnu korelacijsku strukturu genske ekspresije. Točnije, permutiraju se fenotipske oznake, izračunavaju ES vrijednosti iz permutiranih podataka, što generira nultu distribuciju ES vrijednosti. Empirijska nominalna p vrijednost opaženih ES vrijednosti se potom izračunava iz nulte distribucije. Važno je naglasiti da permutacija temeljena na fenotipu održava gensku korelaciju i na taj način omogućava procjenu značajnosti koja je biološki smislenija nego što bi bila permutacija gena.
3. Prilagodba na višestruko testiranje hipoteza: Kada se iz baza podataka analiziraju svi setovi gena, potrebno je prilagoditi procijenjenu razinu značajnosti kako bi se uzelo u obzir testiranje više hipoteza. Prvo se normalizira ES za svaki set gena uzimajući u obzir veličinu seta i takav ES naziva se normalizirana vrijednost obogaćivanja (engl. *Normalized Enrichment Score*, NES). Potom se računa stopa lažnog otkrivanja (engl. *False Discovery Rate*, FDR) za svaku NES vrijednost. FDR označava procijenjenu vjerojatnost da set gena sa izračunatim NES vrijednostima predstavlja lažno pozitivan rezultat, a računa se usporedbom krajnjih granica opažene i nulte distribucije NES vrijednosti [102].

U našim analizama, za analizu obogaćivanja setovima gena koristili smo računalni alat GSEA (verzija 4.0.3.) i kolekciju setova gena C2 Svi kanonički putevi (engl. *All canonical pathways*) koju smo preuzeli iz MSigDB baze podataka (engl. *Molecular Signature Database*, MSigDB). Postavke za analizu postavljene su na slijedeći način: odabrana je analiza obogaćivanja setovima gena iz

ponderirane rangirane liste gena sa 1000 permutacija; odabrani su setovi gena koji imaju više od 10 (minimalno 10 gena) i manje od 500 gena u setu (maksimalno 500 gena). Uzorcima su u zasebnom dokumentu dodijeljene fenotipske oznake tumorskog ili zdravog tkiva. Geni sa promijenjenom ekspresijom su poredani prema *log2FC* vrijednostima, od najveće prema najmanjoj vrijednosti i uneseni u GSEA alat, zajedno sa fenotipskim oznakama. Rezultati analize su vizualizirani u R programu pomoću *ggplot2* paketa, u obliku točkastog grafa. Stanični putevi su smatrani obogaćenima ako im je korigirana p vrijednost $FDR < 0,25$.

3.6.5. Analiza mreže proteinskih interakcija, analiza modula i identifikacija ključnih gena

Interakcije između proteina su važne za mnoge procese u stanici. Mreža proteinskih interakcija (engl. *protein-protein interaction network*, PPI) pomaže u razumijevanju fizičkih i funkcionalnih interakcija proteina u stanici. PPI mreže su uobičajeno predstavljene u obliku grafa u kojima su geni/proteini prikazani kao čvorišta (engl. *nodes*), a povezani su linijama (engl. *edges*) koje predstavljaju interakcije između proteina. Stupanj povezanosti (engl. *degree*) označava broj interakcija jednog čvorišta unutar PPI mreže [103]. Unutar PPI mreže moguće je identificirati klastere gena. Svrha klasteriranja PPI mreže je pronalazak skupina proteina- klastera, koji su u interakciji i sudjeluju u istim biološkim procesima ili obavljaju specifičnu biološku funkciju.

Kako bi istražili potencijalne interakcije između gena sa promijenjenom ekspresijom, napravili smo analizu mreže proteinskih interakcija. Za analizu smo koristili mrežni alat *Search Tool for the Retrieval of Interaction Genes* (STRING, verzija 11,0), dostupan na stranici <http://string-db.org>. Kako bi napravili mrežni model, u STRING alat smo unijeli listu diferencijalno eksprimiranih gena i za vrijednost pouzdanosti interakcije (engl. *confidence score*) odabrali smo vrijednost $\geq 0,4$. Vrijednost pouzdanosti interakcije označava vjerojatnost da interakcija između dva proteina zaista postoji. STRING alat vjerojatnost postojanja interakcije iskazuje na skali 0 do 1, gdje 0 predstavlja najmanju mogućnost za postojanje interakcije, a 1 najveću pouzdanost u postojanje interakcije između proteina. Mrežni model, napravljen u STRING alatu, je potom vizualiziran pomoću *Cytoscape* programa (verzija 3,8,2). Nakon što je PPI mreža vizualizirana, analizirana su njezina svojstva: postojanje visoko umreženih proteina (engl. *hub genes*) i postojanje visoko umreženih regija unutar PPI mreže, odnosno klastera.

Za identifikaciju klastera, regija visoko umreženih proteina unutar PPI mreže, koristili smo aplikaciju MCODE (skraćeno iz engl. *Molecular Complex Detection*) u Cytoscape programu. Granične vrijednosti za pronalazak klastera postavili smo na slijedeći način:

- Stupanj povezanosti čvorišta=2
- Vrijednost čvorišta (engl. *node score*) =0,02
- K-core vrijednost=2
- Maksimalna dubina=100
- MCODE vrijednost $\geq$ 5

Za identifikaciju visoko umreženih proteina unutar PPI mreže (engl. *hub genes*) koristili smo 4 algoritma: Maksimalna centralnost grupe (engl. *Maximal Clique Centrality*, MCC), Stupanj povezanosti čvorišta, Linijski filtrirana komponenta (engl. *Edge Percolated Component*, EPC) i Maksimalna komponenta susjeda (engl. *Maximum Neighborhood Component*, MNC), pomoću aplikacije *CytoHubba* u *Cytoscape* programu. Izračunate vrijednosti iz navedenih algoritama su potom preklopljene, u svrhu identifikacije top 50 proteina sa najvišim izračunatim vrijednostima. One proteine koji se u svim algoritmima svrstavaju u top 50 proteina, prema izračunatim vrijednostima algoritama, smatrali smo visoko umreženim proteinima.

Nakon identifikacije visoko umreženih gena/proteina i identifikacije klastera, htjeli smo vidjeti u kojim staničnim putovima sudjeluju. U tu svrhu napravili smo analizu obogaćenih staničnih puteva uz pomoć mrežnog alata *Enrichr* dostupnog na stranici <https://maayanlab.cloud/Enrichr/>. U alat smo unijeli gene iz visoko umreženih regija PPI mreže (klastera) ili visoko umrežene proteine (hub gene) i za analizu obogaćivanja setovima gena odabrali set gena *MSigDB Hallmark 2020*. Rezultati obogaćivanja vizualizirani su stupičastim grafovima pomoću navedenog alata.

3.6.6. Analiza ponderirane korelacijske mreže (engl. *weighted gene co-expression network analysis*, WGCNA)

Kako bi identificirali module (klastera), koji sadrže međusobno jako korelirane gene, i korelirali ih sa kliničkim parametrima pacijenata oboljelih od tumora pluća skvamoznih stanica, koristili smo ponderiranu mrežnu analizu genske ko-ekspresije. Analiza ponderirane korelacijske mreže, WGCNA, je metoda sistemske biologije za opisivanje korelacijskih odnosa gena u uzorcima. WGCNA se često koristi za pronalazak modula visoko koreliranih gena, sažimanje klastera pomoću *eigengena* modula, za usporedbu modula međusobno ili za usporedbu modula sa

kliničkim parametrima. Za ovu analizu koristili smo matricu sa regularno normaliziranim očitanim sekvencama (*rlog* normalizacija) i WGCNA paket u R programu. Definicija osnovnih pojmova povezanih sa WGCNA analizom prikazana je u Tablici 7.

Tablica 7. Definicija osnovnih pojmova korištenih u WGCNA analizi

Pojam	Definicija
Korelacijska mreža	Korelacijska mreža je neusmjerena ponderirana mreža gena. Čvorišta (engl. <i>nodes</i>) unutar mreže odgovaraju profilu ekspresije gena, a linije između gena (engl. <i>edges</i>) odgovaraju korelaciji ekspresije dva gena. Potenciranjem apsolutne vrijednosti korelacije na snagu $\beta \geq 1$ (engl. <i>Soft-threshold</i>), prilikom konstrukcije ponderirane korelacijske mreže naglašavaju se jače korelacije ekspresije gena na račun slabijih korelacija ekspresije gena. Matrica blizine (engl. <i>Adjacency</i> , a) kod neusmjerene ponderirane korelacijske mreže računa se pomoću formule $a_{ij} = \text{cor}(x_i, x_j) ^\beta$. Kod usmjerene ponderirane korelacijske mreže matrica blizine računa se pomoću formule $a_{ij} = (1 + \text{cor}(x_i, x_j)) / 2 ^\beta$.
Modul	Moduli su klasteri visoko umreženih gena. Kod neusmjerene ponderirane korelacijske mreže, moduli odgovaraju klasterima gena koji pokazuju visoku apsolutnu korelaciju. Kod usmjerene ponderirane korelacijske mreže, moduli odgovaraju klasterima pozitivno koreliranih gena.
Povezanost	Za svaki gen, povezanost (engl. <i>connectivity</i> ili <i>degree</i>) se definira kao zbroj jačina povezanosti sa ostalim genima u mreži. U korelacijskoj mreži povezanost označava razinu korelacije jednog gena sa ostalim genima u mreži.
Povezanost unutar modula k_{IM}	Povezanost unutar modula mjeri razinu povezanosti, ili ko-ekspresiju gena sa ostalim genima određenog modula. Povezanost unutar modula može se interpretirati kao mjera pripadnosti određenom modulu.
Eigengen vrijednost E modula	Eigengen vrijednost E modula definira se kao glavna komponenta određenog modula. Može se smatrati reprezentativnim primjerom profila genske ekspresije u modulu.
Značajnost Eigengena	Ukoliko su dostupni klinički podaci, poput primjerice statusa pušenja, mogu se korelirati sa vrijednostima eigengena određenog modula. Koeficijent korelacije predstavlja značajnost eigengena.
Pripadnost modulu k_{ME}	Za svaki gen, pripadnost određenom modulu računa se korelacijom ekspresije gena sa eigengenom modula. Primjerice, $MM^{blue}(i) = K_{cor,i}^{blue} = \text{cor}(x_i, E^{blue})$ mjeri korelaciju gena i sa eigengenom modula <i>blue</i> . $MM^{blue}(i)$ mjeri pripadnost gena i <i>blue</i> modulu. Ako je $MM^{blue}(i)$ vrijednost blizu nuli, gen i ne pripada modulu <i>blue</i> . Ako je $MM^{blue}(i)$ vrijednost blizu 1 ili -1, gen i je jako povezan sa genima unutar modula <i>blue</i> . Predznak pripadnosti modulu označava pozitivnu ili negativnu korelaciju gena sa eigengenom modula.

Tablica 7. Nastavak

Visoko umreženi gen	Gen koji ima puno interakcija sa ostalim genima. Geni unutar modula mogu imati visoku umreženost.
Značajnost gena GS	Značajnost gena vrijednosti 0 ukazuje da gen nije značajan za biološko pitanje od interesa. Veća vrijednost GS-a podrazumijeva veću biološku značajnost.
Značajnost modula MS	Značajnost modula je srednja vrijednost GS-a svih gena u određenom modulu. Kada je značajnost gena definirana korelacijom profila genske ekspresije i kliničkih parametara, MS je često povezan sa korelacijom eigengena određenog modula i kliničkog parametra.

Funkcije u WGCNA paketu mogu se podijeliti u nekoliko kategorija: konstrukcija mreže; detekcija modula unutar mreže; odabir modula i gena; računanje topoloških svojstva mreže i vizualizacija rezultata [104].

Prvi korak u analizi je konstrukcija mreže. Za konstrukciju mreže koristili smo matricu sa regularno normaliziranim očitanim sekvencama (*rlog* normalizacija), iz koje se računa sličnost profila genske ekspresije uzoraka u obliku matrice sličnosti (engl. *similarity matrix*). Konstrukcija ponderirane mreže gena podrazumijeva odabir snage granične vrijednosti β (engl. *soft threshold power*) koja se koristi za potenciranje matrice sličnosti. Za konstrukciju mreže odabrali smo snagu $\beta = 18$ unutar funkcije *pickSoftThreshold*, jer je na toj vrijednosti $R^2 > 0,8$. Matrica sličnosti se potenciranjem na vrijednost β transformira u matricu blizine (engl. *adjacency*) koja predstavlja snagu povezanosti čvorišta u mreži. Matrica blizine kreirana je pomoću funkcije *adjacency* sa odabranim postavkama *Pearson* korelacije i opcijom usmjerene mreže (engl. *signed*) koja zadržava informaciju o smjeru ekspresije gena. Potom je matrica blizine transformirana u matricu topološke mjere preklapanja (engl. *Topological Overlap Matrix*, TOM) koju smo koristili za računanje matrice različitosti (engl. *dissimilarity*) pomoću 1-TOM formule. Za identifikaciju modula koji sadrže gene sa sličnom razinom ekspresije, koristili smo matricu različitosti i funkciju *geneTree* i funkciju *cutreeDynamic* u kojoj smo minimalnu veličinu modula ograničili na 30 gena a *DeepSplit* parametar postavili na vrijednost 0. Identificirani moduli koji sa drugim modulima koreliraju $> 0,75$ su spojeni, na temelju razlike između *eigengene* vrijednosti dobivenih modula. Potom su ispitane korelacije identificiranih modula i kliničkih parametara koristeći korelaciju *eigengene* vrijednosti identificiranih modula i kliničkih podataka. Također, za svaki gen u identificiranim modulima izračunate su vrijednosti značajnosti gena (engl. *Gene significance*, GS) i vrijednost pripadnosti modulu (engl. *Module membership*, MM). Vrijednost značajnosti gena korelira svaki gen u modulu sa kliničkim

parametrom od interesa, dok vrijednost pripadnosti modulu korelira *eigengene* modula sa profilom ekspresije gena. Na temelju korelacije modula i kliničkih parametara, odabrani su zanimljivi moduli čiji su geni analizirani mrežnom interakcijom proteina (PPI analiza). STRING baza podataka korištena je za modeliranje mreže proteinskih interakcija, a vizualizacija je napravljena u *Cytoscape* programu. Analizirane su karakteristike PPI mreže pomoću funkcije *Analyze Network* u *Cytoscape* programu. Ako gen ima visoku vrijednost stupnja povezanosti čvorišta (engl. *degree*), može imati bitnu ulogu u identificiranom modulu. Za identifikaciju visoko umreženih gena u klinički interesantnim modulima koristili smo granične vrijednosti: stupanj umreženosti čvorišta >5, vrijednost pripadnosti gena modulu (MM) >0,8 i značajnost gena unutar modula (GS) >0,2. Osim identifikacije visoko umreženih gena, funkcija gena u klinički interesantnim modulima analizirana je pomoću analize ontologije gena. Za ovu analizu korištena je mrežna baza podataka *Panther* dostupna na stranici <http://pantherdb.org/tools>, a kao anotacijski set podataka odabrana je *GO biological process complete*. Za računanje prekomjerne zastupljenosti GO pojmova korišten je Fisherov egzaktni test, a za računanje statističke značajnosti stopa lažnog otkrića FDR.

3.6.7. Analiza sveukupnog preživljenja

Kako bi provjerili dali postoji korelacija između sveukupnog preživljenja pacijenata sa razinom ekspresije diferencijalno eksprimiranih gena (*DEG*) ili korelacija preživljenja sa identificiranim visoko umreženim genima, napravili smo analizu sveukupnog preživljenja. Obzirom da je eksperimentalna kohorta relativno malena, za ovu analizu koristili smo validacijsku TCGA LUSC kohortu. Korelacija između sveukupnog preživljenja TCGA LUSC pacijenata i *DEG*-ova, ispitana je za one gene čija je promjena u razini ekspresije potvrđena na eksperimentalnoj i validacijskoj kohorti. Analiza je napravljena pomoću mrežnog alata *Xena* dostupnog na stranici <https://xenabrowser.net/>. U alatu je odabrana kohorta TCGA LUSC (n=626 uzoraka), te su zadržani samo oni uzorci koji su uključeni u našu validacijsku kohortu. Kao varijable za analizu preživljenja odabrani su: za vrstu podataka odabrani su genomski podaci; za set podataka odabrana je genska ekspresija te su upisani geni čiju smo korelaciju sa preživljenjem ispitivali. Za svaki gen od interesa, pacijenti su uvršteni u jednu od dvije grupe, na temelju razine ekspresije (jaka ili slaba ekspresija gena). Kao granična vrijednost za grupiranje (engl. *threshold*) u grupe, korištena je srednja vrijednost ekspresije gena. Za procjenu razlike u preživljenju između

grupa korištene su Kaplan-Meier krivulje preživljenja, a statistička razlika između krivulja procijenjena je pomoću log-rank testa značajnosti. P vrijednosti manje od 0,05 smatrane su statistički značajnima.

3.6.8. Procjena razine infiltriranosti tumora sa imunološkim stanicama i povezanost sa sveukupnim preživljenjem

U tkivu solidnih tumora se, osim tumorskih stanica, nalaze epitelne stanice, stromalne stanice, imunološke stanice i vaskularne stanice. Stromalne stanice mogu doprinijeti rastu tumora i progresiji bolesti [58]. Infiltrirajuće imunološke stanice mogu, ovisno o kontekstu, imati anti-tumorsko djelovanje, poput anti-tumorskog djelovanja infiltrirajućih T limfocita u tumorskom tkivu, ili mogu podupirati rast tumora i proces metastaziranja [105].

Kako bi procijenili razinu infiltriranosti tumora sa imunosnim stanicama u pojedinačnim uzorcima eksperimentalne kohorte, koristili smo paket *ESTIMATE* (engl. *Estimation of Stromal and Immune cells in Malignant Tumors using Expression data, ESTIMATE*) u R programu. *ESTIMATE* analiza se temelji na analizi obogaćivanja setovima gena na razini uzorka (engl. *single-sample GSEA, ssGSEA*) [106]. Obogaćivanje se radi za dva seta gena: set gena čija se ekspresija povezuje sa stromom i set gena specifičnih za imunološke stanice. Rezultat *ESTIMATE* analize iskazuje se u obliku Stromalnih vrijednosti (engl. *Stromal score, SS*), Imunosnih vrijednosti (engl. *Immune score, IS*) te *ESTIMATE* vrijednosti koja predstavlja zbroj *SS* i *IS* vrijednosti. Za računanje je korištena matrica ekspresijskih vrijednosti koje su normalizirane (rlog vrijednosti). Za vizualizaciju toplinske mape *SS*, *IS* i *ESTIMATE* vrijednosti za sve uzorke korištena je funkcija *heatmap.2* unutar *gplot* paketa u R programu.

U svrhu detaljnijeg proučavanja profila tumorskog mikrookoliša, koristili smo još jednu metodu obogaćivanja setovima gena na razini pojedinog uzorka, *GSVA* (engl. *Gene Set Variation Analysis, GSVA*). *GSVA* izračunava *GSVA* vrijednost obogaćivanja (engl. *GSVA enrichment score*), u rasponu -1 do 1, koji predstavlja sveukupnu ekspresiju ispitivanog skupa gena u odnosu na druge gene u pojedinom uzorku. Za računanje *GSVA* vrijednosti obogaćivanja korištena je matrica normalizirane ekspresije gena (rlog) u tumorskim uzorcima i 16 setova gena čija se ekspresija povezuje sa specifičnim imunološkim stanicama. Također, u analizu smo uključili i jedan dodatni set gena nazvan Citotoksične stanice a predstavljaju ga geni specifični za aktivirane T limfocite, gama delta T stanice i NK stanice zajedno, u svrhu procjene ukupne protutumorske aktivnosti

imunoloških infiltrata. U Tablici 8. navedeni su setovi gena korišteni u GSVA analizi, zajedno sa genima za koje se radi obogaćivanje.

Tablica 8. Setovi gena korišteni u GSVA analizi

Imunološke stanice	Geni uključeni u analizu GSVA	Ref.
B limfociti	ABCB4, BACH2, BCL11A, BLK, BLNK, CCR9, CD19, COCH, CR2, DTNB, FCRL2, GLDC, GNG7, HLA_DOB, HLA_DQA1, MEF2C, MICAL3, MS4A1, OSBPL10, PNOC, QRSL1, SCN3A, SLC15A2, SPIB, TCL1A, TNFRSF17	[107]
Mastociti	ABCC4, ADCYAP1, CALB2, CEACAM8, CMA1, CPA3, CTSG, ELANE, GATA2, HDC, HPGD, HPGDS, KIT, LINC01140, MAOB, MLPH, MPO, MS4A2, NR0B1, PPM1H, PRG2, PTGS1, SCG2, SIGLEC6, SLC18A2, SLC24A3, TAL1, TPSAB1, TPSB2, VWA5A	[107]
Makrofazi	APOE, ATG7, BCAT1, CCL7, CD163, CD68, CD84, CHI3L1, CHIT1, COL8A2, COLEC12, CTSK, CXCL5, CYBB, DNASE2B, EMP1, FDX1, FN1, GM2A, GPC4, ANOS1, ME1, MS4A4A, MSR1, PCOLCE2, PTGDS, RAI14, SCARB2, SCG5, SGMS1, SULT1C2	[107]
Nezrele dendritične stanice	ABCG2, BLVRB, CARD9, CD1A, CD1B, CD1C, CD1E, CH25H, CLEC10A, CSF1R, CTNS, F13A1, FABP4, FZD2, GUCA1A, HS3ST2, LMAN2L, NUDT9, PDXK, PPARG, PREP, RAP1GAP, SLC26A6, SLC7A8, SYT17, TACSTD2, DCSTAMP, VASH1	[107]
Aktivirane dendritične stanice	CCL1, EBI3, IDO1, LAMP3, OAS3	[107]
Neutrofili	ALPL, BST1, CD93, CEACAM3, CREB5, CRISPLD2, CSF3R, CYP4F3, DYSE, FCAR, FCGR3B, FPR1, FPR2, G0S2, HIST1H2BC, HPSE, CXCR1, CXCR2, KCNJ15, LILRB2, MGAM, MME, PDE4B, S100A12, SIGLEC5, SLC22A4, SLC25A37, TECPR2, TNFRSF10C, VNN3	[107]
NK CD56dim stanice	S1PR5, GTF3C1, GZMB, IL21R, KIR2DL3, KIR3DL1, KIR3DL2, KIR3DL3, KIR3DS1, SPON2, PMPA1	[107]
NK 56bright stanice	DUSP4, FOXJ1, LPCAT4, MADD, MPPED1, TRAPPC9, PLA2G6, RRAD, XCL1	[107]
Efektorne memorijske stanice	AKT3, SND1_IT1, CCR2, DDX17, EWSR1, FLI1, GDDP5, LTK, MEFV, NFATC4, PRKY, TBC1D5, TBCD, TRA, EZR	[107]
Centralne memorijske stanice	AQP3, ATF7IP, ATM, CASP8, CDC14A, CEP68, CLUAP1, CREBZF, CYLD, DOCK9, FAM153B, FOXP1, FYB, HNRNPH1, INPP4B, KLF12, LOC441155, MAP3K1, KMT2A, N4BP2L2_IT2, NEFL, NFATC3, PCM1, PCNX1, PDXDC2P, PHC3, POLR2J2, PSCP1, REPS1, RPP38, SLC7A6, ST3GAL1, STX16, TIMM8A, TRAF3IP3, TXK, TXLNGY, USP9Y	[107]
Eozinofili	ABHD2, ACACB, TRMO, CAT, CCR3, CLC, CYSLTR2, ADGRE1, EPN2, GALT, PTGDR2, HES1, HIST1H1C, HRH4, CD101, IL5RA, KBTBD11, KCNH2, LRP5L, MYO15B, RCOR3, RNASE2, RRP12, SIAH1, SMPD3, SYNJ1, TGIF1, THBS1, THBS4, TIPARP, TKTL1	[107]
Folikularne B pomoćne stanice	B3GAT1, CXCR5, LDLRAD4, CDK5R1, CHGB, CHI3L2, CXCL13, HEY1, HIST1H4K, ICA1, KCNK5, KIAA1324, MAF, MAGEH1, MKL2, MYO6, MYO7A, PASK, PDCD1, POMT1, PTPN13, PVALB, SH3TC1, SIRPG, SLC7A10, SMAD1, ST8SIA1, STK39, THADA, TOX, TSHR, ZNF764	[107]
Pomoćničke T stanice	ANP32B, ASF1A, ATF2, BATF, BORA, CD28, DDX50, FAM111A, FRYL, GOLGA8A, ICOS, ITM2A, LRBA, NAP1L4, NUP107, PHF10, PPP2R5C, RPA1, SEC24C, SLC25A12, SRSF10, TRA, UBE2L3, YME1L1	[107]
Citotoksične stanice	APBA2, APOL3, CTSW, DUSP2, GNLV, GZMA, GZMH, KLRB1, KLRD1, KLRF1, KLRK1, NKG7, RORA, RUNX3, SIGIRR, WHAMMP3, ZBTB16	[107]
Gama delta T stanice ($\gamma\delta$ T)	ACP5, AQP9, BTN3A2, C1orf54, CARD8, CCL18, CD209, CD33, CD36, CDK5, IL10RB, KLRF1, LGALS1, MAPK7, KLHL7, KRT80, LAMC1, LCORL, LMNB1, MEIS3P1, MPL, FABP1, FABP5, FADD, MFAP3L, MINPP1, RPS24, RPS7, RPS9, DBNL, CCL13	[108]
T regulacijske stanice	CD72, CLEC5A, FOXP3, ITGA4, L1CAM, LIPA, LRP1, LRRC42, MARCO, MMP12, MND4, MRC1, MS4A6A, PELO, PLEK, PRSS23, PTGIR, ST8SIA4, STAB1	[108]
Citotoksične T stanice	ADRM1, AHS1, C1GALT1C1, CCT6B, CD37, CD3D, CD3E, CD3G, CD69, CD8A, CETN3, CSE1L, GEMIN6, GPT2, GZMK, IL2RB, LCK, MPZL1, PIK3IP1, PTRH2, TIMM13, ZAP70	[108]

Izračunate *GSVA* vrijednosti obogaćivanja u tumorskim uzorcima upotrijebljene su za identifikaciju imunskih podtipova pomoću hijerarhijskog aglomerativnog klasteriranja, na temelju kvadratne euklidske udaljenosti i Ward-ove metode povezivanja. Za vizualizaciju toplinske mape *GSVA* vrijednosti za sve uzorke korištena je funkcija *heatmap* unutar R programa.

Za procjenu povezanosti stupnja infiltracije tumora sa imunološkim stanicama i sveukupnog preživljenja pacijenata u eksperimentalnoj kohorti, napravili smo analizu sveukupnog preživljenja. Za ovu analizu koristili smo *GraphPad Prism5* statistički program. Za izradu krivulja preživljenja *Kaplan-Meierovom* metodom, u program smo unijeli informacije o pripadnosti identificiranim podtipovima i vremena preživljenja pacijenata. Smrt se smatrala događajem od interesa. Sveukupna razlika između krivulja preživljenja ispitala se log-rang statističkim testom. Razlike u preživljenju između pojedinačnih krivulja preživljenja ispitane su pomoću analize varijanci (*ANOVA* test), sa korigiranom *p* vrijednosti pomoću Bonferroni metode. Korigirana *p* vrijednost manja od 0,05 smatrala se statistički značajnom.

3.7. Analiza profila ekspresije malih molekula RNA dobivenih sekvenciranjem slijedeće generacije

Rezultati sekvenciranja malih molekula RNA su neobrađene sekvence duljine 51 nukleotida. Prije daljnje analize potrebno je provjeriti kvalitetu neobrađenih sekvenci, kako bi se utvrdila kvaliteta samog sekvenciranja. Nakon provjera kvalitete sekvenci i uklanjanja sekvenci adaptera, slijedi poravnavanje sekvenci na referentni ljudski genom verzije GRCh38_p10 i identifikacija miRNA molekula. U svrhu identifikacije miRNA čija je razina promijenjena u tumorskim uzorcima, u odnosu na zdrave kontrole, napravljena je analiza diferencijalne ekspresije. Potom su identificirani ciljni geni miRNA sa promijenjenom razinom ekspresije, na temelju predviđenih ili validiranih interakcija zabilježenih u bazama podataka. Na temelju identificiranih ciljnih gena miRNA, napravljen je interaktom miRNA sa promijenjenom razinom ekspresije i diferencijalno eksprimiranih gena.

3.7.1. Provjera kvalitete neobrađenih sekvenci

Profil ekspresije malih RNA vrsta u uzorcima tumora pluća skvamoznih stanica, u odnosu na zdrave kontrole, istražili smo pomoću sekvenciranja slijedeće generacije i bioinformatičkih analiza dobivenih podataka. Priprema knjižnice za sekvenciranje ranije je opisan (poglavlje 3.5.). Pripremljene knjižnice, duljine 51 parova baza, sekvencirane su na HiSeq2500 platformi (Illumina,

SAD). Rezultati sekvenciranja dobiveni su u .fastq formatu. Kao i kod transkriptomskih analiza, i u ovom slučaju je prvi korak u analizi provjera kvalitete očitavanja sekvenci. Kvaliteta očitanih sekvenci provjerena je pomoću mrežnog oblaka *BaseSpace* i *FASTQC* aplikacije. Kvaliteta očitavanja sekvenci procijenjena je na temelju dodijeljenih Phred vrijednosti kvalitete, udjela GC baza u sekvenci, zastupljenosti motiva kratkog niza (engl. *k-mers*) i stope dupliciranih sekvenci.

3.7.2. Poravnavanje neobrađenih sekvenci na referentni genom i identifikacija miRNA

Nakon provjere kvalitete očitanih sekvenci, sekvence smo poravnali na referentni genom verzije GRCh38_p10 pomoću mrežnog alata *sRNAbench* u grupnom načinu rada (engl. *batch mode*). U alat su unesene neobrađene sekvence u .fastq obliku za svaki uzorak u analizi. Odabrani su slijedeći parametri za poravnavanje na referentni genom:

- Odabrana vrsta je čovjek i referentni genom verzije GRCh38_p10
- Uklanjanje sekvenci adaptera iz Illumina *TruSeq* kompleta: Odabrano je uklanjanje sekvenci adaptera sa 3' kraja koje odgovaraju nukleotidnom slijedu TGGAATTCTCGGGTGCCAAGGG. Minimalna odabrana duljina adaptera za prepoznavanje je 10 nukleotida; dopušteno je nepoklapanje sekvenci adaptera i referentne sekvence za 1 nukleotid
- Kontrola kvalitete poravnavanja: Za kontrolu kvalitete poravnavanja odabrana je minimalna srednja vrijednost kvalitete; Phred vrijednost; a granična Phred vrijednost kvalitete postavljena je na 20
- Analiza miRNA: Za identifikaciju miRNA odabrana je *miRBase* baza podataka
- Poravnanje pomoću *bowtie* alata: Odabrana duljina *seed* sekvence je 20 nukleotida, minimalni broj očitanih sekvenci:2; minimalna duljina očitane sekvence:15; dopušteni broj nepoklapanja sa referentnim genomom u *seed* regiji: 1 nukleotid

sRNAbench alat detektira sekvence adaptera na slijedeći način: dio očitanih sekvenci (100 000 sekvenci) se poravnava na ljudski genom pomoću uključene opcije *Bowtie seed*. Potom se sekvenca adaptera definira kao sekvenca koja počinje sa prvim nepoklapanjem sa genomom. Sve sekvence adaptera se potom grupiraju i najčešća sekvenca se smatra adapterskom sekvencom. Sekvence kojima su uklonjeni adaptorske sekvence i koje imaju *Phred* vrijednosti očitane kvalitete veće od 20 se poravnavaju na genom pomoću alata *bowtie1*. Poravnavanje *seed* regije pomoću *bowtie1* alata koristi se u svrhu detekcije isomiR-a, sa duljinom *seed* sekvence od 20 nukleotida, a očitane sekvence se koriste ako imaju najviše 10 poravnanja na genom.

Poznato je da miRNA postoje u kanonskom obliku i u obliku izoformi. Stoga se postavlja pitanje dali se ekspresija zrele miRNA treba procijeniti samo na temelju ekspresije kanonske forme ili zbrojem kanonske forme sa izoformama. *sRNAbench* alat dodjeljuje očitane sekvence miRNA koje se poravnavaju u blizini kanonske forme. Blizina je definirana kao [početak kanonske forme miRNA – 3nt ; završetak kanonske forme miRNAe + 5 nt].

3.7.3. Analiza promijenjeno eksprimiranih miRNA

Analiza diferencijalne ekspresije identificiranih miRNA napravljena je pomoću mrežnog alata *sRNAdet*. Navedeni alat za analizu diferencijalne ekspresije miRNA između ispitivane i kontrolne skupine koristi pet algoritama: *DESeq*, *DESeq2*, *EdgeR*, *NOISeq* i *ttest*. Za naše analize, kao algoritam za računanje razine promijenjene ekspresije miRNA odabrali smo *DESeq2* metodu, jer smo nju koristili i u transkriptomskim analizama. Za analizu su korištene poravnate sekvence na referentni genom sa uklonjenim adapterima, dobivene alatom *sRNAbench*. miRNA čija je ekspresija u tumorskom tkivu dva puta veća nego u zdravom tkivu, $|\log_2FC| \geq 1$, sa korigiranom p vrijednošću $p < 0,05$ smatrane su promijenjeno eksprimiranima.

3.7.4. Predikcija ciljnih gena miRNA

miRNA reguliraju ekspresiju ciljnih transkripata promovirajući njihovu degradaciju ili ograničavajući njihovu translaciju. Informacije o vezanju miRNA na ciljne transkripte pohranjene su u bazama podataka, a dobivene su na temelju eksperimentalne validacije interakcije miRNA sa ciljnim genima ili na temelju računalne predikcije pomoću različitih algoritama. Predikcije interakcije miRNA sa ciljnim genima temelje se na procjeni komplementarnosti baza miRNA u području *seed* regije (prvih osam nukleotida miRNA) sa ciljnim transkriptom, očuvanosti ciljnog mjesta između vrsti, pristupačnosti ciljnog mjesta vezanja, slobodnoj energiji miRNA-mRNA duplekta ili kombinaciji navedenih karakteristika.

U svrhu identifikacije gena za koje se miRNA sa promijenjenom razinom ekspresije potencijalno vežu, nakon analize diferencijalne ekspresije miRNA u tumorskim u odnosu na zdrave uzorke, napravljena je analiza predikcije ciljnih gena miRNA. Analiza je napravljena u R programu pomoću paketa *multiMir*. U paketu su integrirane informacije o potvrđenim i prediktivnim interakcijama miRNA sa ciljnim genima iz sveukupno 14 baza podataka. Informacije o potvrđenim interakcijama miRNA/ciljni geni dolaze iz baza podataka *miRecords*, *miRTarBase* i *TarBase*. Informacije o predviđenim interakcijama dolaze iz baza podataka *DIANA-microT*, *ElMMo*,

MicroCosm, *miRanda*, *miRDB*, *PicTar*, *PITA* i *TargetScan*. Informacije o povezanosti miRNA sa bolestima i lijekovima preuzete su iz baza podataka *miR2Disease*, *Pharmaco-miR VerSe* i *PhenomiR*. Informacije o interakcijama promijenjeno eksprimiranim miRNA i ciljnih gena dobivene su pomoću funkcije *get_multimir*. Kako bi smanjili broj potencijalno lažno pozitivnih rezultata, ciljnim genima smatrali smo one gene čija je interakcija sa miRNA ili eksperimentalno potvrđena ili je interakcija predviđena u više od tri navedene prediktivne baze podataka.

Kako bi bolje razumjeli ulogu identificiranih ciljnih gena miRNA, njihovu funkciju provjerili smo pomoću analize genske ontologije. Za ovu analizu korišten je mrežni alat *PANTHER* dostupan na mrežnoj stranici <http://geneontology.org/>. Kao anotacijski set GO pojmova odabran je *PANTHER GO-Slim* set. Za izračun statističke značajnosti obogaćivanja GO pojmovima, odabran je *Fisher* egzaktni test, a za korigiranu p vrijednost odabrana je *FDR* korekcija. Ponavljajući GO pojmovi uklonjeni su pomoću alata *Revigo*, dostupnog na mrežnoj stranici <http://revigo.irb.hr/>, sa odabranom dopuštenom sličnošću pojmova 0,7. Obogaćeni GO pojmovi sa prilagođenom p vrijednošću manjom od 0,05 su vizualizirani su u *Cytoscape* programu.

3.7.5. Analiza interakcija miRNA-mRNA

Nakon identifikacije ciljnih gena miRNA, zanimalo nas je koji ciljni geni diferencijalno eksprimirane miRNA su diferencijalno eksprimirani u suprotnom smjeru, odnosno koji ciljni geni miRNA sa pozitivno promijenjenom ekspresijom u tumorskom tkivu imaju negativno promijenjenu ekspresiju na mRNA razini i obrnuto. Lista diferencijalno eksprimiranih miRNA i lista identificiranih ciljnih gena, eksperimentalno potvrđenih ili predviđenih u više od tri baze podataka, korištene su za pronalazak preklapanja između listi pomoću *multiMir* paketa u R programu. Ukoliko je u bazama podataka potvrđena interakcija miRNA/mRNA ili je interakcija predviđena u najmanje 3 baze podataka, smatrali smo da postoji interakcija između diferencijalno eksprimirane miRNA i ciljnog gena. Pronađene miRNA-mRNA interakcije su potom vizualizirane u *Cytoscape* programu. U *Cytoscape* su unesene liste promijenjeno eksprimiranih miRNA i mRNA te su ciljni geni miRNA označeni kao *target-i* u mreži. Zasebno su vizualizirane interakcije pozitivno promijenjenih miRNA i negativno promijenjenih ciljnih gena u jednoj mreži, a u drugoj mreži su vizualizirane interakcije negativno promijenjenih miRNA i pozitivno promijenjenih mRNA.

Također nas je zanimalo koji GO pojmovi su obogaćeni za one ciljne gene za koje je pokazana interakcija sa miRNA. GO analizu napravili smo pomoću mrežno dostupnog alata *Enrichr* za kategoriju Biološki procesi.

3.8. Analiza proteoma

U sklopu ove doktorske disertacije napravljena je i analiza proteoma tumora pluća skvamoznih stanica i zdravih kontrolnih uzoraka, na razini pilot studije. Iz FFPE rezova tkiva izolirani su proteini, određena im je koncentracija, potom su proteini pocijepani u peptidne fragmente. Ispiranjem peptidne smjese pomoću hidroksilnih magnetskih kuglica, uzorci su pripremljeni za analizu na spektrometru masa. Na spektrometru masa peptidi su razdvojeni pomoću gradijenta acetonitrila te su zabilježeni njihovi spektri. Peptidni sljedovi su potom analizirani na temelju omjera mase i naboja, u svrhu identifikacije proteinske ekspresije pomoću *MaxQuant* programa. Pomoću *Perseus* programa matrica identificiranih proteina je filtrirana, u svrhu uklanjanja nepouzdanih rezultata identificiranih proteina. Potom su dobivene vrijednosti intenziteta transformirane ($\log_2$ skala) i uspoređeni su rezultati bioloških replika, u svrhu provjere kvalitete analize. Razina promjene ekspresije identificiranih proteina je potom uspoređena između dvije skupine uzoraka, tumorskih i zdravih kontrola, pomoću t-testa.

3.8.1. Izolacija proteina iz FFPE rezova tkiva

U svrhu profiliranja proteinske ekspresije, napravljena je analiza proteoma tumora pluća skvamoznih stanica i zdravih kontrolnih uzoraka. U ovu analizu uključena su tri uzorka tumora pluća skvamoznih stanica i tri uzorka zdravih kontrola, sa tri biološke replike od svakog uzorka. Za izolaciju proteina korištena su dva reza FFPE tkiva (debljine 10 μ m). Prije same izolacije proteina iz FFPE uzoraka, prvo im je uklonjen parafin dodavanjem organskog otapala ksilena. Nakon 2 minute inkubacije, ksilen je uklonjen centrifugiranjem. Ispiranje ksilenom je ponovljeno tri puta, nakon čega su rezovi tkiva dva puta isprani sa 100 % etanolom, jedanput sa 70 % etanolom, jedanput sa 56 % etanolom i u konačnici sa destiliranom vodom. Između svakog ispiranja, rezovi tkiva su centrifugirani te im je uklonjen supernatant. Proteini u uzorcima su potom izolirani dodavanjem pufera za ekstrakciju proteina (100 mM HEPES pH 7,5; 4% SDS; 50mM DTT), uz inkubaciju od 1 sata na 95°C, uz laganu rotaciju na 300 rpm, nakon čega su kratko ohlađeni na ledu, uz uklanjanje ekstrakcijskog pufera centrifugiranjem. Potom je uzorcima dodana otopina ohlađenog acetona i metanola (omjer 9:1) te su inkubirani 2 sata na temperaturi -80°C. Nakon

inkubacije, supernatant je uklonjen centrifugiranjem, te su uzorci isprani ohlađenim metanolom četiri puta. Pelet je resuspendiran u 50 µl ohlađene otopine 100 mM NaOH ultrazvučnom sonikacijom na temperaturi 4°C. Otopini je po potrebi dodan 200 mM HEPES, kako bi se podesila pH vrijednost na 7.5-8.0. Svi koraci centrifugiranja bili su podešeni na 12 000 x g.

3.8.2. Određivanje koncentracije proteina

Koncentracija izoliranih proteina u FFPE uzorcima određena je pomoću kompleta *Pierce™ BCA Protein Assay* (Thermo Scientific, SAD). BCA je kolorimetrijska metoda koja detektira reducirani bakrov kation (Cu^+) u otopini proteina pomoću bicinkonične kiseline. Za određivanje nepoznate koncentracije proteina u otopini koristi se proteinski standard poznate koncentracije, a koncentracija proteina u otopini se određuje usporedbom vrijednosti apsorbancije sa standardnom krivuljom poznatih koncentracija.

Prvo se proteinski standard albumina iz goveđeg seruma (engl. *bovine serum albumin*, BSA) koncentracije 2 mg/ml (BSA u otopini 0,9% NaCl i 0,05% NaN_3), serijski razrijedi u mQH_2O , u tri jažice. Potom se uzorci kojima se određuje koncentracija razrijede 10 puta u mQH_2O , također u tri jažice. Pripremi se radna otopina reagensa na način da se otopina B 50x razrijedi u otopini A. U jažice sa priređenim razrjeđenjima standarda i uzoraka se potom doda 80 µl radne otopine te se pločica sa uzorcima inkubira 30 minuta na temperaturi 37°C, nakon čega se mjeri asporbancija na 562 nm. Izmjerene vrijednosti asporbancija koriste se za računanje koncentracije proteina u uzorcima. Prvo se iz triplikata standarda poznate koncentracije (BSA) izračunaju srednje vrijednosti za svako razrjeđenje. Potom se napravi baždarna krivulja, čija se jednadžba pravca koristi za računanje nepoznatih koncentracija proteina u uzorcima. Dobiveni rezultat se pomnoži sa 10, jer je prilikom pripreme uzorak razrijeđen 10 puta. Dobivena koncentracija proteina u uzorku izražena je u mg/ml.

3.8.3. Priprema uzoraka i analiza na spektrometru masa

Nakon određivanja koncentracije proteina u uzorcima, proteini se prvo reduciraju kako bi se pocijepale disulfidne veze. Proteini su potom alkilirani da kovalentno modificiraju –SH skupinu cisteina, kako ne bi došlo do stvaranja novih disulfidnih veza. Na kraju se proteini fragmentiraju u peptidne sljedove dodavanjem enzima tripsina, koji proteine cijepa u peptidne fragmente iza aminokiselina lizin ili arginin (osim kada nakon njih slijedi prolin). Nakon tripsinizacije proteina, proteini se ispiru pomoću magnetskih perli, kako bi se u potpunosti uklonili reagensi (deterdženti)

u otopini proteina koji nisu kompatibilni sa spektrometrom masa. Magnetske perle, korištene za pripremu uzoraka za analizu na spektrometru masa, na svojoj površini imaju vezane karboksilne skupine na koje se kovalentno vežu proteini/peptidi prilikom ispiranja.

Za pripremu uzoraka za analizu na spektrometru masa uzeto je 20 µg izoliranih ukupnih proteina po uzorku. Uzorcima je dodan 160mM kloracetamid u svrhu alkilacije proteina, a volumen je podešen na 40µl sa puferom za ekstrakciju koji sadrži reducirajući spoj ditiotritol (DTT). Uzorci su potom inkubirani 5 minuta na 95°C u svrhu redukcije i alkilacije disulfidnih veza u proteinima te su ohlađeni na sobnu temperaturu. Prije dodavanja magnetskih perli u uzorak, potrebno je pripremiti radnu otopinu perli. Korištene su magnetske perle *Sera-Mag™ Carboxylate-Modified Magnetic Beads* (Sigma-Aldrich, SAD). Set se sastoji od *Sera-Mag* otopine magnetskih perli A (hidrofilne) i *Sera-Mag* otopine magnetskih perli B (hidrofobne). Perle A i B (koncentracije 50 mg/ml) se pomiješaju u omjeru 1:1, a u odnosu na proteine u uzorku omjer perle/proteini treba biti 10:1. Nakon dodavanja magnetskih perli, u otopinu proteina dodan je i 100% etanol (kako bi se pospješilo vezanje proteina na perle). Nakon inkubacije od 5 minuta na sobnoj temperaturi (25°C) uz trešnju 1000 rpm, nevezani proteini se uklanjanju iz supernatanta uzorka pomoću magnetskog stalka. Perle sa proteinima se potom uzastopno ispiru sa 80 % etanolom. Nakon uklanjanja supernatanta, uzorcima je dodana otopina tripsina (0,8µg) u 100 mM amonijevom bikarbonatu (pH 8). Proteini su inkubirani na 37 °C 12-18 sati uz trešnju od 1000 rpm. Uzorcima je potom uklonjen supernatant pomoću magnetskog stalka, nakon čega su isušeni vakuumom. Ovako pripremljeni uzorci pohranjeni su na -80°C, do analize na spektrometru masa.

Za analizu uzoraka na spektrometru masa, isušene peptide u uzorcima potrebno je resuspendirati u 20µl otopine 2% acetonitrila i 0,1% formijatne kiseline, a potom sonicirati u ultrazvučnoj vodenoj kupelji 5 minuta i inkubirati 5 minuta na sobnoj temperaturi uz trešnju 1000 rpm. Uzorci su potom centrifugirani na 5000 x g i supernatanti prebačeni u tubice za analizu na spektrometru masa Q Exactive HF-X (*Thermo Fisher*, SAD). Volumen za iniciranje u uređaj podešen je na 5µl. Peptidi su razdvojeni na kolonama promjera 75 µm, te gradijentu acetonitrila (ACN) tijekom 2 sata (100 minuta u 24 % ACN, 20 minuta u 30 % ACN). MS spektri postavljeni su za *m/z* raspon 350-1600 sa MS1 rezolucijom od 70 000 za 200 *m/z*, i MS2 rezolucijom od 17 500.

3.8.4. Identifikacija proteina

Omjeri mase i naboja peptida korišteni su za identifikaciju proteina u uzorcima pomoću *MaxQuant* programa (verzija 1.6.3.3.). *MaxQuant* detektira peptidne pikove, ocjenjuje peptide, pretražuje baze podataka u svrhu identifikacije i kvantifikacije proteina u uzorku te daje sažetu statistiku eksperimenta. U našim analizama, za potrebe identifikacije i kvantifikacije proteina u uzorcima, u *MaxQuant* programu većina parametara je ostavljeno u zadanim vrijednostima (engl. *default parameters*). Promijenjeni su slijedeći parametri:

- Odabrani enzim za cijepanje proteina prilikom pripreme uzoraka- Trypsin
- Način kvantifikacije proteina- kvantifikacija bez oznaka (engl. *label- free quantification*)
- Maksimalna tolerancija pogreške (ppm)- 20
- FDR vrijednost za identifikaciju peptida i proteina- 0,01
- Minimalna potrebna duljina peptida korištena za analizu- 7 aminokiselina
- Modifikacije uključene u kvantifikaciju proteina- oksidacija i acetilizacija
- Za kvantifikaciju potrebna je detekcija minimalno- 2 peptida
- Referentni proteom- odabran je Uniprot *HomoSapiens reviewed* proteom

Identificirani proteini spremljeni su u obliku *proteinGroups.txt* datoteke, koja je korištena za daljnje analize.

3.8.5. Provjera kvalitete proteomskih podataka i identifikacija diferencijalno eksprimiranih proteina

Kvaliteta proteomskih podataka provjerena je pomoću programa *Perseus* (verzija 1.6.0.2.). U program se prvo unesu rezultati identifikacije proteina (*proteinGroups.txt*), napravljeni pomoću *MaxQuant* programa. Prvo je potrebno iz matrice ukloniti one proteine koji su identificirani samo prema mjestu (engl. *identified only by site*). Ti proteini su identificirani samo prema peptidima koji nose jednu ili više modificiranih aminokiselina. Identifikacija proteina iz takvih peptida je često manje pouzdana, pa se proteini identificirani isključivo prema takvim peptidima stoga uklanjaju iz daljnje analize.

Nakon uklanjanja proteina identificiranih samo prema modificiranim peptidima, slijedeći korak je logaritamska transformacija *LFQ* intenziteta. Ovaj korak radi se kako bi distribucija *LFQ* intenziteta imala što sličniju distribuciju normalnoj distribuciji. Distribucija *LFQ* intenziteta prije i nakon transformacije provjerena je vizualno, pomoću histograma u *Perseus* programu.

Nakon transformacije *LFQ* intenziteta, potrebno je grupirati uzorke. U našu analizu uključena su tri uzorka tumora pluća skvamoznih stanica i tri uzorka zdravih kontroli. Svaki uzorak zastupljen je sa tri biološke replike, stoga je ukupan broj uzoraka u analizi 18 (9 uzoraka tumora i 9 uzoraka kontroli). Uzorci su grupirani na temelju fenotipa i označeni su kao Tumor ili Kontrola.

Kako bi provjerili kvalitetu dobivenih rezultata u svim biološkim replikama, napravljena je korelacija *LFQ* intenziteta bioloških replika svakog uzorka u analizi. Korelacija *LFQ* intenziteta između uzoraka napravljena je pomoću Pearson korelacije, te su rezultati vizualizirani pomoću *Multi-scatter* grafa. Rezultati korelacije bliže jedinici, označavaju jaču korelaciju *LFQ* intenziteta između bioloških replika.

Grupiranje uzoraka na temelju *log2 LFQ* intenziteta provjereno je vizualno, pomoću PCA analize i hijerarhijskog klasteriranja unutar *Perseus* programa, odabirom ikonice PCA/hijerarhijsko klasteriranje na sučelju programa.

Za identifikaciju diferencijalno eksprimiranih proteina u tumorskom tkivu, u odnosu na zdrave kontrole, korišten je *Perseus* program i *two sample t-test*. T test pretpostavlja da su varijance *LFQ* intenziteta jednake unutar svake skupine uzoraka. Ukoliko je razlika u razini ekspresije proteina između dvije skupine 2 puta veća ($\log_2FC=1$) i FDR vrijednost manja od 0,05 smatrali smo da protein ima statistički značajno promijenjenu razinu ekspresije u tumorskom tkivu u odnosu na zdrave kontrole. Rezultati diferencijalne ekspresije vizualizirani su pomoću *Volcano* grafa odabirom ikonice na korisničkom sučelju programa *Perseus*.

3.8.6. Usporedba razine ekspresije mRNA, proteina i miRNA

U svrhu korelacije ekspresije proteina sa ekspresijom mRNA i miRNA, rezultati *omics* metoda su preklapljeni i uspoređeni pomoću Excel programa. Kako bi vidjeli dali su geni koji su diferencijalno eksprimirani na mRNA razini, diferencijalno eksprimirani i na proteinskoj razini, uspoređivane su vrijednosti njihove ekspresije. Za ovu usporedbu korišteni su samo oni geni čija je ekspresija potvrđena u obje testirane kohorte. Nadalje, kako bi provjerili potencijalni utjecaj na utišavanje proteina povećanom ekspresijom miRNA, usporedili smo listu ciljnih gena prekomjerno eksprimiranih miRNA sa smanjeno eksprimiranim proteinima. Identificirane potencijalne interakcije protein-miRNA su prikazane tablično, dok je utjecaj miRNA na proteine sa prognostičkim kapacitetom vizualiziran pomoću *Cytoscape* programa u obliku mrežnih interakcija.

4. REZULTATI

60 slika, 28 tablica

4. REZULTATI

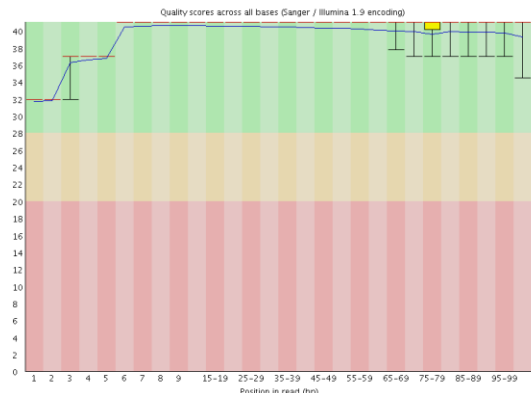
4.1. Rezultati analize transkriptoma

Kako bi utvrdili razliku u ekspresijskom profilu gena u tumoru pluća skvamoznih stanica, u odnosu na zdrave kontrole, pripremili smo knjižnice za sekvenciranje transkriptoma metodom sekvenciranja slijedeće generacije. Knjižnice za analizu su pripremljene za 23 FFPE uzoraka pacijenata oboljelih od tumora pluća skvamoznih stanica i tri FFPE uzorka zdravih kontrola. Knjižnice su sekvencirane na platformi Illumina HiSeq4000.

4.1.1. Rezultati analize provjere kvalitete neobrađenih sekvenci

Prije daljnje bioinformatičke analize rezultata potrebno je procijeniti kvalitetu očitanih sekvenci dobivenih sekvenciranjem. U našim analizama za procjenu kvalitete neobrađenih sekvenci koristili smo mrežni oblak *BaseSpace* i aplikaciju *FASTQC*. Aplikacija koristi rezultate sekvenciranja u .fastq obliku datoteke za analizu kvalitete. Kvaliteta očitanih neobrađenih sekvenci se ispituje kako bi se osiguralo da su baze koje je uređaj za sekvenciranje dodijelio sekvenci pouzdano očitane, da udio GC baza u očitanim sekvencama odgovara očekivanoj distribuciji, da motivi kratkog niza nisu previše zastupljeni i da je razina dupliciranih sekvenci prihvatljiva. Analiza kvalitete određuje se po modulima.

U modulu *Kvaliteta očitanih sekvencija po bazama*, kvaliteta očitanih sekvenci se izražava Phred vrijednostima kvalitete. Phred vrijednost kvalitete 10 (Q10) označava 90% vjerojatnost da je baza dobro očitana, Phred vrijednost kvalitete 20 (Q20) označava 99% vjerojatnost da je baza dobro očitana, a Phred vrijednost kvalitete 30 (Q30) označava 99,9% vjerojatnost da je baza dobro očitana. Rezultati ovog modula su pokazali da su Phred vrijednosti kvalitete očitanih sekvenci u svim sekvenciranim uzorcima visoke, odnosno da možemo sa visokom pouzdanošću reći da su baze dobro očitane. Na Slici 11. prikazan je primjer rezultata ovog modula za jednu očitanu sekvencu u uzorku. Na ovom primjeru vidljivo je da su vrijednosti kvalitete nešto niže na početku i na kraju očitane sekvence, što je očekivani rezultat na većini platformi za sekvenciranje, uključujući i platformu HiSeq4000. Usprkos tome, vidljivo je da su vrijednosti kvalitete očitane sekvence visoke za sve baze u sekvenci, odnosno da sve baze u očitanoj sekvenci imaju Phred vrijednost kvalitete >Q30.



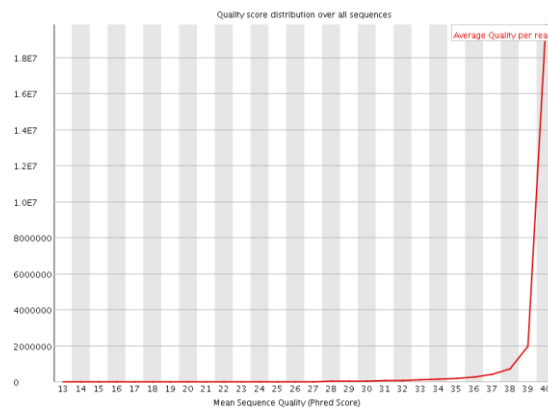
Slika 11. Primjer rezultata u modulu *Kvaliteta očitanih sekvenca po bazama*. Pozadina grafa označava visoku Phred vrijednost kvalitete očitanih sekvenca, žuta boja označava prihvatljive vrijednosti kvalitete, crvena boja označava niske vrijednosti kvalitete očitanih sekvenca.

U Tablici 9. prikazane su Phred vrijednosti kvalitete za sve uzorke u analizi. Vrijednosti navedene u tablici označavaju postotak baza u očitanim sekvencama koje imaju Q20 i Q30 vrijednosti kvalitete. Kao što je vidljivo iz tablice, više od 90% baza u očitanim sekvencama ima visoku (Q30), a prihvatljivu vrijednost kvalitete (Q20) ima >95% baza u očitanim sekvencama.

Tablica 9. Sažetak rezultata kvaliteta očitanih sekvenca po bazama za sve uzorke u analizi. Vrijednosti prikazane u drugom i trećem stupcu označavaju postotak baza u sekvencama koje imaju vrijednosti kvalitete veće od Q20 i Q30.

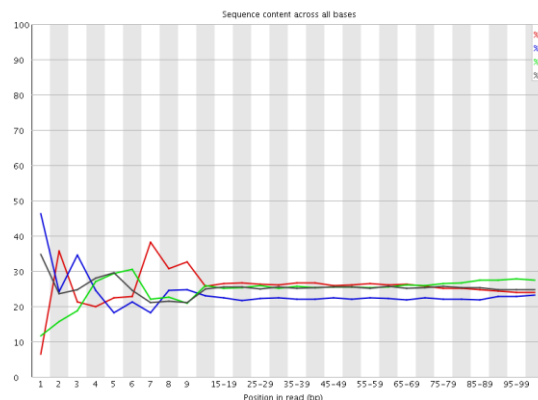
Uzorak	Q20 (%)	Q30 (%)
12210	98,71	96,45
12364	98,73	96,58
12545	97,66	94,15
12698	97,79	94,43
16734	97,79	94,24
17428	98,01	94,86
17444	98,48	95,75
17664	98,77	96,58
18356	98,73	96,58
19694	98,24	95,16
20109	97,58	93,52
20432	97,77	94,07
21671	97,81	94,15
2304	97,25	93,44
244	97,94	94,46
25486	98,78	96,49
25609	96,29	92,06
25850	98,49	96,06
3260	97,57	93,70
44	95,82	89,30
4613	98,0	94,94
4737	97,60	93,86
55	97,63	93,99
6452	98,85	96,81
7202	97,88	94,61
8323	98,57	96,01

Rezultati modula *Sažetak vrijednosti kvalitete po očitanoj sekvenci* potvrđuju dobre rezultate modula Kvaliteta očitanih sekvenca po bazama za sve uzorke u analizi. Na Slici 12. prikazan je primjer rezultata ovog modula za jednu očitanu sekvencu u uzorku. Na ovom primjeru vidljivo je da većina sekvenci u uzorku ima Phred vrijednosti kvalitete >35 i da ne postoji podskupina sekvenci sa malim vrijednostima kvalitete.



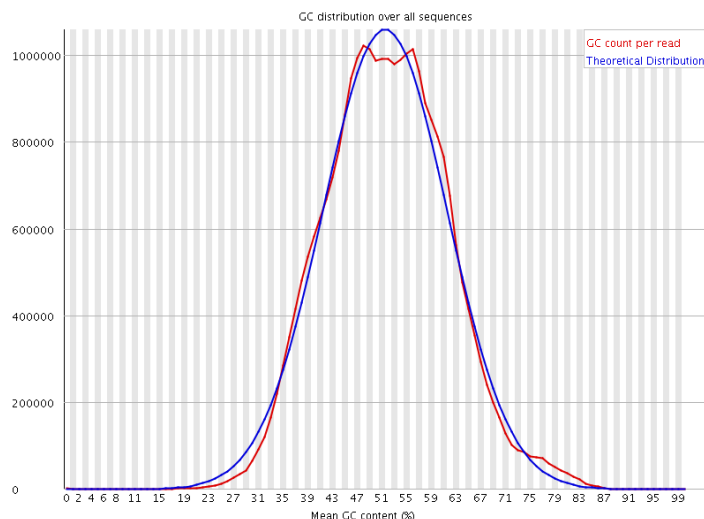
Slika 12. Primjer rezultata za modul *Sažetak vrijednosti kvalitete po očitanoj sekvenci*.

Rezultati modula *Provjera očitanih sekvenci po bazama* također ukazuju da su knjižnice dobro pripremljene za sve uzorke u analizi. U ovom modulu očekivana je podjednaka zastupljenost svake baze u sekvenci. Kao što je vidljivo na primjeru rezultata na Slici 13. u prvih 10-ak baza vidljivi su nasumični sljedovi, što je i očekivano zbog toga što se prilikom pripreme knjižnice koriste nasumični heksameri. Također, vidljivo je da je u ostatku očitane sekvence zastupljenost svake od četiri baza očekivano podjednaka.



Slika 13. Primjer rezultata za modul *Provjera očitanih sekvenci po bazama*

Rezultati modula *Udio GC baza u očitanim sekvencama* pokazuju da u svim uzorcima u analizi očekivani udio GC baza u sekvencama ne odstupa puno od očekivane distribucije GC baza. Na Slici 13. prikazan je primjer rezultata ovog modula za jednu sekvencu uzorka, a sažetak udjela GC baza u sekvencama svih uzoraka u analizi prikazan je u Tablici 10.

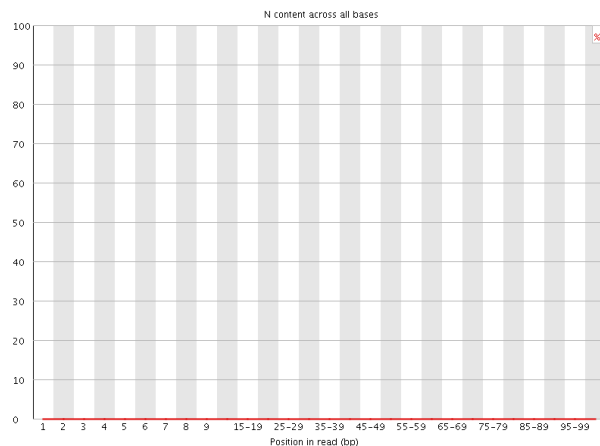


Slika 14. Primjer rezultata za modul *Udio GC baza u očitanim sekvencama*. Crvenom krivuljom prikazan je udio GC baza u sekvenci uzorka, dok je plavom bojom označena očekivana distribucija GC baza u sekvenci.

Tablica 10. Sažetak rezultata za modul *Udio GC baza u očitanim sekvencama* za sve uzorke u analizi

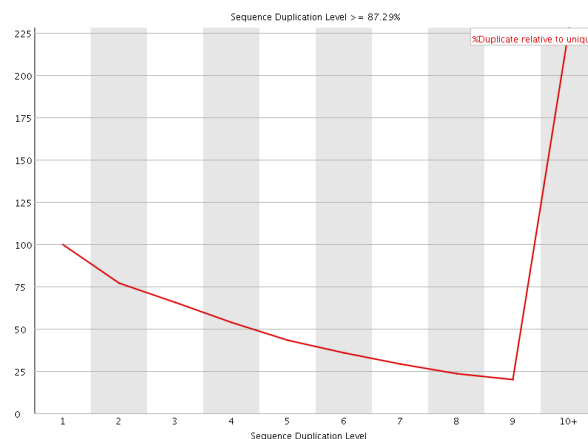
Uzorak	GC (%)
12210	49,25
12364	48,44
12545	48,71
12698	48,46
16734	51,54
17428	48,97
17444	49,04
17664	49,39
18356	47,03
19694	49,07
20109	51,51
20432	49,64
21671	51,02
2304	47,78
244	46,88
25486	51,41
25609	51,24
25850	48,37
3260	50,75
44	49,42
4613	47,43
4737	51,57
55	47,52
6452	48,17
7202	48,72
8323	48,62

Rezultati modula *Udio N očitavanja u sekvenci* nadopunjuju rezultate modula *Sažetak kvalitete sekvence po bazama* i modula *Sažetak vrijednosti kvalitete po očitanoj sekvenci*. Ukoliko uređaj za sekvenciranje ne može odrediti bazu u sekvenci, označiti će je sa N. Sve očitane sekvence u našim uzorcima prošle su provjeru kvalitete u ovom modulu, što nije začuđujuće budući da su vrijednosti kvalitete sekvenci u našim uzorcima visoke, pa je i očekivano da je udio N očitavanja u sekvenci zanemarivo malen. Kao što je vidljivo na primjeru rezultata ovog modula (Slika 15.), sve baze u sekvenci su uspješno određene, na što ukazuje ravna crvena crta.



Slika 15. Primjer rezultata za modul *Udio N očitavanja u sekvenci*.

Svi uzorci u analizi nisu prošli provjeru kvalitete u modulu *Duplicirane sekvence*. Obzirom da protokol za pripremu knjižnice uključuje obogaćivanje kodirajućih regija gena te njihovo umnažanje PCR-om, ovi rezultati su očekivani i prihvatljivi. Primjer rezultata ovog modula prikazan je na Slici 16.



Slika 16. Primjer rezultata za modul *Duplicirane sekvence*

Rezultati provjere kvalitete očitanih sekvenci svih uzoraka u analizi ukazuju da su sekvence očitane sa visokom pouzdanošću i da zadovoljavaju kriterije kvalitete za nizvodne analize.

4.1.2. Rezultati analize poravnavanja neobrađenih sekvenci na referentni genom GRCh38

Kako bi mogli utvrditi kojim genima očitane sekvence pripadaju, potrebno ih je poravnati na referentni genom. U ovu svrhu korišten je mrežni oblak *BaseSpace* i aplikacija za poravnavanje *RNA-Seq Alignment*. Za poravnavanje je odabran referentni genom čovjeka verzije GRCh38 *Decoy*. Također, prilikom poravnavanja se detektiraju fuzije gena pomoću *Manta* alata, kvantificiraju se očitane sekvence na razini transkripta i gena pomoću alata *Salmon*, a navedena kvantifikacija se koristi u nizvodnim analizama.

Sažetak rezultata poravnavanja prikazan je u Tablici 11. Broj očitanih sekvenci koji su se poravnavali na referentni genom je 40-90 milijuna po uzorku, sa duljinom sekvence za poravnavanje od 101 nukleotida. Kao što je vidljivo iz tablice, na referentni genom uspješno je poravnato >90% ukupnog broja sekvenci u uzorcima, neovisno o smjeru očitavanja sekvenci. Udio prekomjerno zastupljenih sekvenci, poput rRNA ili mtRNA, je relativno malen i iznosi oko 5% za većinu uzoraka.

Tablica 11. Sažetak rezultata poravnavanja očitanih sekvenci na referentni genom.

Uzorak	Broj očitanih sekvenci po smjeru očitavanja	Duljina očitanih sekvenci (bp)	Udio prekomjerno zastupljenih sekvenci		Poravnate sekvence (% od ukupnog broja)	
			Očitavanje 1	Očitavanje 2	Očitavanje 1	Očitavanje 2
44	28 360 358	101	5,56	5,5	97,34	97,34
55	28 801 112	101	5,16	4,9	94,52	94,52
244	32 515 305	101	5,04	4,89	96,38	96,38
2304	22 927 759	101	3,65	2,88	90,68	90,68
3260	46 367 135	101	3,76	3,74	93,73	93,73
4613	30 379 692	101	2,21	2	95,75	95,75
4737	29 851 793	101	5,28	5,15	94,77	94,77
6452	22 975 324	101	3,88	3,69	96,52	96,52
7202	24 082 635	101	6,61	5,67	96,34	96,34
8323	24 269 105	101	3,46	3,06	92,94	92,94
12210	31 413 751	101	4,37	4,29	95,68	95,68
12364	29 608 982	101	2,47	2,5	95,69	95,69
12545	23 669 438	101	3,37	3,34	95,93	95,93
12698	24 628 592	101	3,48	3,39	95	95
16734	25 818 217	101	4,73	4,65	96,28	96,28
17428	25 102 499	101	5,98	5,83	95,35	95,35
17444	24 679 279	101	2,55	2,59	96,48	96,48
17664	21 505 354	101	1,81	1,74	96,2	96,2
18356	36 307 796	101	2,31	2,08	95,16	95,16
19694	30 540 069	101	4,4	4,11	93,98	93,98
20109	19 873 111	101	2,8	2,82	96,17	96,17
20432	23 237 424	101	3,02	3	96,88	96,88
21671	29 481 460	101	2,55	2,58	96,68	96,68
25486	24 637 729	101	3,27	3,32	96,71	96,71
25609	26 695 213	101	18,83	17,51	90,51	90,51
25850	20 230 477	101	5,8	4,86	89,27	89,27

Iz Tablice 12. vidljivo je da se najveći udio poravnatih sekvenci (otprilike 40-60 % baza) poravnao na kodirajuće regije. Manji udio sekvenci, oko 20-25%, se poravnao na UTR regije; oko 15% na intronske regije, dok je generalno najmanji udio sekvenci poravnat na intergenske regije.

Tablica 12. Sažetak rezultata poravnavanja očitanih sekvenci na određene regije genoma.

Uzorak	% baza poravnatih na kodirajuće regije	% baza poravnatih na UTR regije	% baza poravnatih na intronske regije	% baza poravnatih na intergenske regije
44	59,2	25,04	14,57	1,19
55	61,86	21,22	14,45	2,47
244	56,67	24,76	16,53	2,04
2304	46,25	23,18	16,58	13,99
3260	52,24	24,23	13,64	9,89
4613	54,46	23,08	14,46	8,00
4737	48,77	21,57	13,35	16,31
6452	53,54	25,63	13,19	7,64
7202	50,72	26,79	13,26	9,24
8323	48,91	25,19	14,61	11,26
12210	55,81	25,76	15,89	2,54
12364	53,07	26,47	15,27	5,19
12545	57,81	26,03	15,19	0,97
12698	50,88	25,74	14,83	8,55
16734	51,72	26,06	13,75	8,46
17428	47,78	26,14	15,25	10,83
17444	56,89	25,18	14,06	3,87
17664	44,89	21,06	13,56	20,49
18356	56,20	25,4	14,45	3,95
19694	58,62	21,82	14,47	5,1
20109	47,4	23,77	12,7	16,14
20432	54,47	24,87	13,45	7,22
21671	51,1	22,81	12,85	13,24
25486	50,95	21,94	12,37	14,74
25609	57,34	23,2	15,54	3,92
25850	41,55	18,07	12,74	27,64

Prilikom poravnavanja na referentni genom detektirane su i fuzije gena u uzorcima. U Tablici 13. prikazane su najčešće detektirane fuzije u tumorskim uzorcima. Kao što je vidljivo iz tablice, u najvećem broju uzoraka (26,92%) detektirana je fuzija LOC107985542 i IGLL5 gena, zatim LOC107984438 i MED13L (23,08 %) i KANSL1 i ARL17A/B (19,23 %). U najmanjem broju uzoraka (3,85 %) detektirana je fuzija SNORD3D i SNORD3C gena.

Tablica 13. Detektirane fuzije gena u tumorskim uzorcima

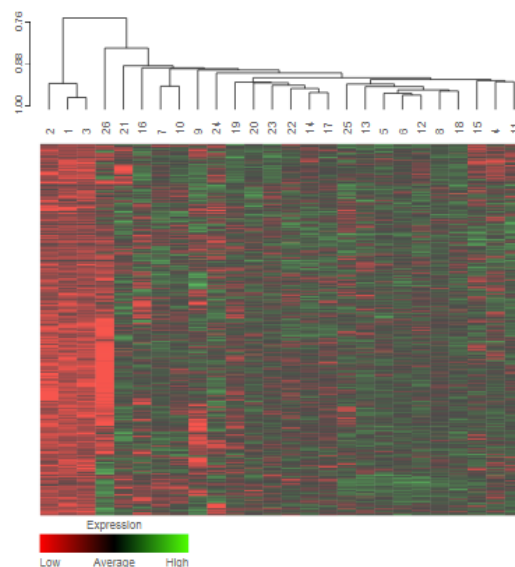
Gen 1	Mjesto spajanja 1	Gen 2	Mjesto spajanja 2	Fuzije u % uzoraka
LOC107985542	Chr22: 22759199	IGLL5	Chr22:22893782	26,92
LOC107984438	Chr12:116230532	MED13L	Chr12:116237704	23,08
KANSL1	Chr17:46094558	ARL17A/B	Chr17:46570867	19,23
SNORD3A	Chr17:19188232	SNORD3B-1	Chr17:19061967	19,23
MKNK1	Chr1:46585303	NEAT1	Chr11:65440843	11,54
TFG	Chr3:100720057	ADGRG7	Chr3:100629597	11,54
HLA-C	Chr6:31268946	HLA-B	Chr6:31354063	11,54
ITFG1	Chr16:47218868	LONP2	Chr16:48270013	7,69
SNORD3B-2	Chr17:19063920	SNORD3C	Chr17:19190239	7,69
SNORD3D	Chr17:19112457	SNORD3C	Chr17:19190230	3,85

Prilikom poravnavanja očitanih sekvenci na referentni genom, alat *Salmon* kvantificira očitavanja na razini gena i transkripata. Obzirom da je rezultat kvantifikacije velika matrica sa 38552 redova (geni) i 26 kolona (uzorci), rezultati nisu prikazani.

4.1.3. Rezultati analize diferencijalne ekspresije gena

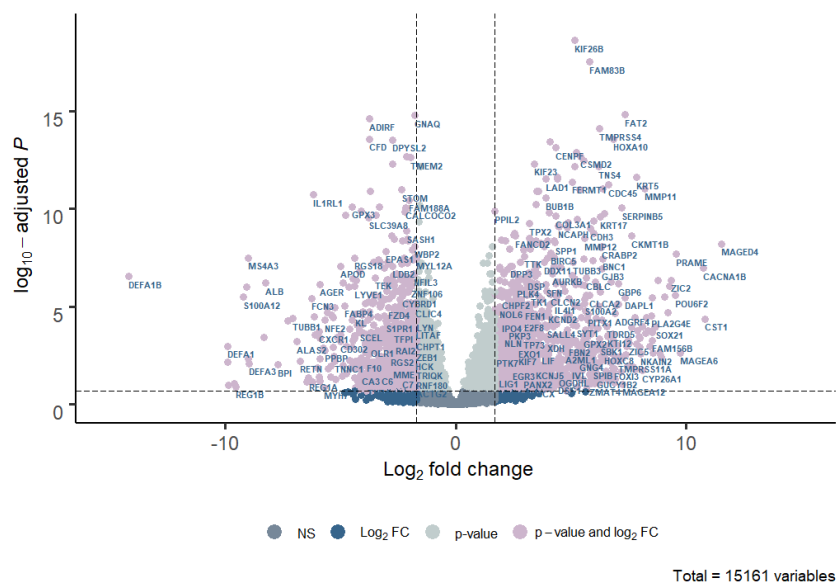
Kako bi ispitali profil genske ekspresije koji doprinosi fenotipu tumora pluća skvamoznih stanica, u odnosu na zdrave kontrole, napravljena je analiza diferencijalne ekspresije gena. Za ovu analizu korištena je matrica kvantificiranih očitavanja na razini gena, mrežni oblak *BaseSpace* i aplikacija *RNA-Seq Differential Expression*, koja diferencijalnu ekspresiju gena identificira pomoću *DESeq2* alata. Geni su okarakterizirani kao diferencijalno eksprimirani na temelju postavljene granične vrijednosti $|\log_2FC| \geq 2$ i prilagođene statističke značajnosti $p < 0,05$, prema metodi Benjamin-Hochberg [4]. Osim analize diferencijalne ekspresije, aplikacija omogućava vizualizaciju uzoraka u analizi pomoću toplinske mape, koja pokazuje relativnu sličnost uzoraka u analizi na temelju profila genske ekspresije.

Vizualizacija uzoraka u analizi putem toplinske mape, prikazane na Slici 17., pokazuje da se uzorci zdravih kontrola (uzorci 1-3) grupiraju zajedno jer imaju homogeniji profil ekspresije gena. Tumorski uzorci (uzorci 4-26) se također grupiraju zajedno, ali je profil genske ekspresije u tumorskim uzorcima heterogeniji u odnosu na kontrolu.



Slika 17. Toplinska mapa ekspresije gena u uzorcima. Redovi predstavljaju razinu ekspresije gena, a kolone predstavljaju uzorke u analizi. Dendrogram iznad toplinske mape ukazuje na Spearman- Rho korelaciju profila genske ekspresije između uzoraka

Rezultati analize diferencijalne ekspresije gena pokazali su da 1535 gena u uzorcima tumora pluća skvamoznih stanica ima promijenjenu ekspresiju u odnosu na zdrave kontrole. Od 1535 gena, 961 gen ima pozitivnu, a 574 gena negativno promijenjenu ekspresiju, u odnosu na zdrave kontrole. Na Slici 18. prikazani su rezultati diferencijalne ekspresije u obliku *volcano* plota. Kao što je vidljivo iz slike, statistički najznačajnije promijene detektirane su za gene *KIF26B* (*Kinesin Family Member 26B*), *FAM83B* (*Family With Sequence Similarity 83 Member B*), *FAT2* (*FAT Atypical Cadherin 2*), *GNAQ* (*G Protein Subunit Alpha Q*) i *ADIRF* (*Adipogenesis Regulatory Factor*).



Slika 18. Rezultati diferencijalne ekspresije prikazani *Volcano* plotom

20 gena sa najvećom pozitivno i negativno promijenjenom razinom ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole, prikazani su u Tablici 13. Među top 20 gena sa najvećom pozitivnom promjenom ekspresije (Tablica 14. A)), najveći udio su geni koji se povezuju sa promocijom tumorigeneze, proliferacijom i migracijom tumorskih stanica. Također, mnogi geni pripadaju MAGE (engl. *Melanoma associated genes*) i CTA (engl. *Cancer testis antigens*) obitelji gena. Među genima sa najvećom negativnom promjenom u razini ekspresije (Tablica 14. B)), u odnosu na zdrave kontrole, su geni koji sudjeluju u regulaciji imunološkog odgovora, poput DEFA1 (engl. *Defensin Alpha 1*), DEFA3 (engl. *Defensin Alpha 3*) i DEFA4 (engl. *Defensin Alpha 4*) gena.

Tablica 14. Prikaz 20 gena sa najvećom pozitivnom (A) i negativnom (B) promjenom u razini ekspresije u tumorskim uzorcima u odnosu na zdrave kontrole

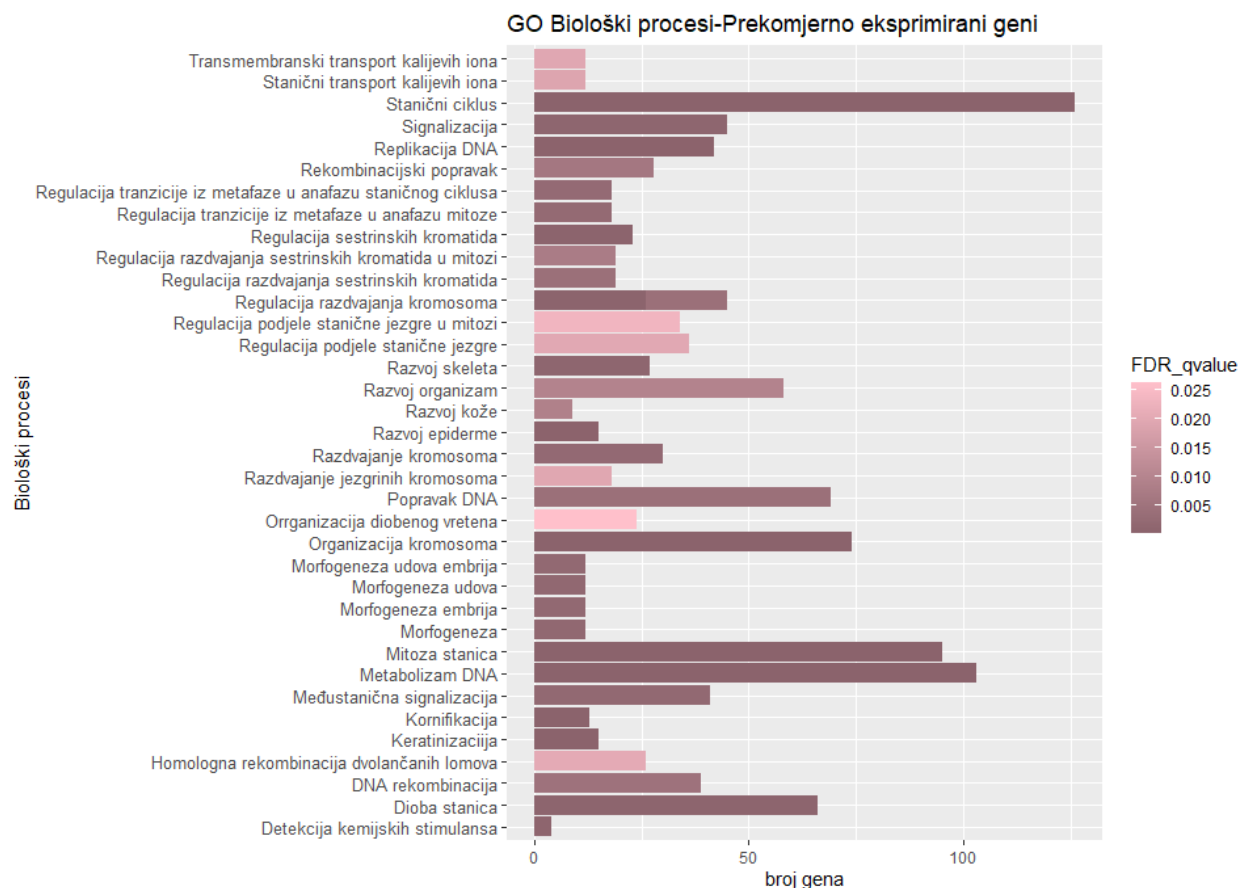
A)				B)			
Gene Symbol	Gene description	Log2FC	padj. value	Gene Symbol	Gene description	Log2FC	padj. value
MAGEA9B	MAGE Family Member A9B	22.82	1.7 x10 ⁻⁷	TUBB1	Tubulin Beta 1 Class VI	-7.08	4.20 x10 ⁻⁵
STAG3L3	Stromal Antigen 3-Like 3	22.07	1.13 x10 ⁻¹²	CA1	Carbonic Anhydrase 1	-7.28	5.60 x10 ⁻⁵
CSAG1	Chondrosarcoma Associated Gene 1	21.68	2.35 x10 ⁻⁸	ELANE	Elastase, Neutrophil Expressed	-7.58	2.54 x10 ⁻⁴
MAGED4	MAGE Family Member D4	11.52	6.26 x10 ⁻⁹	BPI	Bactericidal permeability increasing protein	-7.72	9.43 x10 ⁻³
CST1	Cystatin SN	10.83	4.28 x10 ⁻⁵	PGLYRP1	Peptidoglycan Recognition Protein 1	-7.73	6.70 x10 ⁻⁵
CACNA1B	Calcium Voltage-Gated Channel Subunit Alpha1 B	10.75	1.09 x10 ⁻⁷	GP9	Glycoprotein IX Platelet	-7.76	3.07 x10 ⁻²
MAGEA6	MAGE Family Member A6	9.75	2.26 x10 ⁻³	DEFA4	Defensin Alpha 4	-8.04	1.36 x10 ⁻²
PRAME	Protein Preferentially Expressed Antigen in Melanoma	9.58	2.13 x10 ⁻⁸	APOA2	Apolipoprotein A2	-8.16	7.56 x10 ⁻³
POU6F2	POU Class 6 Homeobox 2	9.53	3.00 x10 ⁻⁶	CELA3B	Chymotrypsin Like Elastase 3B	-8.23	3.40 x10 ⁻²
ZIC2	Zic Family Member 2	9.37	4.54 x10 ⁻⁷	ALB	Albumin	-8.25	1.00 x10 ⁻⁶
B4GALNT4	Beta-1,4-N-Acetyl-Galactosaminyltransferase4	9.3	1.00 x10 ⁻⁶	CYP1A1	Cytochrome P450 Family 1 Subfamily A Member 1	-8.32	3.56 x10 ⁻⁴
CALML3	Calmodulin Like 3	9.21	2.10 x10 ⁻⁵	MYL2	Myosin Light Chain 2	-8.35	1.11 x10 ⁻²
C12orf56	Chromosome 12 Open Reading Frame 56	9.07	3.00 x10 ⁻⁶	DEFA3	Defensin Alpha 3	-8.98	8.03 x10 ⁻³
MAGEA11	MAGE Family Member A11	8.93	3.73 x10 ⁻³	CELA3A	Chymotrypsin Like Elastase 3A	-8.99	4.71 x10 ⁻³
MAGEA4	MAGE Family Member A4	8.87	2.31 x10 ⁻³	MS4A3	Membrane Spanning 4-Domains A3	-9.03	3.16 x10 ⁻⁸
VPS9D1-AS1	VPS9D1 Antisense RNA 1	8.77	4.49 x10 ⁻⁷	MMP8	Matrix Metalloproteinase 8	-9.08	1.00 x10 ⁻⁶
SOX21	SRY-Box Transcription Factor 21	8.69	1.07 x10 ⁻⁴	S100A12	S100 calcium-binding protein A12	-9.21	3.00 x10 ⁻⁶
SLC24A2	Solute Carrier Family 24 Member 2	8.64	1.16 x10 ⁻⁴	DEFA1	Defensin Alpha 1	-9.89	1.06 x10 ⁻³
HOXD11	Homeobox D11	8.56	2.24 x10 ⁻³	CEACAM8	Carcinoembryonic antigen-related cell adhesion molecule 8	-9.9	7.17 x10 ⁻³
DUSP9	Dual Specificity Phosphatase 9	8.54	2.07 x10 ⁻⁴	DEFA1B	Defensin Alpha 1B	-14.21	2.94 x10 ⁻⁷

4.1.4. Rezultati analize genske ontologije i analize GSEA

Kako bi dobili uvid u biološku značajnost promjene razine ekspresije gena, i kako bi genima sa promijenjenom razinom ekspresije odredili ulogu u biološkim procesima, molekularnim funkcijama ili staničnim komponentama, dodijelili smo im anotacije genske ontologije (GO). Analiza genske ontologije napravljena je pomoću mrežnog alata *Gorilla*, zasebno za 961 gen sa pozitivno promijenjenom ekspresijom i zasebno za 574 gena sa negativno promijenjenom ekspresijom u tumorskim uzorcima. Analiza je napravljena za sve tri GO kategorije: kategoriju Biološki procesi, kategoriju Stanične komponente i kategoriju Molekularne funkcije. Prekomjerno zastupljeni GO pojmovi sa statističkom značajnošću korigirane p vrijednosti (FDR) manje od 0,05, prema metodi Benjamini-Hochberg, smatrani su značajnima.

Rezultati analize genske ontologije za gene sa pozitivnom promjenom u razini ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole, pokazuju da je u kategoriji Biološki procesi (Slika 19.) značajno obogaćeno 37 termina genske ontologije sa FDR<0,05. Pet najviše rangiranih obogaćenih termina genske ontologije, prema FDR vrijednostima, koji su povezani sa biološkim procesima su: Procesi staničnog ciklusa (GO:0022402, FDR=5,59x10⁻⁶), Keratinizacija (GO:0031424, FDR=4,9x10⁻⁶), Procesi povezani sa mitozom (GO:1903047, FDR=4,6x10⁻⁶), Metabolički procesi koji uključuju DNA (GO:0006259, FDR=3,72x10⁻⁵) i Organizacija kromosoma (GO:0051276, FDR=4,94x10⁻⁵). U GO termin Procesi staničnog ciklusa uključeno je

najviše pozitivno promijenjenih gena, njih 126. Među genima koji sudjeluju u regulaciji staničnog ciklusa je i STAG3L3 (engl. *Stromal Antigen 3-Like 3*), gen koji je u našim analizama pokazao drugu najveću promjenu razine ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole ($\log_2FC=22,07$).



Slika 19. Analiza genske ontologije za gene sa pozitivno promijenjenom ekspresijom u kategoriji Biološki procesi. Na x-osi prikazan je broj gena sa pozitivno promijenjenom ekspresijom koji su obogaćeni u određenom GO pojmu, prikazanim na y-osi. Boja označava prilagođenu statističku značajnost obogaćivanja GO pojmova.

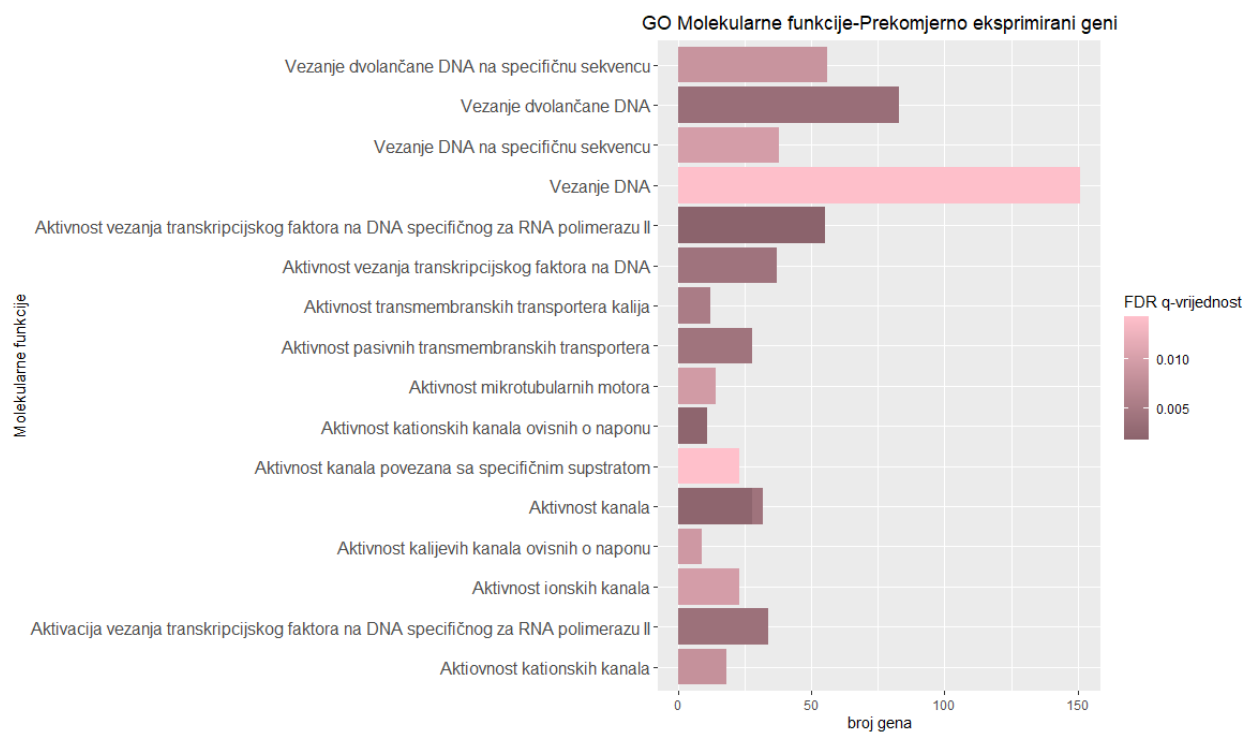
Kao što je vidljivo na Slici 19., osim u regulaciji staničnog ciklusa, najviše gena sa pozitivno promijenjenom razinom ekspresije ($n=103$) sudjeluje u metaboličkim procesima koji uključuju DNA (GO:0006259), 95 gena sudjeluje u regulaciji mitoze (GO:1903047), 74 gena utječe na organizaciju kromosoma (GO:0051276), a 69 gena ima ulogu u popravku DNA (GO:0006281). Svi geni koji se ubrajaju u termin Procesi povezani sa mitozom, zajednički su i terminu Procesi staničnog ciklusa, a veliki broj zajedničkih gena može se pronaći i u terminima Metabolički procesi

koji uključuju DNA i terminu Organizacija kromosoma. 17 gena zajedničko je svim navedenim terminima (Tablica 15.), uključujući i četiri pripadnika MCM obitelji gena (engl. (*Minichromosome Maintenance Complex Component*) MCM2, MCM4, MCM7 i MCM8 koji su esencijalni za inicijaciju procesa replikacije DNA. Među pet najviše rangiranih obogaćenih termina genske ontologije, povezanih sa biološkim procesima, je i termin Keratinizacija. Iako uključuje manji broj gena (n=13), svi geni uključeni u ovaj termin značajno su promijenjeni u tumorskim uzorcima u odnosu na kontrole, u rasponu $\log_2FC=4,64$ za Keratin 13 (KRT13) do $\log_2FC=7,84$ za Keratin 5 (KRT5). Nadalje, manji broj gena ima ulogu i u (među) staničnom transportu kalija (n=12, GO:0071805 i GO:0071804).

Tablica 15. Geni sa pozitivno promijenjenom ekspresijom koji se pojavljuju u svim pet najznačajnijim pojmovima genske ontologije u kategoriji Biološki procesi

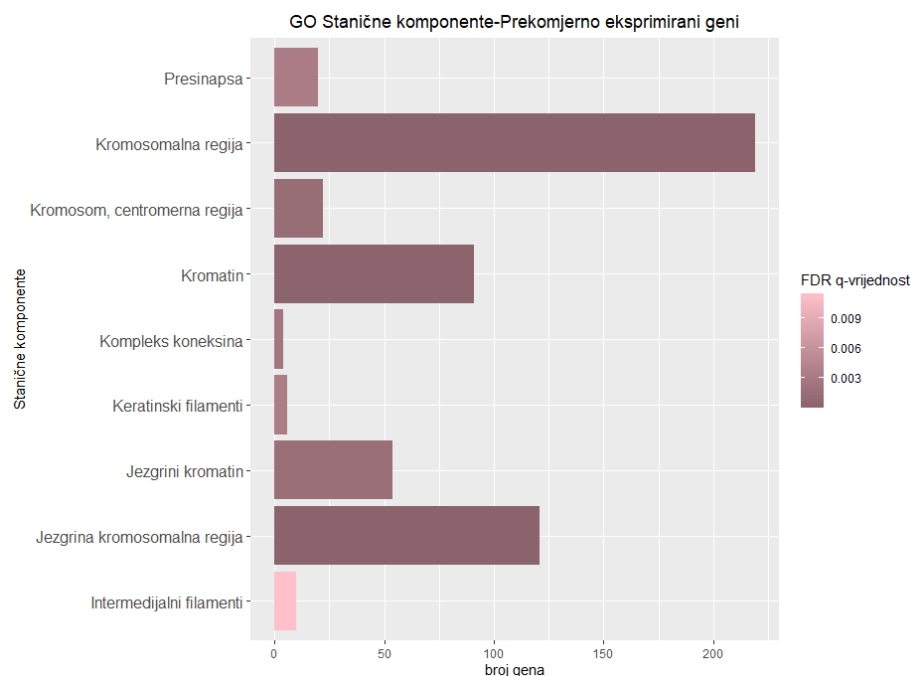
Zajednički geni GO:0022402, GO:0006259, GO:1903047, GO:0051276 i GO:0006281
FANCD2, BLM, CCNE1, CDCA5, CDK1, CDT1, CENPF, DSCC1, KIF22, MCM2, MCM4, MCM7, MCM8, PCNA, RAD51, TOP2A, TRIP13

U kategoriji Molekularne funkcije, većina gena sa pozitivno promijenjenom ekspresijom u tumoru, u odnosu na zdrave kontrole, ima ulogu u vezanju DNA (n=151, GO:0003677), vezanju DNA koje je ovisno o specifičnoj sekvenci (GO:0043565), vezanju na dvostruku uzvojnici DNA (GO:0003690) ili oboje (GO:1990837). Također, obogaćeni su GO pojmovi koji podrazumijevaju vezanje transkripcijskih faktora na DNA (GO:0003700, GO:0000981 i GO:0001228). Primjerice, geni ZIC2 (engl. *Zic Family Member 2*) i SOX21 (engl. *SRY-Box Transcription Factor 21*), koji su među 20 gena sa najvećom pozitivnom promijenom razine ekspresije, su transkripcijski faktori koji su obogaćeni u sva tri navedena GO pojma. Također, funkcija gena sa pozitivnom promijenom razine ekspresije povezuje se i sa aktivnošću kanala za transport (GO:0015267), ili aktivnošću transmembranskog transporta kalija (GO:0015079). Na Slici 20. prikazani su rezultati GO analize za kategoriju Molekularne funkcije.



Slika 20. Analiza genske ontologije za gene sa pozitivno promijenjenom ekspresijom u kategoriji Molekularne funkcije. Na x-osi prikazan je broj gena sa pozitivno promijenjenom ekspresijom koji su obogaćeni u određenom GO pojmu, prikazanim na y-osi. Boja označava prilagođenu statističku značajnost obogaćivanja GO pojmova.

Iz rezultata kategorije Stanične komponente (Slika 21.) vidljivo je da je većina gena u obogaćenim GO pojmovima vezana za područja staničnih komponenti (jezgreni) kromatin i kromosomi, što je u skladu sa rezultatima kategorije Biološki procesi u kojoj je najviše gena povezano sa regulacijom staničnog ciklusa.



Slika 21. Analiza genske ontologije za gene sa pozitivno promijenjenom ekspresijom u kategoriji Stanične komponente. Na x-osi prikazan je broj gena sa pozitivno promijenjenom ekspresijom koji su obogaćeni u određenom GO pojmu, prikazanim na y-osi. Boja označava prilagođenu statističku značajnost obogaćivanja GO pojmova.

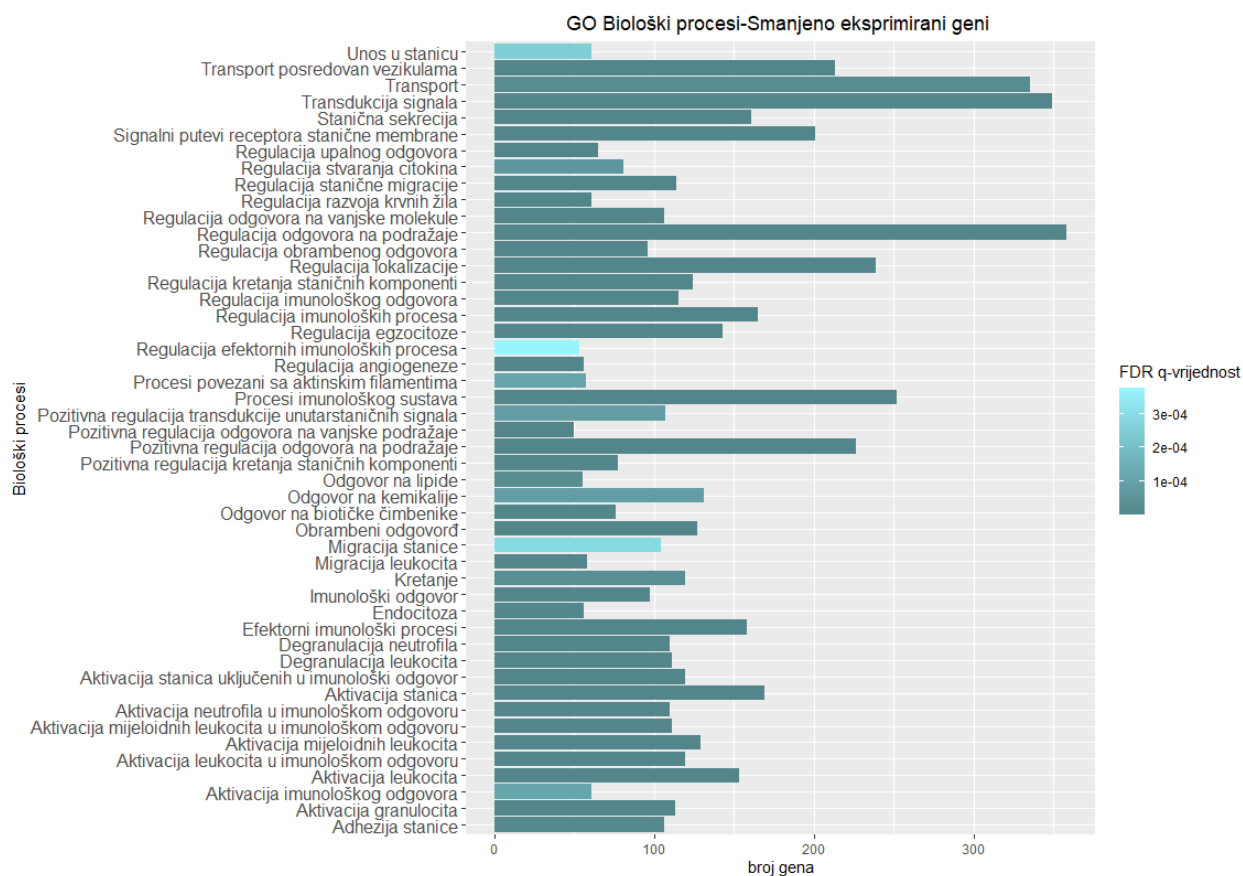
Rezultati analize genske ontologije za gene sa negativnom promjenom u razini ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole, pokazuju da je u kategoriji Biološki procesi (Slika 22.) značajno obogaćen 61 termin genske ontologije sa $FDR < 0,05$. Pet najviše rangiranih obogaćenih termina genske ontologije, prema FDR vrijednostima, koji su povezani sa biološkim procesima su: Regulacija egzocitoze (GO:0045055, $FDR = 4,33 \times 10^{-22}$), Aktivacija mijeloidnih leukocita (GO:0002274, $FDR = 1,61 \times 10^{-20}$), Stanična sekrecija (GO:0032940, $FDR = 5,15 \times 10^{-19}$), Aktivacija stanica (GO:0001775, $FDR = 1,42 \times 10^{-17}$) i Efektorni imunološki procesi (GO:0002252, $FDR = 2,08 \times 10^{-17}$). Većina gena, među pet najviše rangiranih obogaćenih termina genske ontologije, je zajedničko svim navedenim terminima ($n=110$) (Tablica 16.). Među 110 zajedničkih gena najviše obogaćenih GO termina, nalazi se i 7 gena na najvećem negativnom promijenom ekspresije u tumorskim uzorcima: BPI (engl. *Bactericidal permeability-increasing protein*), MS4A3 (engl. *membrane-spanning 4-domains, subfamily A, member 3*), MMP8 (engl. *matrix metalloproteinase-8*), S100A12 (engl. *Calgranulin C*), DEFA1 (engl. *Defensin Alpha 1*), DEFA1B (engl. *Defensin Alpha 1B*) i CEACAM8 (engl. *CEA Cell Adhesion Molecule 8*). BPI ima posredničku ulogu u fagocitozi i neutralizaciji bakterijskog upalnog odgovora. MMP8, osim što

ima ulogu u degradaciji vanstaničnog matriksa, posreduje u medijaciji upale preko neutrofila [109]. Lučenjem iz citosola neutrofila u vanstanični prostor, S100A12 može posredovati u urođenom imunosnom odgovoru. Primjerice, vezanjem S100A12 na AGER (engl. *Advanced Glycosylation End-Product Specific Receptor*) aktiviraju se nizvodni putevi koji dovode do aktivacije *NF-κB* signalnog puta i lučenja proupalnih citokina [110]. CEACAM8 je marker granulocita, dok DEFA1/1B također imaju ulogu u posredovanju urođenog imunološkog odgovora [111].

Tablica 16. Geni sa negativno promijenjenom ekspresijom koji se pojavljuju u svih pet najznačajnijih pojmova genske ontologije u kategoriji Biološki procesi

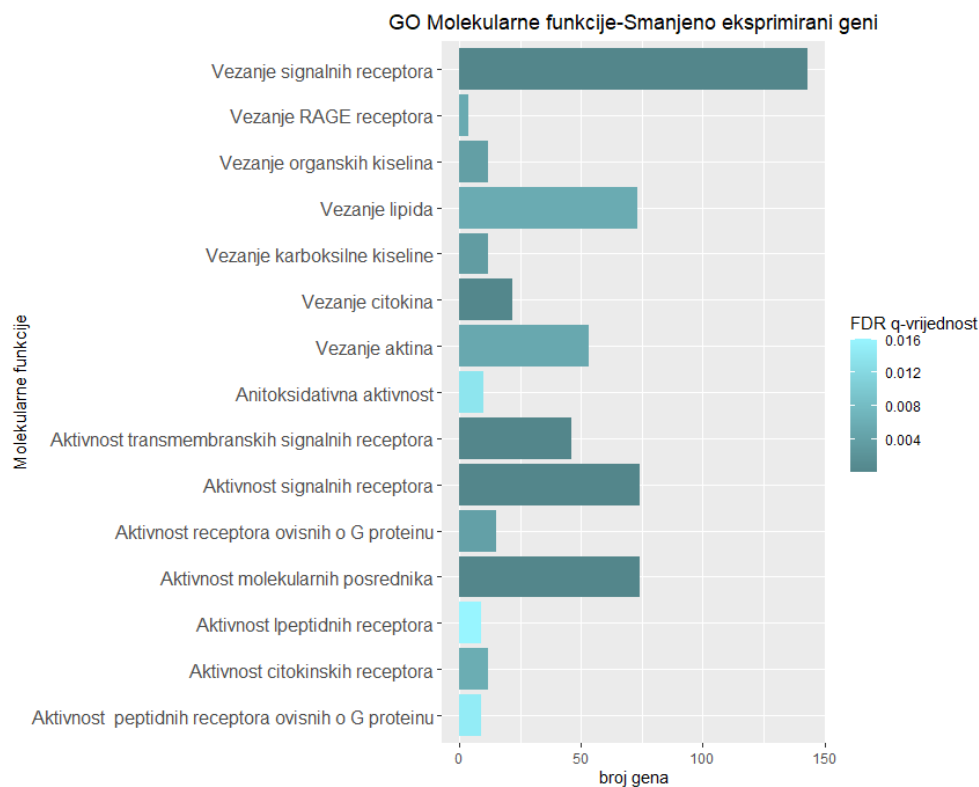
Zajednički geni GO:0045055, GO:0002274, GO:0032940, GO:0001775 i GO:0002252
ALOX5, ANO6, ANXA3, ARPC5, ASAH1, ATP8B4, BIN2, BPI, BRI3, BST1, C19orf59, C5AR1, CAB39, CAT, CD177, CD300A, CD36, CD47, CD53, CD59, CD63, CD68, CD93, CDA, CEACAM3, CEACAM6, CEACAM8, CFD, CFP, CLEC4D, CLU, COTL1, CPPED1, CR1, CST3, CTSS, CXCR1, CXCR2, CYBB, CYSTM1, DEFA1, DEFA1B, EEF1A1, ENPP4, ERP44, FCAR, FCER1G, FCGR2A, FCGR3B, FCN1, FGL2, FGR, FOLR3, FPR1, FPR2, FTH1, FTL, GCA, GMFG, HK3, HSPA1B, IQGAP2, ITGAM, ITGAX, ITGB2, LILRB3, LTA4H, LYZ, METTL7A, MGAM, MME, MMP8, MNDA, MPO, MS4A3, OLR1, PECAM1, PLEKHO2, PPBP, PTPRB, PTPRC, PTX3, QPCT, RAB27A, RAP1A, RETN, RHOA, RNASE2, ROCK1, S100A12, S100A8, S100A9, SDCBP, SELL, SERPINA1, SERPINA3, SERPINB1, SIGLEC14, SIRPB1, SLC11A1, SLC2A3, SLPI, SNAP23, STOM, TLR2, TYROBP, UNC13D, VAPA, VNN1, HSPA1A

Osim 5 najviše rangiranih obogaćenih termina genske ontologije, i većina ostalih GO termina su povezana sa regulacijom imunološkog odgovora. Geni sa smanjenom razinom ekspresije primjerice sudjeluju u procesima poput aktivacije stanica uključenih u imunološki odgovor (GO:0002263), aktivacije (GO:0045321) i degranulacije leukocita (GO:0043299), aktivacije (GO:0002283) i degranulacije neutrofila (GO:0043312), aktivacije granulocita (GO:0036230) i regulacije lučenja citokina (GO:0001817) (Slika 22.).



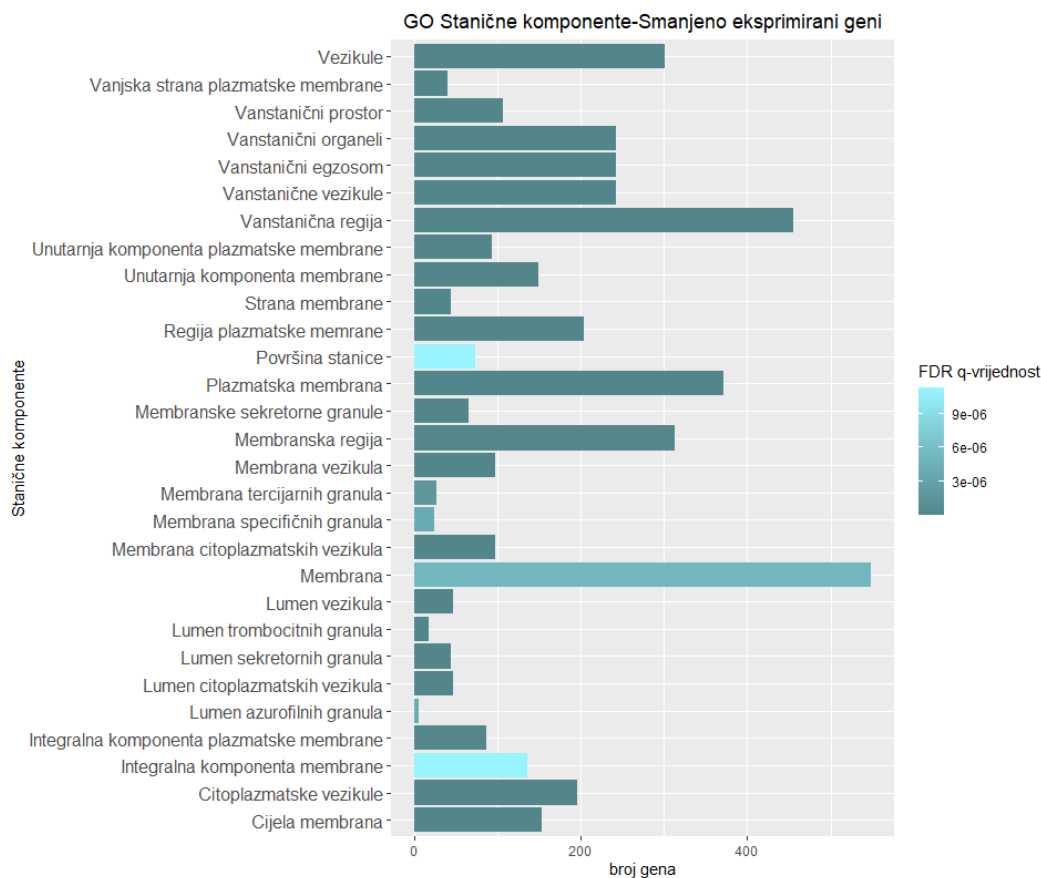
Slika 22. Rezultati analize genske ekspresije za gene sa negativnom promijenom razine ekspresije u tumoru u odnosu na zdrave kontrole za kategoriju Biološki procesi.

Rezultati analize genske ontologije za kategoriju Molekularne funkcije ukazuju da geni sa negativnom promijenom ekspresije u tumorskim uzorcima imaju funkcije u vezanju (GO:0005102) i aktivaciji signalnih receptora (GO:0038023), vezanju citokina (GO:0019955) i aktivaciji njihovih receptora (GO:0004896) te vezanju aktina (GO:0003779) i lipida (GO:0008289) (Slika 23.). Obzirom da je većina GO termina u kategoriji Biološki procesi povezano sa regulacijom imunološkog odgovora, rezultati za kategoriju Molekularne funkcije su očekivani.



Slika 23. Rezultati analize genske ekspresije za gene sa negativnom promijenom razine ekspresije u tumoru u odnosu na zdrave kontrole za kategoriju Molekularne funkcije.

Rezultati analize genske ontologije za kategoriju Stanične komponente ukazuju da najveći broj gena sa smanjenom razinom ekspresije u tumoru svoju aktivnost vrše u području membrana, području međustaničnog matriksa, u vezikulama i izvanstaničnim eksosomima (Slika 24.).



Slika 24. Rezultati analize genske ekspresije za gene sa negativnom promijenom razine ekspresije u tumoru u odnosu na zdrave kontrole za kategoriju Stanične komponente. Prikazano je 30 najznačajnijih rezultata

Kako bi dobili bolji uvid u kojim staničnim putevima i procesima geni sa promijenjenom razinom ekspresije sudjeluju, osim analize genske ontologije, napravljena je i analiza obogaćivanja setovima gena (GSEA). Rezultati GSEA analize pokazuju obogaćenje za 325 staničnih puteva, sa postavljenom graničnom vrijednosti $FDR < 0,25$. Od 325 obogaćenih staničnih puteva, 156 staničnih puteva pokazuje povišenu regulaciju, a 169 staničnih puteva pokazuje sniženu regulaciju u tumorskim uzorcima. Na Slici 25. prikazano je 40 najznačajnijih staničnih puteva, 20 sa povišenom i 20 sa sniženom regulacijom u tumorskim uzorcima. 5 najznačajnijih staničnih puteva s povišenom regulacijom uključuju Keratinizaciju (R-HSA-6805567, $FDR < 1 \times 10^{-150}$), Stvaranje kornificirane omotnice (R-HSA-6809371, $FDR < 1 \times 10^{-150}$), Elongacija DNA lanca (R-HSA-69190, $FDR = 7 \times 10^{-4}$), Aktivacija ATR-a u odgovoru na replikacijski stres (R-HSA-176187, $FDR = 5,25 \times 10^{-4}$) i NP63 delta stanični put (PMID: 18832364, $FDR = 4,2 \times 10^{-4}$). Kao što je vidljivo

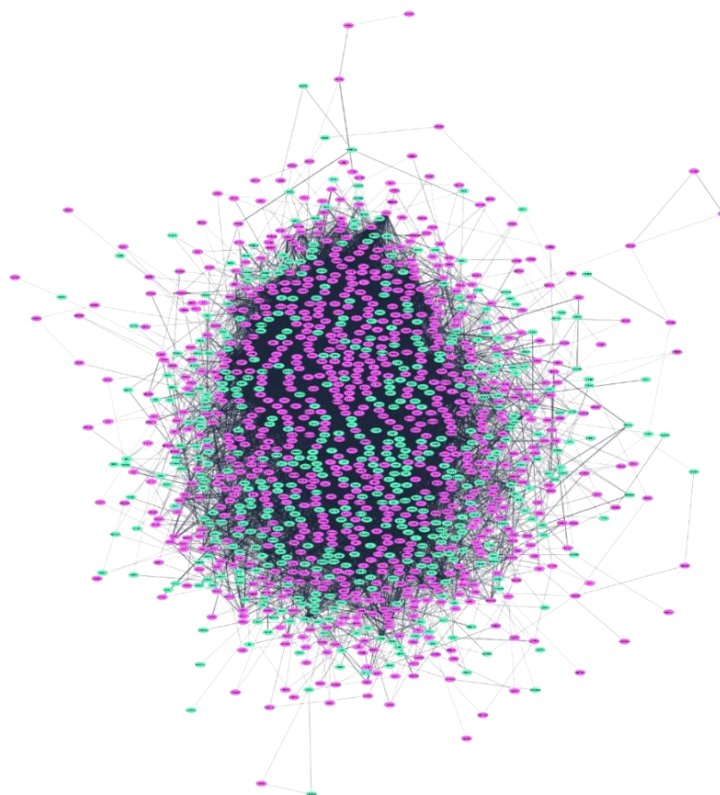
na Slici 25., osim Elongacije DNA lanca i Aktivacije ATR-a u odgovoru na replikacijski stres, većina prikazanih staničnih puteva je također povezana sa regulacijom staničnog ciklusa. Primjerice, stanični putevi AURORA B, FOXM1 i PLK1 se povezuju sa regulacijom staničnog ciklusa. 5 najznačajnijih staničnih puteva sa sniženom regulacijom uključuju Degranulaciju neutrofila (R-HSA-6798695, $FDR < 1 \times 10^{-150}$), Ribosom (hsa03010, $FDR < 1 \times 10^{-150}$), Odgovor na povišenu razinu Ca^{2+} u citosolu trombocita (R-HSA-76005, $FDR < 1 \times 10^{-150}$), NMD (engl. *Nonsense-mediated decay*) neovisan o eksonskom spojnem kompleksu (R-HSA-975956, $FDR < 1 \times 10^{-150}$) i Antimikrobni peptidi (R-HSA-6803157, $FDR < 1 \times 10^{-150}$). Kao što je vidljivo na Slici 24. nekoliko staničnih puteva je povezano sa regulacijom imunološkog odgovora, poput primjerice CXCR2, TRAF6 posredovana aktivacija *NF- κ B* i Regulacija TLR (engl. *Toll-like receptors*) endogenim ligandima.



Slika 25. Prikaz 20 najznačajnijih staničnih puteva s povišenom i 20 najznačajnijih puteva s sniženom regulacijom u tumorskim uzorcima.

4.1.5. Rezultati analize mreže proteinskih interakcija, analize modula i identifikacije ključnih gena

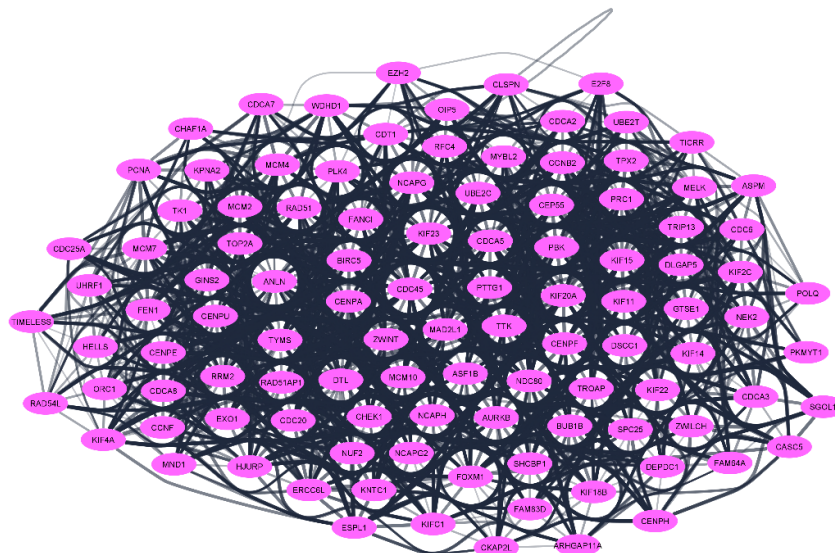
Kako bi istražili potencijalne interakcije između gena sa promijenjenom ekspresijom, napravili smo analizu mreže potencijalnih interakcija gena na razini proteina (engl. *Protein-protein interaction network, PPI*). U mrežni alat *Search Tool for the Retrieval of Interaction Genes* (STRING, verzija 11.0) uneseni su geni sa promijenjenom razinom ekspresije u tumorskom tkivu i odabrana je vrijednost pouzdanosti interakcije (engl. *confidence score*) $\geq 0,4$. Mrežni model, napravljen u STRING alatu, je potom vizualiziran pomoću Cytoscape programa (verzija 3.8.2.). Vizualizirana mreža proteinskih interakcija sadrži 1470 čvorišta (engl. *nodes*) i 13 575 linija (engl. *edges*) koje predstavljaju interakcije između gena. Za vizualizaciju je odabran stil (engl. *layout*) *yFiles Organic* i *yFiles Remove Overlaps*. 172 gena uklonjeno je iz vizualizacije jer za njih nije pokazana interakcija sa ostalim diferencijalno eksprimiranim genima te tako konačna mreža interakcija sadrži 1298 čvorišta i 13 575 linija (Slika 26.).



Slika 26. Mreža predviđene interakcije proteina na temelju razine genske ekspresije. Tirkiznom bojom označeni su geni sa smanjenom ekspresijom u tumorskom tkivu, a ljubičastom bojom geni sa povećanom razinom ekspresije.

Nakon vizualizacije PPI mreže u *Cytoscape* programu, analizirana su njezina svojstva: postojanje visoko umreženih regija, odnosno klastera unutar PPI mreže, i postojanje visoko umreženih proteina (engl. *hub genes*).

Za identifikaciju klastera, regija visoko umreženih proteina unutar PPI mreže, koristili smo aplikaciju MCODE (skraćeno iz engl. *Molecular Complex Detection*) u Cytoscape programu. Granične vrijednosti za pronalazak klastera postavili smo na slijedeći način: stupanj povezanosti čvorišta=2; vrijednost čvorišta (engl. *node score*) =0.02; K-core vrijednost=2; maksimalna dubina=100 i MCODE vrijednost ≥ 5 . Na temelju postavljenih parametara identificirano je 11 regija visoko umreženih gena/proteina. Najznačajniji identificirani klaster, Klaster 1 (MCODE vrijednost 89,25), sadrži 106 čvorišta povezanih sa 4686 linija i ujedno je i najveći od identificiranih klastera. Kao što je vidljivo na Slici 27. svi geni u ovom klasteru su geni sa pozitivnom promjenom razine ekspresije u tumorskim uzorcima. Kao što je vidljivo iz Tablice 17., 6 od 8 obogaćenih staničnih puteva u Klasteru 1 povezano je sa kontrolom staničnog ciklusa, sa najznačajnije obogaćenim signalnim putem Kontrolna točka iz G2 u M fazu staničnog ciklusa ($p=5,05 \times 10^{-54}$), u kojem sudjeluje 41 gen sa pozitivno promijenjenom ekspresijom u tumorskim uzorcima. Obzirom da rezultati analize obogaćivanja setovima gena GSEA (iz odjeljka 4.1.4.), također ukazuju na veliki broj obogaćenih staničnih puteva povezanih sa regulacijom staničnog ciklusa, ne iznenađuje da su se geni povezani sa regulacijom staničnog ciklusa grupirali u najznačajniji klaster, Klaster 1.



Slika 27. Prikaz gena iz Klastera 1 identificiranih pomoću MCODE aplikacije u Cytoscape programu. Ljubičasta boja ukazuje na gene sa pozitivno promijenjenom razinom ekspresije u tumorskim uzorcima

Tablica 17. Prikaz obogaćenih staničnih puteva za Klaster 1.

	Stanični putevi	Prilagođena P vrijednost	Vrijednost obogaćivanja	Geni
Klaster 1	Kontrolna točka iz G2 u M fazu staničnog ciklusa	$5,03 \times 10^{-54}$	9825,19	TOP2A, TROAP, CCNF, TTK, KIF11, CENPA, AURKB, KIF15, CDC20, CCNB2, CHAF1A, CDC45, PTTG1, EXO1, CHEK1, PBK, RAD54L, NEK2, MYBL2, KPNA2, PLK4, POLQ, GINS2, UBE2C, KIF23, KNL1, CDC6, KIF22, CDC25A, NDC80, TPX2, CENPE, CENPF, ESPL1, PRC1, KIF4A, BIRC5, KIF2C, EZH2, MCM2, MAD2L1
	Mete E2F transkripcijskih faktora	$1,93 \times 10^{-52}$	9056,42	TOP2A, PCNA, CDCA3, MCM7, DSCC1, BUB1B, CDCA8, AURKB, CDC20, RAD51AP1, CCNB2, PTTG1, CHEK1, MYBL2, TK1, KPNA2, DLGAP5, E2F8, PLK4, HELLS, RRM2, KIF22, CDC25A, CENPE, MELK, KIF18B, ESPL1, UBE2T, KIF4A, TIMELESS, DEPDC1, BIRC5, MCM4, KIF2C, TRIP13, ASF1B, EZH2, MCM2, MAD2L1, SPC25
	Diobeno (mitotičko) vreteno	$1,69 \times 10^{-19}$	1150,09	TOP2A, KIF23, TTK, KIF11, KIF22, NDC80, KIF15, CCNB2, TPX2, ANLN, CENPE, CENPF, ESPL1, PRC1, KIF4A, BIRC5, KNTC1, NEK2, KIF2C, DLGAP5
	Mete Myc obitelji gena	$4,42 \times 10^{-7}$	173,03	CDC20, RFC4, PCNA, CDC45, MCM7, MCM4, KPNA2, TYMS, MCM2, MAD2L1
	Spermatogeneza	$2,44 \times 10^{-5}$	128,65	CCNB2, RFC4, NEK2, KIF2C, TTK, NCAPH, EZH2
	DNA popravak	$3,95 \times 10^{-4}$	72,55	FEN1, RFC4, PCNA, RAD51, TYMS, ZWINT
	Kasni odgovor na estrogen	0,0015	44,34	CDC20, PLK4, TOP2A, GINS2, KIF20A, CDC6
	mTORC1 signalni put	0,0085	27,29	RRM2, CCNF, MCM4, CDC25A, MCM2

Klaster 2 (MCODE vrijednost 16,75) sadrži 49 gena i 402 njihove potencijalne interakcije (Slika 28.). 17 gena u ovom klasteru ima pozitivno, a 32 gena imaju negativno promijenjenu razinu ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole. Prema analizi obogaćivanja setovima gena, geni koji su se grupirali u ovaj klaster najznačajnije su povezani sa regulacijom upalnog odgovora ($p=0,0026$). Također, u ovom klasteru je većina staničnih puteva povezana sa regulacijom imunološkog odgovora, što je vidljivo iz Tablice 18.

Tablica 19. Prikaz obogaćenih staničnih puteva za Klustere 3-11

Klaster	Stanični putevi	Prilagođena P vrijednost	Vrijednost obogaćivanja	Geni
Klaster 3 (74 gena i 413 interakcija)	Koagulacija	$6,59 \times 10^{-8}$	424,99	CFD, F8, APOA1, SERPING1, LTA4H, MMP8, A2M, CLU, THBS1 (9/138)
	Sustav komplementa	$9,81 \times 10^{-5}$	123,88	CDA, F8, GCA, SERPING1, LTA4H, MMP8, CLU (7/200)
	Epitelno mezenhimalna tranzicija	$9,81 \times 10^{-5}$	123,88	IGFBP3, TNC, VEGFC, PTX3, THY1, THBS2, THBS1 (7/200)
	KRAS Signalni put	0,007	51,61	SLP1, IGFBP3, SPARCL1, RETN, GNG11 (5/200)
	Miogeneza	0,0295	28,99	CFD, ITGB4, IGFBP3, CLU (4/200)
	Apikalna spojnica	0,0295	28,99	NEGR1, ITGB4, ITGA10, THY1 (4/200)
	Odbacivanje presatka	0,0295	28,99	SRGN, THY1, CFP, ELANE (4/200)
	Apikalna površina	0,0462	58,65	SRGN, THY1, CFP, ELANE (4/200)
Klaster 4 (7 gena i 20 interakcija)	Mete Myc obitelji gena	0,0201	227,67	RRP12 (1/58)
	Androgeni odgovor	0,0344	112,77	RRP12 (1/100)
Klaster 5 (58 gena i 183 interakcija)	Epitelno mezenhimalna tranzicija	$9,41 \times 10^{-12}$	780,66	POSTN, COL1A2, PCOLCE2, COL5A1, MMP1, ITGB3, COL11A1, WNT5A, SPP1, MXRA5, DCN, PTHLH (12/200)
	Upalni odgovor	$1,93 \times 10^{-5}$	186,11	MSR1, PTGIR, CSF3R, CALCL, ITGB3, AQP9, CD48 (7/200)
	Angiogeneza	0,0011	288,79	POSTN, SPP1, LPL (13/36)
	Sustav komplementa	0,0015	78,03	FCN1, SERPINA1, MMP13, S100A12, S100A9 (5/200)
	Koagulacija	0,0030	79,70	FGA, SERPINA1, MMP1, ITGB3 (4/138)
	Geni sa smanjenom ekspresijom kao odgovor na UV zračenje	0,0309	36,56	COL1A2, ITGB3, COL11A1 (3/144)
Klaster 6 (23 gena i 61 interakcija)	Epitelno mezenhimalna tranzicija	$1,14 \times 10^{-6}$	587,57	LAMA2, LAMA1, COL7A1, COL5A2, TIMP3, SDC1 (6/200)
	Rani odgovor na estrogen	0,0105	97,87	KRT19, KRT15, KRT13 (3/200)
Klaster 7 (32 gena i 85 interakcija)	Kasni odgovor na estrogen	0,0018	117,88	TRIM29, PERP, CAV1, PKP3 (4/200)
	Odbacivanje presatka	0,0018	117,88	CD79A, SPI1, IGSF6, IRF4 (4/200)
Klaster 8 (9 gena i 20 interakcija)	Mete E2F transkripcijskih faktora	$4,70 \times 10^{-6}$	1103,16	GINS1, POLA2, RNASEH2A, NCAPD2 (4/200)
Klaster 9 (5 gena i 10 interakcija)	Kasni odgovor na estrogen	0,0490	74,99	XRCC3 (1/200)
Klaster 10 (5 gena i 10 interakcija)	Nema obogaćenih staničnih puteva			

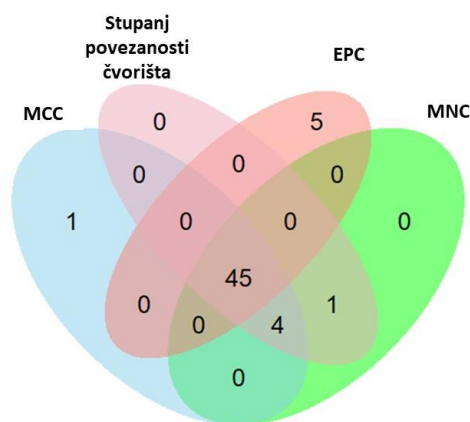
Tablica 19. Nastavak

Klaster 11 (5 gena i 10 interakcija)	Adipogeneza	0,0087	462,10	FABP4, PLIN2 (2/200)
	Peroksisom	0,0490	176,70	ACSL1 (1/104)
	Metabolizam žučne kiseline	0,0490	160,62	ACSL1 (1/112)
	Metabolizam masnih kiselina	0,0490	102,57	ACSL1 (1/158)
	IL-2/STAT5 Signalni put	0,0490	75,50	PLIN2 (1/199)
	Miogeneza	0,0490	74,99	ACSL1 (1/200)
	Metabolizam ksenobiotika	0,0490	74,99	PDK4 (1/200)
	Hipoksija	0,0490	74,99	PLIN2 (1/200)
	Oksidativna fosforilacija	0,0490	74,99	PDK4 (1/200)

Također smo identificirali izvorišne gene (čvorišta) sa najvećim MCODE vrijednostima iz kojih klasteri potječu (engl. *Seed genes*). Od 11 identificiranih klastera, izvorišni geni su identificirani u njih 10. TRIP13 (engl. *Thyroid Hormone Receptor Interactor 13*), čija je uloga zabilježena u regulaciji stanične proliferacije i invazije u nekoliko vrsta tumora, uključujući tumor pluća [112], izvorišni je gen za Klaster 1. Sličnu ulogu ima i gen S1PR1 (engl. *Sphingosine-1-phosphate receptor 1*) koji je izvorišni gen za Klaster 2 [113]. Nadalje izvorišni gen za Klaster 3 je GNG11 (engl. *Guanine Nucleotide-Binding Protein G(I)/G(S)/G(O) Gamma-11 Subunit*), za Klaster 4 DQX1 (engl. *DEAQ-Box RNA Dependent ATPase 1*), Klaster 5 S100A9 (engl. *S100 Calcium Binding Protein A9*), Klaster 6 COL6A6 (engl. *Collagen Type VI Alpha 6 Chain*), Klaster 7 POU2AF1 (engl. *POU Class 2 Homeobox Associating Factor 1*), Klaster 8 NCAPD2 (engl. *Non-SMC Condensin I Complex Subunit D2*), Klaster 9 SLX4 (engl. *SLX4 Structure-Specific Endonuclease Subunit*) i za Klaster 10 ADAMTS18 (engl. *ADAM Metallopeptidase With Thrombospondin Type 1 Motif 18*).

Za identifikaciju visoko umreženih gena unutar PPI mreže (engl. *hub genes*) koristili smo 4 algoritma: Maksimalna centralnost grupe (engl. *Maximal Clique Centrality*, MCC), Stupanj povezanosti čvorišta, Linijski filtrirana komponenta (engl. *Edge Percolated Component*, EPC) i Maksimalna komponenta susjeda (engl. *Maximum Neighborhood Component*, MNC), pomoću aplikacije *CytoHubba* u *Cytoscape* programu. Izračunate vrijednosti za sve diferencijalno eksprimirane gene pomoću navedenih algoritama su potom preklapljene, u svrhu identifikacije top 50 gena sa najvišim izračunatim vrijednostima. Kao što je vidljivo iz Slike 29. preklapom

izračunatih vrijednosti za sve gene svakog algoritma, pronađeno je 45 zajedničkih gena sa najvišim vrijednostima. Popis svih 45 gena sa izračunatim vrijednostima prikazan je u Tablici 20.



Slika 29. Preklap top 50 visoko umreženih gena prema izračunatim vrijednostima 4 algoritma i *CytoHubba* aplikacije. MCC-Maksimalna centralnost grupe; EPC-Linijski filtrirana komponenta; MNC-Maksimalna komponenta susjeda.

Tablica 20. Identificirani visokoumreženi geni unutar PPI mreže

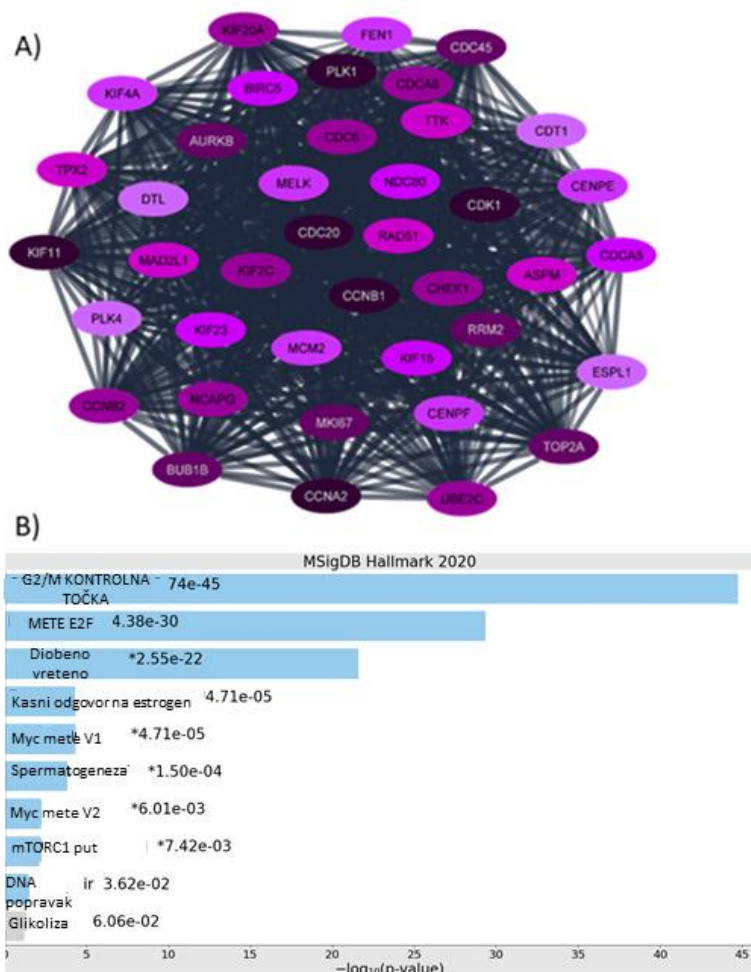
Simbol gena	Puni naziv gena (iz engl.)
CENPA	<i>Centromere protein A</i>
AURKB	<i>Aurora Kinase B</i>
CDCA8	<i>Cell Division Cycle Associated 8</i>
CHEK1	<i>Checkpoint Kinase 1</i>
DLGAP5	<i>Disks large-associated protein 5</i>
DTL	<i>Denticleless E3 Ubiquitin Protein Ligase Homolog</i>
KIF20A	<i>Kinesin Family Member 20A</i>
KIF23	<i>Kinesin Family Member 23</i>
MAD2L1	<i>Mitotic Spindle Assembly Checkpoint Protein MAD2A</i>
MCM10	<i>Minichromosome Maintenance 10 Replication Initiation Factor</i>
MCM2	<i>Minichromosome Maintenance Complex Component 2</i>
MKI67	<i>Marker Of Proliferation Ki-67</i>
PLK4	<i>Polo Like Kinase 4</i>
RAD51	<i>DNA Repair Protein RAD51 Homolog 1</i>
ASPM	<i>Assembly Factor For Spindle Microtubules</i>
BIRC5	<i>Baculoviral inhibitor of apoptosis repeat-containing 5</i>
BUB1B	<i>Mitotic Checkpoint Serine/Threonine-Protein Kinase BUB1 Beta</i>

Tablica 20. Nastavak

CCNA2	<i>Cyclin A2</i>
CDC20	<i>Cell Division Cycle 20</i>
CDC45	<i>Cell Division Cycle 45</i>
CDC6	<i>Cell Division Cycle 6</i>
CDCA5	<i>Cell Division Cycle Associated 5</i>
CDT1	<i>Chromatin Licensing And DNA Replication Factor 1</i>
CENPE	<i>Centromere Protein E</i>
CENPF	<i>Centromere Protein F</i>
ESPL1	<i>Extra Spindle Pole Bodies Like 1, Separase</i>
EXO1	<i>Exonuclease 1</i>
FEN1	<i>Flap Structure-Specific Endonuclease 1</i>
KIF11	<i>Kinesin Family Member 11</i>
KIF15	<i>Kinesin Family Member 15</i>
KIF2C	<i>Kinesin Family Member 2C</i>
MELK	<i>Maternal Embryonic Leucine Zipper Kinase</i>
NCAPG	<i>Non-SMC Condensin I Complex Subunit G</i>
NDC80	<i>NDC80 Kinetochore Complex Component</i>
PBK	<i>PDZ Binding Kinase</i>
RRM2	<i>Ribonucleotide Reductase Regulatory Subunit M2</i>
TOP2A	<i>DNA Topoisomerase II Alpha</i>
TTK	<i>TTK Protein Kinase</i>
UBE2C	<i>Ubiquitin Conjugating Enzyme E2 C</i>
CCNB1	<i>Cyclin B1</i>
CCNB2	<i>Cyclin B2</i>
CDK1	<i>Cyclin Dependent Kinase 1</i>
KIF4A	<i>Kinesin Family Member 4A</i>
PLK1	<i>Polo Like Kinase 1</i>
TPX2	<i>TPX2 Microtubule Nucleation Factor</i>

Kao što je vidljivo iz Tablice 20. 10 gena sa najvišim vrijednostima su CDK1, CDC20, CCNA2, CCNB1, PLK1, KIF11, CDC45, BUB1B i MKI67. Obzirom da ti geni imaju ulogu u regulaciji staničnog ciklusa, imaju potencijalno važnu ulogu u karcinogenezi tumora pluća skvamoznih stanica (Slika 30A).

Kako bi ispitali ulogu identificiranih visokoumreženih gena u staničnim procesima, napravljena je analiza obogaćenih staničnih puteva. Kao što je vidljivo na Slici 30 B), uloga *hub* gena je uglavnom vezana za regulaciju staničnog ciklusa, poput Kontrolna točka iz G2 u M fazu staničnog ciklusa, Mete E2F transkripcijskih faktora ili Diobeno (mitotičko) vreteno. Ovi stanični putevi su također obogaćeni u Klasteru 1, što nije iznenađujuće obzirom da većina identificiranih *hub* gena (35/45, 77%) također pripada identificiranom Klasteru 1.



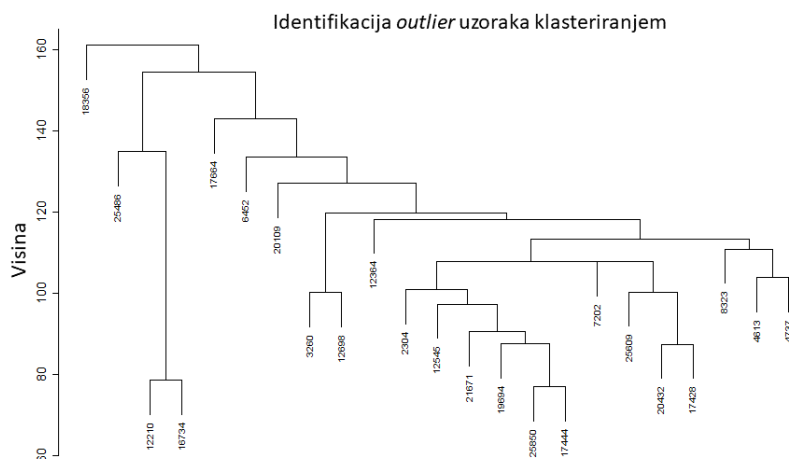
Slika 30. Identificirani visokoumreženi geni unutar PPI mreže. A) Prikaz interakcija *hub* gena. Boja čvorišta (gena) odgovara vrijednostima dobivenim pomoću *CytoHubba* aplikacije. Tamnija boja označava više rangirani gen, u odnosu na svjetliju boju. B) Rezultati analize obogaćenih staničnih puteva prikazani su stupčastim grafom. Statistički značajni stanični putevi prikazani su sa plavom bojom, sa istaknutom prilagođenom p-vrijednošću

4.1.6. Rezultati analize ponderirane korelacijske mreže (engl. *Weighted gene co-expression network analysis, WGCNA*)

WGCNA analiza koristi informacije o razini ekspresije svih gena, ne samo diferencijalno eksprimiranih, u svrhu identifikacije skupova gena od interesa i njihove povezanosti s fenotipovima. Pretvaranjem tisuća gena u desetine genskih skupova povezanih s fenotipovima, eliminiran je problem testiranja i ispravljanja više hipoteza. Zahvaljujući metodi i principima WGCNA, neki znanstvenici su pronašli potencijalne biomarkere za nekoliko različitih vrsta tumora, što ukazuje da ova metoda može biti koristan alat za pronalazak novih saznanja o genima i njihovoj povezanosti sa raznim kliničkim parametrima.

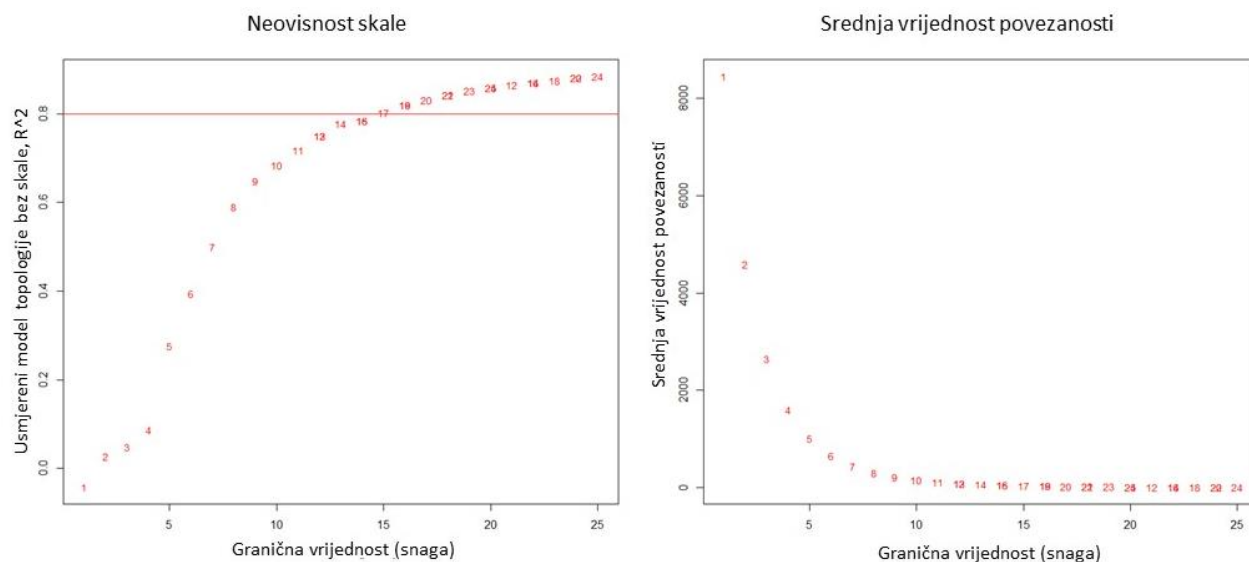
Kako bi identificirali klasterne koji sadrže međusobno jako korelirane gene, prema sličnoj razini ekspresije, i kako bi ih povezali sa kliničkim parametrima pacijenata oboljelih od tumora pluća skvamoznih stanica, koristili smo ponderiranu mrežnu analizu genske ko-ekspresije. Sama analiza sastoji se od nekoliko koraka koji podrazumijevaju konstrukciju mreže na temelju izračunate sličnosti profila genske ekspresije uzoraka u analizi i odabrane granične vrijednosti snage za potenciranje matrice sličnosti. Nakon identifikacije klastera ispituje se korelacija klastera i odabranih kliničkih parametara.

Prije same WGCNA analize provjereno je dali neki od 23 uzorka tumora pluća skvamoznih stanica odstupa profilom genske ekspresije od ostalih uzoraka u analizi (engl. *outlier sample*) i postoje li uzorci kojima nedostaje veći broj podataka. Kao što je vidljivo na dendogramu (Slika 31.), nije uočeno veće odstupanje u profilima genske ekspresije u tumorskim uzorcima, te je analiza nastavljena sa svim predviđenim uzorcima.



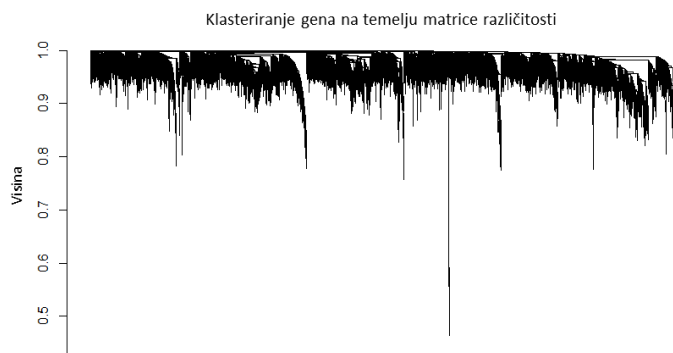
Slika 31. Detekcija *outlier* uzoraka u WGCNA analizi klasteriranjem.

Konstrukcija ponderirane mreže podrazumijeva odabir granične vrijednosti β (engl. *soft threshold power*) koja se koristi za potenciranje matrice sličnosti. Odabrana je granična vrijednost na kojoj bi konstruirana mreža bez skale pratila polinomsku distribuciju stupnja povezanosti čvorišta. Prema sugestijama autora WGCNA paketa, odabire se prva vrijednost na kojoj je $R^2 > 0,8$, što zadovoljava kriterije mreže bez skale. Za konstrukciju mreže u našem je slučaju odabrana granična vrijednost $\beta=18$ (Slika 32.)



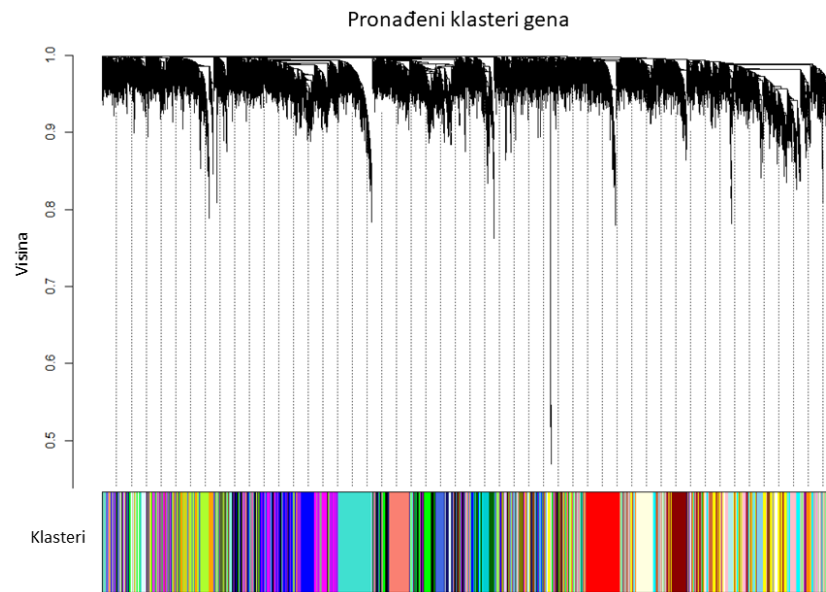
Slika 32. Odabir granične vrijednosti (β) za konstrukciju mreže bez skale. Na lijevom dijelu slike prikazan je usmjereni model topologije bez skale u ovisnosti o graničnoj vrijednosti β . Na desnom dijelu slike prikazana je ovisnost srednje vrijednosti povezanosti čvorišta u mreži i granične vrijednosti β .

Nakon konstrukcije mreže bez skale, identificirali smo klastere koji sadrže visoko korelirane gene, na temelju sličnosti u razini njihove ekspresije. Za identifikaciju klastera koji sadrže gene sa sličnom razinom ekspresije, koristili smo matricu različitosti (Slika 33.).



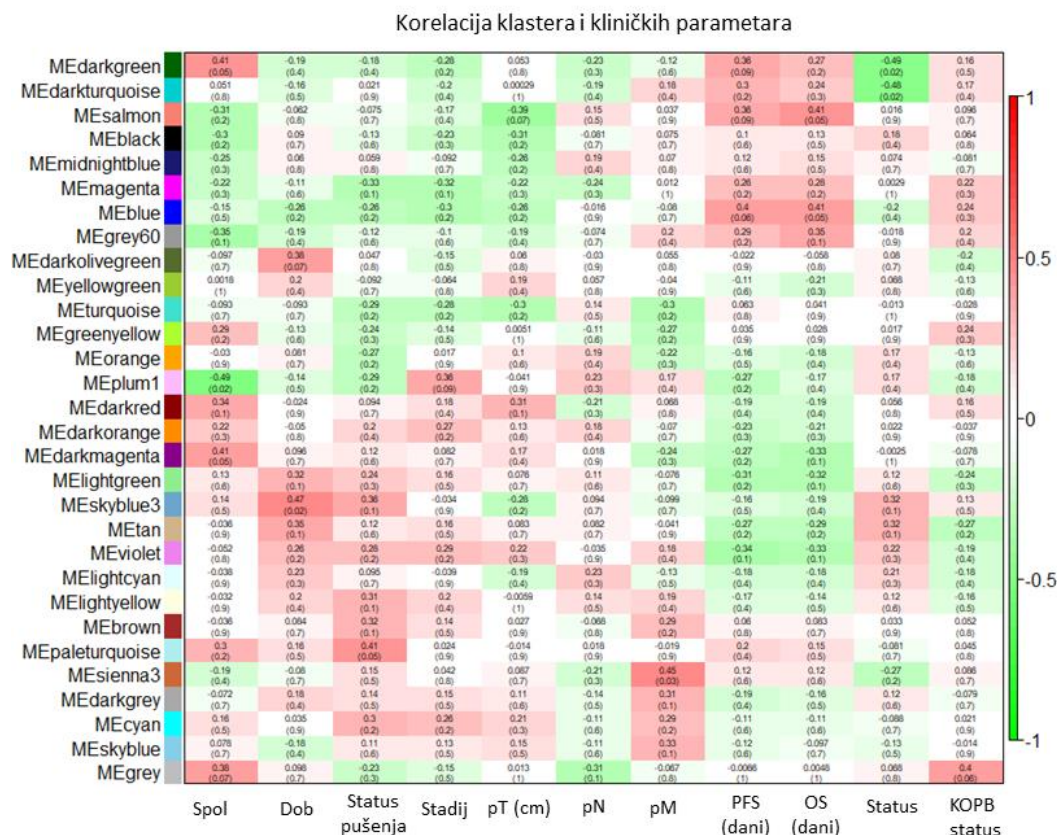
Slika 33. Identifikacija klastera klasteriranjem gena

U ovom koraku je identificirano ukupno 38 klastera. Kako bi smanjili broj klastera, spojili smo one klastere koji imaju visoki stupanj sličnosti sa drugim klasterima, na temelju vrijednosti korelacije $R^2 > 0,75$. Spajanjem visoko koreliranih klastera, dobili smo konačni broj klastera, njih 30. Svakom klasteru pripojene su različite boje kao oznake klastera. Na primjer, klasteru sa najvećim brojem gena dodijeljena je oznaka *Turquoise*, drugom po broju gena oznaka *Blue*, trećem *Brown*. Geni koji se nisu grupirali u klastere, objedinjeni su u klasteru *Grey*. Ukupno je 16 nedodijeljenih gena koji se nalaze u *Grey* klasteru (Slika 34.).



Slika 34. Identificirani klasteri gena WGCNA analizom. Dendrogram prikazuje klasterirane gene, a boja na liniji ispod dendograma prikazuje pripadnost gena određenom klasteru.

Potom su ispitane korelacije identificiranih klastera i kliničkih parametara pomoću korelacije *eigen* vrijednosti svakog klastera i kliničkih podataka. Klinički parametri koje smo korelirali sa klasterima su: spol ($\check{Z}=0$, $M=1$), dob (godine života), status pušenja (nepušač=0, aktivni pušač=1 i bivši pušač=2), stadij (1a/b, 2/a/b, 3a/b, 4a), veličina tumora izražena u cm (pT(cm)), zahvaćenost limfnih čvorova tumorom (pN; tumor nije zahvatio limfne čvorove=0, tumor je zahvatio obližnje limfne čvorove=1, tumor je zahvatio distalne limfne čvorove=2), prisutnost metastaza (pM, bez metastaza=0, prisutnost metastaza=1), preživljenje bez progresije bolesti (PFS (dani)), sveukupno preživljenje (OS (dani)), Status (živ=0, mrtav=1) i KOPB status (bez KOPB-a=0, sa KOPB-om=1).



Slika 35. Korelacija identificiranih klastera i kliničkih parametara pacijenata oboljelih od tumora pluća skvamoznih stanica. Rezultat korelacije prikazan je kao toplinska mapa na čijoj y osi su identificirani klasteri, a na x osi ispitivani klinički parametri. Boja u poljima toplinske mape ukazuje na jačinu ispitivane korelacije (crvena boja- jaka pozitivna korelacija, zelena-jaka negativna korelacija), čiji iznos je prikazan u poljima mape, zajedno sa statističkom značajnosti korelacije prikazanoj u zagradama. ME-Module eigengene

Kao što je vidljivo na Slici 35., rezultati korelacije klastera i kliničkih podataka pokazuju negativnu korelaciju *Plum* klastera sa spolom (korelacija=-0,49, $p=0,02$), dok je klaster *Skyblue* pozitivno koreliran sa dobi pacijenata (korelacija=0,47, $p=0,02$). Klasteri *Darkgreen* i *Darkturquoise* pokazuju negativnu korelaciju sa statusom pacijenta (korelacija=-0,49 i -0,48, $p=0,02$). *Sienna3* klaster pokazuje pozitivnu korelaciju sa prisutnošću metastaza kod pacijenata (korelacija=0,45, $p=0,03$). Nisu pronađene statistički značajne korelacije klastera i kliničkih parametara: status pušenja, stadij bolesti, veličina tumora, zahvaćenost limfnih čvorova tumorom, preživljenjem bez progresije bolesti, sveukupnim preživljenjem i KOPB statusom.

Za daljnje analize odabrali smo nama najzanimljiviji klaster, *sienna3*. Obzirom da je ovaj klaster pokazao korelaciju sa procesom metastaziranja, htjeli smo provjeriti ulogu gena u

navedenom klasteru, u kojem se grupiralo ukupno 115 gena. Geni su uvršteni u *PANTHER* bazu podataka gdje su pretraženi obogaćeni pojmovi genske ontologije. Popis svih gena koji su se grupirali u ovom klasteru prikazan je u Tablici 21.

Tablica 21. Geni grupirani u klasteru *sienna3*

Klaster	Geni
<i>Sienna3</i>	GIMAP7, BTN3A3, HMHA1, SH2D3C, AKNA, SYTL3, GIMAP4, FGL2, UNC13D, PRF1, GIMAP8, RASSF5, PDCCD4, PREX1, GBP4, HLA-DPB1, GIMAP5, SMAP2, FERMT3, STAT5B, APOL3, GNGT2, SPOCK2, HLA-E, KLRF1, C16orf86, OPHN1, STK4, FYN, SNRK, GBP2, IRF1, EPB41L3, CYTIP, RAPGEF3, ISG20, PPP1R16B, STAT5A, PTPN6, LSP1, ADGRE5, HAVCR2, CELF2, FGD3, IRF8, SPN, ZNF792, ITK, HLA-B, TARP, RUFY2, SULT1A1, SSH2, PRKCH, ZNF791, GIMAP6, GK, TRIM21, TMOD2, PLEKHO1, GBP1, ZNF439, SMARCC2, RAB37, CEP85L, ARMC2, IL27RA, FUT7, KIR2DL3, CDC42SE2, IZUMO1, FRY, NBEAL2, RORA, CARD8, HLA-C, MVB12B, PDE4B, ARHGAP30, LYSDM2, PEBP1, C11orf21, KLRC3, ZNF876P, HLA-J, SAFB2, REM2, TDRD7, ZCCHC6, PDE3B, CRLF3, SOX5, XCL2, SPTB, KIR3DL2, TTLL11, HLF, DNASE1L3, TSPAN32, HLA-A, PRKAB1, ABCC8, ZSWIM6, TCF4, PRCD, TNFRSF1B, CNTN4, DSCAML1, DGKE, ABCG2, RIC3, LRTOMT, CAMK4, XPC, CPEB3

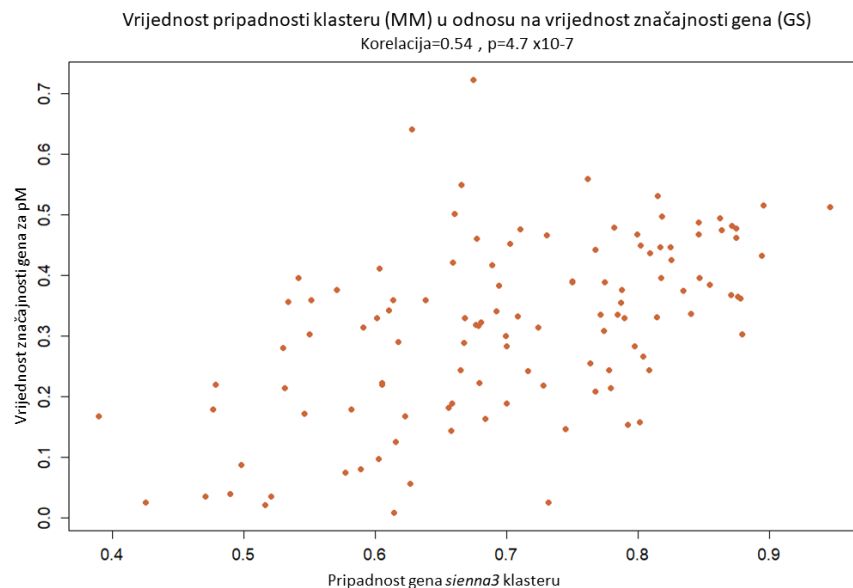
U klasteru *sienna3* pronađeno je obogaćenje za 144 pojma genske ontologije povezanih sa biološkim procesima. Većina obogaćenih pojmova genske ontologije u *sienna3* klasteru povezana je sa regulacijom imunološkog odgovora, poput primjerice regulacije imunološkog odgovora posredovane leukocitima (GO:0002703, $p=3,57 \times 10^{-5}$) ili limfocitima (GO:0002706, $p=1,71 \times 10^{-4}$); procesuiranje i prezentacija endogenih (GO:0002484, $p=1,78 \times 10^{-4}$ i GO:002486, $p=1,78 \times 10^{-4}$) i egzogenih (GO:0002480, $p=2,42 \times 10^{-4}$) antigena pomoću MHC kompleksa klase I; indukcija NK stanične tolerancije (GO:0002519, $p=3,47 \times 10^{-2}$) i zaštita od citotoksičnog djelovanja NK stanica (GO:0042270, $p=2,44 \times 10^{-3}$) ili INF γ signalni put (GO:0060333, $p=1,38 \times 10^{-7}$; GO:0071346, $p=5,35 \times 10^{-7}$; GO:0034341, $p=1,65 \times 10^{-6}$). Ostali obogaćeni pojmovi povezani su sa aktivacijom stanica, provođenjem signala, obrambenim odgovorom i regulacijom stanične adhezije (Slika 36.)



Slika 36. Rezultati analize genske ontologije u kategoriji Biološki procesi za *sienna3* klaster. Prikazan je sažetak rezultata, napravljenog pomoću Revigo mrežnog alata i R programa za vizualizaciju.

Geni koji su grupirani u *sienna3* klasteru imaju uglavnom negativno promijenjenu razinu ekspresije u odnosu na zdravo tkivo. Obzirom da se uloga tih gena povezuje sa regulacijom imunološkog odgovora, a poznato je da imunosupresivni mikrookoliš tumora pogoduje njegovom daljnjem razvoju i procesu metastaziranja, ovakvi rezultati korelacije *sienna3* klastera i prisutnosti metastaza nisu iznenađujuće.

Nadalje, izračunate su vrijednosti značajnosti gena (engl. *Gene significance*, GS) i vrijednosti pripadnosti klasteru (engl. *Module membership*, MM) (Slika 36.). Vrijednost značajnosti gena (GS) korelira svaki gen u modulu sa kliničkim parametrom od interesa, dok vrijednost pripadnosti modulu (MM) korelira eigengene modula sa profilom ekspresije gena.



Slika 37. Korelacija vrijednosti pripadnosti *sienna3* klasteru (MM) i vrijednosti značajnosti gena (GS)

Kao što je vidljivo sa Slike 37. geni koji se nalaze u *sienna3* klasteru pozitivno koleriraju sa prisutnošću metastaza u pacijentima oboljelim od tumora pluća skvamoznih stanica (korelacija=0,54, $p=4,7 \times 10^{-7}$).

Također, napravljena je mreža genskih interakcija za *sienna3* klaster. Za modeliranje mreže genskih interakcija korištena je STRING baza podataka, a mreža je vizualizirana u *Cytoscape* programu (Slika 38.). Karakteristike vizualizirane mreže provjerene su pomoću funkcije *Analyze Network* u *Cytoscape* programu.

Vrijednosti značajnosti gena i vrijednosti pripadnosti klasteru, zajedno sa mrežom interakcija gena u *sienna3* modulu napravljeni su u svrhu identifikacije visokoumreženih gena koji potencijalno imaju krucijalnu ulogu u procesu metastaziranja tumora pluća skvamoznih stanica. Za identifikaciju visoko umreženih gena u *sienna3* klasteru koristili smo granične vrijednosti: stupanj umreženosti čvorišta >5, vrijednost pripadnosti gena modulu (MM) >0,8 i značajnost gena unutar modula (GS) >0,2. Na temelju postavljenih parametara identificirano je 12 visokoumreženih gena prikazanih u Tablici 22.

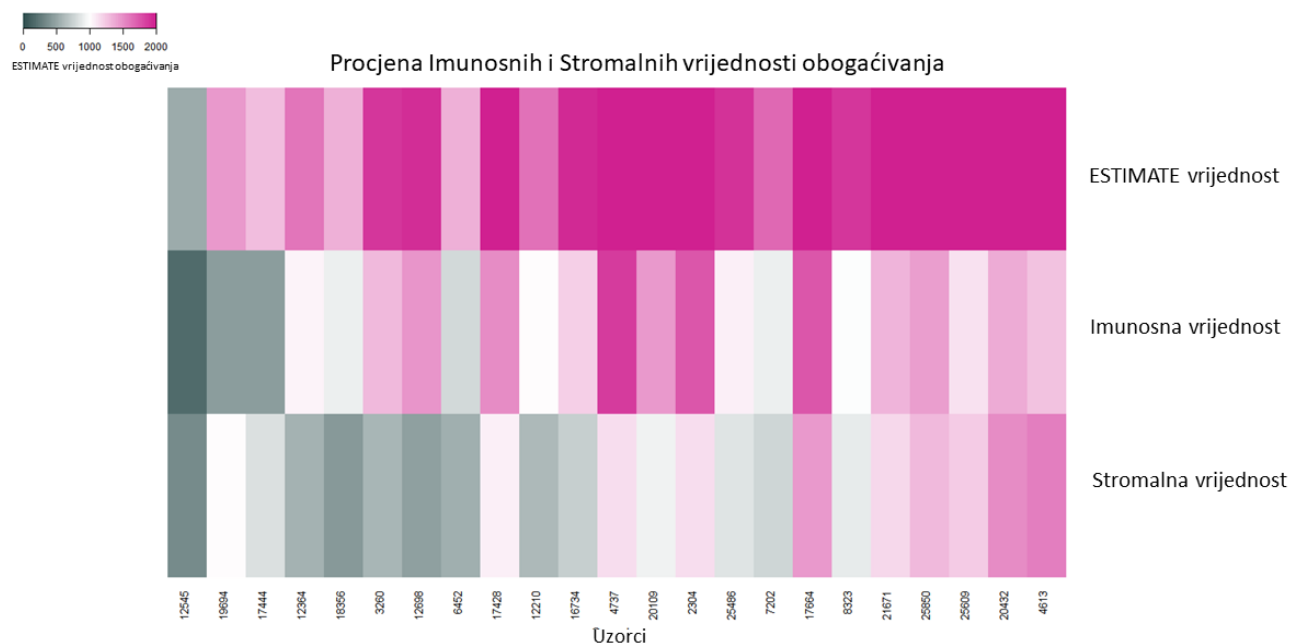
Tablica 22. Identificirani visokoumreženi geni u *sienna3* klasteru

Gen	GS.pM	MM.sienna3	Stupanj umreženosti čvorišta
FGL2	0,4779	0,8746	6
FYN	0,4499	0,8018	9
GIMAP4	0,4617	0,8748	5
GIMAP5	0,3373	0,8405	5
GIMAP7	0,5122	0,9462	5
GIMAP8	0,4743	0,8640	6
HLA-DPB1	0,4869	0,8465	7
HLA-E	0,5316	0,8152	15
KLRF1	0,3311	0,8147	5
PRF1	0,3676	0,8709	15
STAT5B	0,4465	0,8245	9
UNC13D	0,4823	0,8717	7

4.1.7. Rezultati procjene razine infiltriranosti tumora sa imunološkim stanicama i povezanosti sa sveukupnim preživljenjem

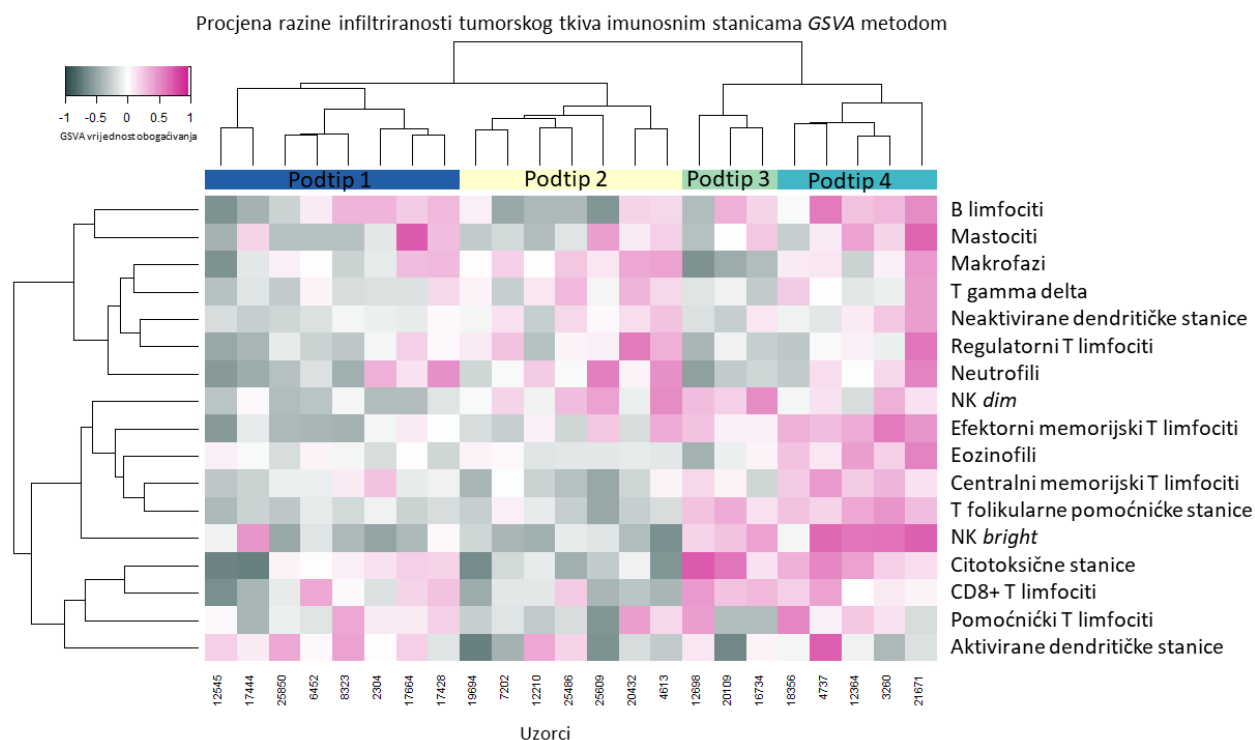
Tkivo solidnih tumora je heterogeno te se u njegovom okruženju, osim tumorskih stanica, nalaze epitelne stanice, stromalne stanice, imunološke stanice i vaskularne stanice. Stromalne stanice mogu doprinijeti rastu tumora i progresiji bolesti [9], dok infiltrirajuće imunološke stanice mogu imati anti-tumorsko djelovanje, ili mogu podupirati rast tumora i proces metastaziranja [10]. Kako bi provjerili razinu infiltriranosti tumora sa imunosnim stanicama u uzorcima eksperimentalne kohorte, napravljena je analiza obogaćivanja setovima gena na razini pojedinačnih uzoraka pomoću *ESTIMATE* paketa u R programu.

Rezultati ove analize ukazuju na različite razine infiltriranosti tumorskog tkiva sa imunosnim i stromalnim stanicama. Tri uzorka pokazuju najveću razinu stupnja infiltriranosti imunosnim stanicama (Slika 39.), dok tri uzorka pokazuju najmanju razinu stupnja infiltriranosti. Prisutnost stromalnih stanica u većini tkiva je umjerena, odnosno slična za većinu uzoraka u analizi.



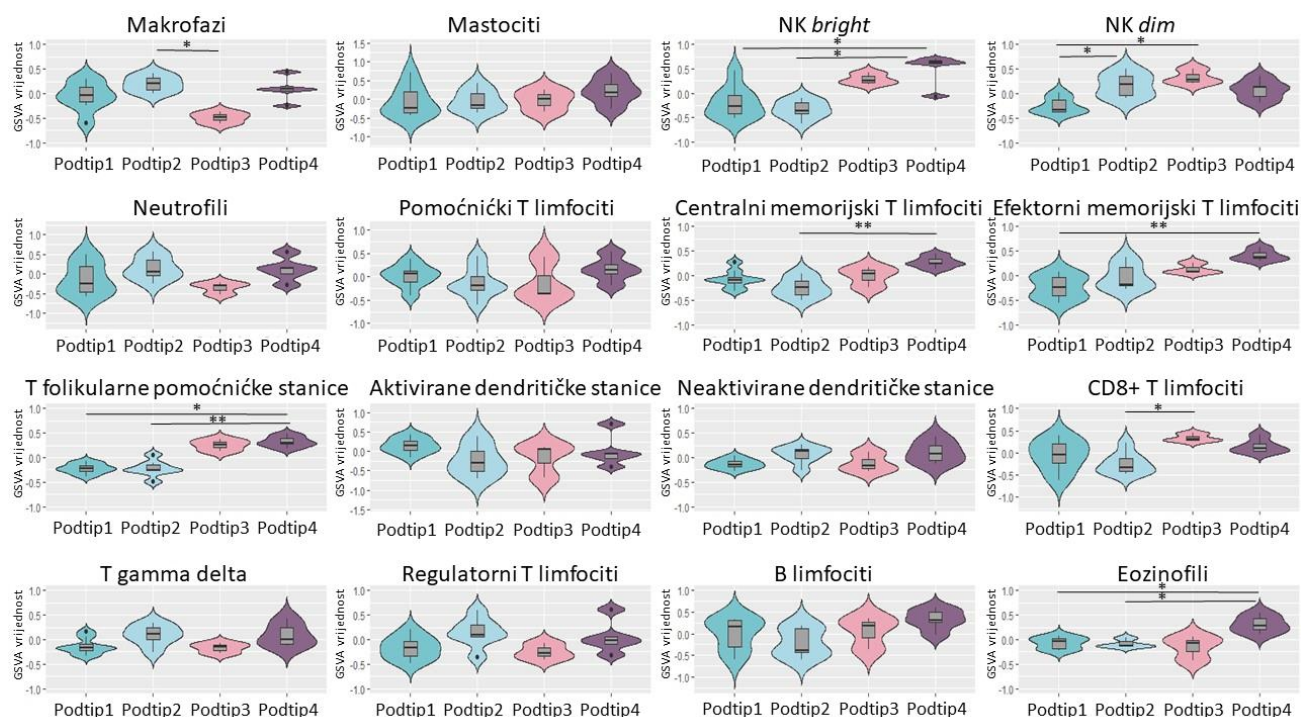
Slika 39. Procjena razine infiltriranosti tumora imunosnim i stromalnim stanicama. Rezultat je prikazan u obliku toplinske mape u kojoj tamnija roza boja označava veću vrijednost obogaćivanja, a tamno siva najmanje vrijednosti obogaćivanja.

Obzirom da rezultati *ESTIMATE* analize ukazuju na postojanje različite razine infiltracije tumorskog tkiva sa imunosnim stanicama u uzorcima, htjeli smo detaljnije proučiti profil tumorskog mikrookoliša. U tu svrhu napravljena je *GSVA* analiza pomoću koje su procijenjene razine infiltriranosti tumora za 16 vrsta imunosnih stanica, kao i stanica sa citotoksičnim djelovanjem. Izračunate su *GSVA* vrijednosti obogaćivanja za sve uzorke u analizi i za sve setove gena za obogaćivanje. Na temelju izračunatih *GSVA* vrijednosti obogaćivanja uzorci su grupirani klasteriranjem u imunosne podtipove koji imaju različiti profil stanica koje infiltriraju tumorsko tkivo. Kao što je vidljivo na Slici 40., na temelju profila ekspresije gena povezanih sa specifičnim imunološkim stanicama, u našim uzorcima su identificirana 4 imunosna podtipa. Od 23 tumorska uzorka u analizi, 8 uzoraka se grupiralo u Podtip 1 (34% uzoraka), 7 uzoraka se grupiralo u Podtip 2 (30% uzoraka), 3 uzorka u Podtip 3 (13% uzoraka) i 5 uzoraka u Podtip 4 (21% uzoraka).



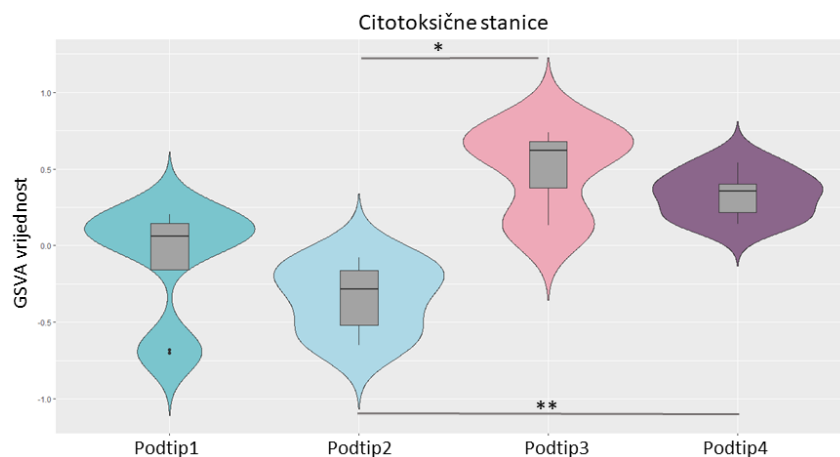
Slika 40. Procjena razine infiltriranosti tumorskog tkiva sa specifičnim imunskim stanicama pomoću GSVA metode. Prikazana je toplinska mapa u kojoj boja polja označava razinu ekspresije gena pojedinačnog seta za obogaćivanje, u odnosu na ostale uzorke u analizi. Linija iznad toplinske mape ukazuje na pripadnost određenom imunskom podtipu.

Rezultati ove analize, na temelju profila infiltracije tumora imunskim stanicama, ukazuju da su Podtip 3 i Podtip 4 potencijalno imunogeničniji jer je u njima zabilježena veća razina infiltracije sa NK *bright* i CD8+ T limfocitima u odnosu na Podtip 1 i Podtip 2. Nadalje Podtip 3 i Podtip 4 pokazuju veću razinu procijenjene infiltriranosti tumorskog tkiva sa antigen iskusnim stanicama (engl. *antigen experienced cells*), poput centralnih memorijskih T limfocita, efektnih memorijskih T limfocita i T folikularnih pomoćničkih stanica, u usporedbi sa Podtipom 1 i Podtipom 2. Nadalje, najmanja razina infiltracije tumorskog tkiva sa makrofazima zabilježena je za Podtip 3, dok je najviša razina zabilježena za Podtip 2 (Slika 41.).



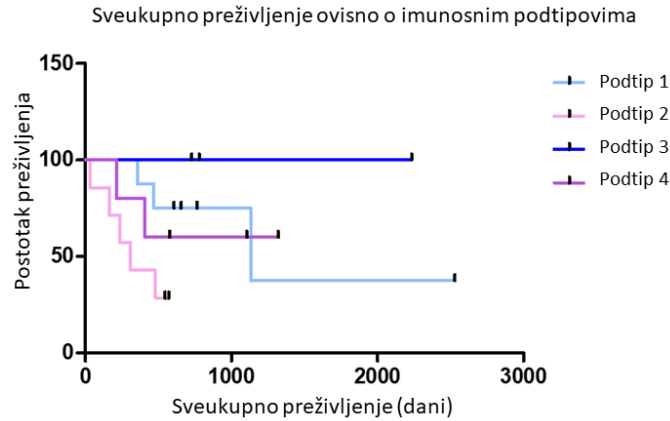
Slika 41. Procjena razine infiltriranosti tumorskog tkiva sa specifičnim imunskim stanicama prema pripadnosti imunskim podtipovima. Prikazani su violinski dijagrami u kojima pravokutnici označavaju interkvartilni raspon GSEA vrijednosti. Horizontalna linija unutar pravokutnika označava centralnu vrijednost (medijan). Oblik violinskog dijagrama ovisi o distribuciji GSEA vrijednosti. Statistička značajnost izračunata je Kruskal-Wallis testom sa *post hoc* Dunn-ovim testom

Rezultati GSEA obogaćivanja za set gena kojeg predstavljaju geni specifični za aktivirane T limfocite, gama delta T stanice i NK stanice zajedno (set nazvan Citotoksične stanice), također ukazuje da su Podtip 3 i Podtip 4 potencijalno imunogeničniji od preostala dva podtipa, odnosno da je ekspresija gena povezana sa ovim setom veća u navedenim podtipovima, u odnosu na ostale podtipove (Slika 42.). Ovi rezultati mogu ukazivati na potencijalno izraženiju protutumorsku aktivnost imunoloških infiltrata u Podtipu 3 i Podtipu 4.



Slika 42. Procjena razine infiltriranosti tumorskog tkiva sa citotoksičnim imunološkim stanicama (T limfociti, $T\gamma\delta$ i NK stanicama) prema pripadnosti imunosnim podtipovima. Prikazani su violinski dijagrami u kojima pravokutnici označavaju interkvartilni raspon GSVa vrijednosti. Horizontalna linija unutar pravokutnika označava centralnu vrijednost (medijan). Oblik violinskog dijagrama ovisi o distribuciji GSVa vrijednosti. Statistička značajnost izračunata je Kruskal-Wallis testom sa *post hoc* Dunn-ovim testom

Obzirom da se u literaturi veća razina infiltracije tumorskog tkiva sa imunološkim stanicama često može povezati sa preživljenjem pacijenata, zanimalo nas je dali pripadnost uzoraka u našoj kohorti određenom imunosnom podtipu ima utjecaj na sveukupno preživljenje pacijenata. Napravljena je analiza sveukupnog preživljenja Kaplan-Meierovom metodom u *GraphPad* programu. Razlike u preživljenju između pojedinačnih krivulja preživljenja ispitane su pomoću analize varijanci (ANOVA test), sa korigiranom p vrijednosti pomoću Bonferroni metode. Kao što je vidljivo sa Slike 43. rezultati ove analize pokazali da pripadnost određenom imunosnom podtipu u našem slučaju nema utjecaj na sveukupno preživljenje pacijenata ($p=0,09$), odnosno da razina infiltracije tumora sa imunosnim stanicama nije glavni čimbenik koji bi utjecao na sveukupno preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica.

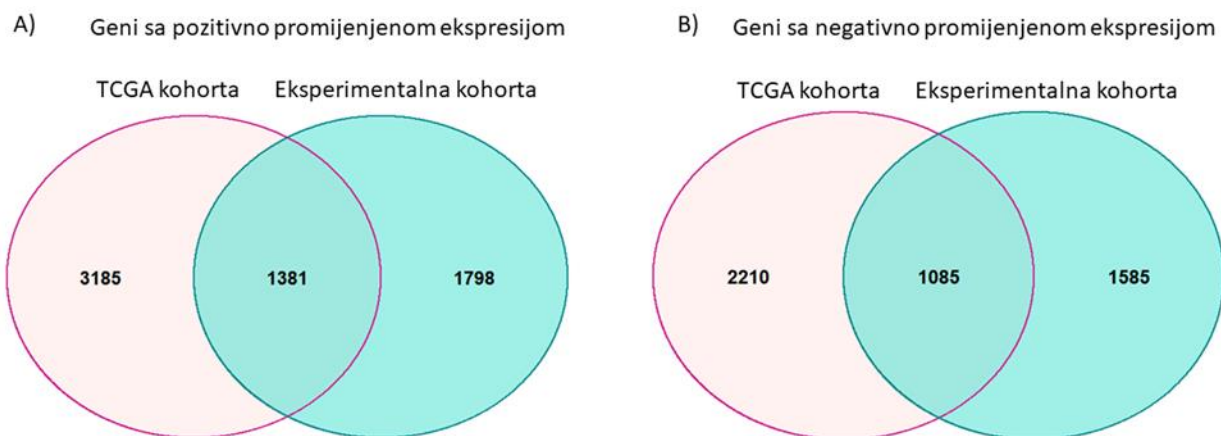


Slika 43. Utjecaj razine infiltracije tumora imunosnim stanicama na sveukupno preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica

4.1.8. Rezultati validacije transkriptomskih analiza na TCGA setu podataka

Kako bi validirali pronađene diferencijalno eksprimirane gene u eksperimentalnoj kohorti na većem broju uzoraka, ponovili smo analizu diferencijalne ekspresije na validacijskoj kohorti. Validacijska kohorta uključivala je ukupno 511 uzoraka, od kojih je 223 uzorka pacijenata oboljelih od tumora pluća skvamoznih stanica iz TCGA LUSC kohorte, a 288 je zdravih kontrola iz GTEx seta podataka. Za usporedbu razine ekspresije gena u obje kohorte, postavili smo graničnu vrijednost promjene ekspresije na $\log_2FC \geq |1|$.

Na temelju postavljene granične vrijednosti promjene ekspresije, u validacijskoj kohorti je pronađeno 4566 gena sa pozitivno promijenjenom ekspresijom, u odnosu na eksperimentalnu kohortu u kojoj je pozitivno promijenjeno 3179 gena sa $\log_2FC \geq |1|$ (Slika 43. A)). Rezultati za gene sa pozitivnom promjenom u razini ekspresije prikazani su tablično u Tablici 23. U obje testirane kohorte pronađeno je ukupno 1381 gen sa pozitivno promijenjenom razinom ekspresije u tumorskim uzorcima. U validacijskoj kohorti pronađeno je 3295 gena sa negativno promijenjenom razinom ekspresije, u odnosu na eksperimentalnu kohortu u kojoj je negativno promijenjeno 2670 gena (Slika 43. B)). Rezultati za gene sa negativnom promjenom u razini ekspresije prikazani su u Tablici 22. U obje ispitivane kohorte pronađeno je ukupno 1085 gena sa negativno promijenjenom razinom ekspresije u tumorskim uzorcima.



Slika 44. Vennov dijagram za gene sa pozitivnom (A) i negativnom (B) promijenom u razini ekspresije, u obje ispitivane kohorte pacijenata oboljelih od tumora pluća skvamoznih stanica

Tablica 23. Popis gena čija je promjena u razini ekspresije potvrđena u obje ispitivane kohorte.

Geni su prikazani abecednim redom

Geni sa pozitivnom promjenom u razini ekspresije u eksperimentalnoj i validacijskoj kohorti
A2ML1, ABCA12, ABCA13, ABCB4, ABCB6, ABCC1, ABCF3, ACAD9, ACKR3, ACOT11, ACTL6A, ACTR3C, ACVR1C, ADA, ADAM28, ADAMTS12, ADAMTS16, ADAMTSL5, ADCK2, ADCY10, ADD2, ADGRB2, ADGRF4, ADGRG1, ADORA2B, AFAP1-AS1, AFAP1L2, AGAP2-AS1, AGL, AGMAT, AGMO, AGRN, AHCY, AHNK2, AIMP2, AJUBA, AK4, AKR1B10, AKR1C2, AKR1C3, ALDH18A1, ALDH1A3, ALDH1B1, ALDH4A1, ALDOC, ALG1, ALG10, ALG1L, ALG3, ALKBH4, ALPK2, AMACR, AMER1, AMMECR1, ANAPC1, ANKEF1, ANKH, ANKRD18A, ANKRD19P, ANKRD20A8P, ANKRD22, ANO7P1, ANXA8, AP1S1, AP5S1, APLP1, APOBEC3B, APOBEC3F, APOBEC3G, AQP3, ARHGAP11A, ARHGAP11B, ARHGEF16, ARHGEF19, ARHGEF34P, ARHGEF4, ARHGEF5, ARL4C, ARMC6, ASF1B, ASIC1, ASPHD1, ASPHD2, ATAD2, ATAD5, ATF5, ATIC, ATP12A, ATP5G1, ATRN, ATRNL1, AUNIP, AURKB, B3GNT3, B3GNT4, B4GALNT1, B4GALNT3, B4GALNT4, B4GALT2, B4GALT4, B9D1, BBOX1, BCAS4, BCL11A, BCL11B, BCL2L12, BCL9, BDH1, BDKRB1, BDKRB2, BEND3, BEND6, BFSP1, BHLHA15, BIRC5, BLCAP, BLM, BLNK, BMP8A, BNC1, BOLA2B, BOP1, BORA, BPHL, BRCA1, BRCA2, BRCC3, BRF2, BRIP1, BUB1, BYSL, C10orf2, C11orf80, C12orf56, C14orf169, C15orf41, C15orf48, C16orf59, C16orf91, C17orf53, C17orf58, C17orf75, C19orf24, C19orf48, C1orf106, C1orf112, C1orf131, C1orf53, C1orf74, C1QTNF6, C2orf27, C2orf29, C3orf58, C3orf67, C5orf46, C9orf43, C9orf69, C9orf84, CABLES2, CACNA1B, CALB1, CALML3, CAPN1, CAPNS2, CARD14, CARM1, CASC5, CASKIN1, CATSPER2P1, CBR3, CBX2, CBX8, CCDC34, CCDC58, CCDC86, CCL8, CCNA2, CCNE1, CCNE2, CCNE, CCR6, CCR7, CCR8, CCSER1, CD19, CD207, CD27, CD276, CD3EAP, CD79A, CD79B, CDC20, CDC25C, CDCA2, CDCA3, CDCA4, CDCA7, CDH1, CDH2, CDH3, CDH8, CDHR1, CDK4, CDK5R1, CDT1, CEACAM5, CEL, CELSR1, CELSR2, CELSR3, CENPE, CENPH, CENPM, CENPN, CENPO, CENPW, CEP72, CEP85, CERCAM, CERS3, CES3, CFHR3, CGREF1, CHAF1A, CHAF1B, CHEK1, CHML, CHPF, CHRM3, CHRNA1, CHRNA7, CHRN4, CHST14, CHST2, CHST6, CHST7, CIB2, CILP2, CKAP2, CKAP2L, CKAP4, CLCA2, CLCA4, CLCN2, CLDN1, CLEC17A, CLN6, CLSTN1, CNFN, CNTNAP2, COA7, COCH, COL11A1, COL12A1, COL17A1, COL19A1, COL1A1, COL1A2, COL22A1, COL4A5, COL5A1, COL5A2, COL8A2, COLCA2, COPS7A, COQ3, COQ9, CORO2A, CPA4, CPNE4, CPSF3, CPXM1, CPXM2, CREB3L4, CSMD1, CSMD2, CSRN3, CST2, CTAGE11P, CTAGE15, CTAGE4, CTAGE8, CTHRC1, CTSK, CTSV, CXCL10, CXCL14, CXCL9, CYB5R2, CYC1, CYP26A1, CYP27B1, CYP4F3, DAPL1, DARS2, DBR1, DCAF13, DCBLD2, DCPS, DCTPP1, DDR1, DDX11, DDX28, DEPDC1, DEPDC4, DERL3, DFNA5, DGUOK-AS1, DHCR7, DHFRL1, DHRS4L1, DHX57, DIAPH3, DIEXF, DIO2, DIP2B, DISP2, DKK1, DLK2, DLX5, DLX6, DMRT2, DMRT3, DNAAF5, DNAH17, DNAJC22, DNAJC9, DNPH1, DOLK, DPFI, DPH2, DPP3, DQX1, DROSHA, DRP2, DSC2, DSC3, DSCC1, DSG1, DSG2, DSG3, DSN1, DTNB, DTYMK, DUXAP8, E2F2, E2F8, EDARADD, EDN2, EEF2KMT, EFNA4, EFN3, EFS, EGLN3, EHHADH, EHMT2, ELFN1, ELMO3, ELOVL2, ELOVL6, EME1, ENC1, ENTPD2, ENTPD3, ENTPD7, EPGN, EPHA1, EPHA4, EPHA7, EPHB2, EPHB3, EPHX3, EPN3, EPPK1, EPT1, ERAL1, ESCO2, ESPL1, ESPN, ESRP1, ETNK2, EXO1, EXOSC5, F2RL2, F8A3, FAAP24, FAM101A, FAM110C, FAM111B, FAM169A, FAM171A2, FAM189B, FAM213A, FAM57A, FAM64A, FAM69A, FAM71E1, FAM72A, FAM72B, FAM72D, FAM83A, FAM83B, FAM83C, FAM83D, FAM83F, FAM83G, FAM83H, FAM84A, FAM86C1, FAM86JP, FANCA, FANCF, FANCG, FANCI, FANCM, FAP, FARSA, FASTKD3, FAT1, FAT2, FBXL19, FBXO16, FBXW9, FCRL5, FDXR, FECH, FEN1, FERMT1, FETUB, FEZF1, FGD1, FGF11, FGF12, FGFBP1, FGFR2, FGFR3, FHL2, FHOD3, FIGNL1, FJX1, FLAD1, FLJ35934, FMO1, FNDC1, FNTB, FOXD1, FOXD3, FOXE1, FOXI3, FOXM1, FOXN1, FOXP3, FOXQ1, FOXRED2, FRMD5, FRMD6, FRRS1L, FSCN1, FSD1, FST, FSTL4, FUT2, FUT3, FUT8, FZD10, FZD6, FZD7, G6PD, GABRA3, GABRG3, GABRP, GABRR1, GAD1, GALNT1, GALNT2, GALNT4, GALNT6, GBP6, GCLC, GCNT3, GCSH, GDA, GDAP1, GDFP1, GEMIN7, GEN1, GFRA3, GGH, GGT6, GINS3, GINS4, GJB2, GJB3, GJB5, GJB6, GLI2, GLMP, GMDS, GMNN, GOLGA7B, GORASP2, GPC1, GPC2, GPNMB, GPR1, GPR143, GPR161, GPR19, GPR37,

Tablica 23. Nastavak.

GPR82, GPR87, GPRIN1, GPT2, GPX2, GPX7, GREM1, GRHL2, GRHL3, GRIK2, GRIN2D, GRPR, GRWD1, GSG2, GSS, GSTM3, GSTO2, GSTZ1, GTF2H3, GTF3C4, GYG2, GYLTL1B, H2AFY2, HAP1, HARB1, HAS3, HAUS1, HCAR1, HDHD3, HEATR1, HELLS, HEPHL1, HES6, HEY1, HGH1, HIST1H1B, HIST1H2A, HIST1H2AI, HIST1H2AM, HIST1H2BH, HIST1H2BL, HIST1H3F, HIST1H3G, HIST2H3A, HIST2H4A, HIST3H2A, HJURP, HK2, HMBS, HMGA1, HMGC51, HMGXB4, HN1L, HOMEZ, HOOK1, HOXA1, HOXA11, HOXA13, HOXA9, HOXC4, HOXC6, HOXD4, HOXD8, HS3ST1, HS6ST1, HTR3A, HTR7, ICE1, IDH2, IGF2BP1, IGF2BP2, IGFBP2, IGFBP3, IGFBP5, IGFBP1, IGHV1OR15-1, IGSF1, IGSF3, IGSF9, IKBKE, IL1A, IL1RL2, IL22RA1, IL2RA, IL36G, IL4I1, ILDR1, IMP3, IMPA2, INCENP, INO80C, INTS5, IPO4, IQGAP3, ISOC2, ITGA11, ITGA6, IVL, JAG1, JAG2, JAKMIP3, JMD4, JUP, KCNG3, KCNH1, KCNH8, KCNK1, KCNMB3, KCNQ5, KCP, KCTD1, KDELRL3, KDF1, KDM4D, KEAP1, KHDC1, KHK, KIAA1024, KIAA1211, KIAA1217, KIAA1522, KIAA1524, KIAA1549L, KIF14, KIF15, KIF18A, KIF20B, KIF23, KIF24, KIF26B, KIF2C, KIF4A, KIFC1, KLF5, KLHL13, KLRG2, KNOP1, KNTC1, KPNA2, KPTN, KREMEN1, KRT14, KRT31, KRT5, KRT6A, KRT6C, KRT80, KRTCAP3, KTI12, L1CAM, LAMA1, LAMB3, LAMB4, LAMC2, LAMP5, LANCL2, LAPTM4B, LAX1, LDOC1L, LEMD1, LGALS2, LGALS7, LGALS7B, LGALS9B, LGR4, LGR5, LIMK1, LINGO1, LINGO2, LGLL1, LMNB2, LNX1, LOXL1, LOXL2, LOXL4, LPAR5, LRP1B, LRP8, LRRC1, LRRC17, LRRC61, LRRC8E, LSG1, LTBP, LUZP2, LY6D, LY6K, LYPD3, LYPD6, LYPD6B, MAGED1, MAGED4B, MAMSTR, MANEAL, MAP10, MAP7, MAPK6, MARCKSL1, MARS2, MAZ, MCCC1, MCM10, MCM2, MCM5, MCM6, MCM8, MDFI, MDK, MED12L, MED20, MEGF10, MET, METTL1, METTL13, METTL8, MEX3D, MFSD3, MFSD5, MFSD9, MGME1, MICAL3, MICALL1, MIR205HG, MIS18A, MKI67, MLLT11, MMACHC, MMP1, MMP11, MMP12, MMP13, MMP14, MMP16, MMP9, MMS22L, MOCS3, MPI, MPPED1, MPPED2, MPV17L, MRGBP, MRPL12, MRPL15, MRPL2, MRPL4, MRPS12, MRPS26, MSC-AS1, MSH2, MSH6, MSI1, MSI2, MSX2, MT1H, MTBP, MTCH2, MTCL1, MTRF2, MTX1, MXRA5, MYBL2, MYCL, MYO3A, MZB1, NAA20, NABP2, NACC1, NCAPD3, NCAPG, NCAPG2, NCS1, NDC1, NDRG1, NEFH, NEFL, NELL2, NETO1, NETO2, NFE2L3, NIP7, NIPAL1, NIPAL4, NIPSNAP1, NLGN1, NLGN4Y, NLN, NME1, NMNAT3, NMRAL1, NMU, NOC2L, NOP2, NOTCH3, NPM2, NPM3, NQO1, NR2C2AP, NRARP, NRRCAM, NRIP3, NRRK, NRXN2, NSUN2, NTS, NUDT19, NUF2, NUP107, NUP155, NUP205, NUP37, NUP62, NUSAP1, OCA2, ODC1, OGDHL, OLFM2, OPA1, ORC1, OVOL1, OXGR1, P3H4, PADI3, PAFAH1B3, PAK1IP1, PAK6, PANX2, PAQR4, PAQR7, PARD6G, PARBP, PAX5, PAX6, PAX9, PC, PCLO, PCNA, PCSK9, PCYT1B, PDCD2L, PDCD6, PDK1, PDLIM4, PES1, PFKP, PFN2, PGAM4, PGF, PHF24, PHKA1, PIF1, PIGM, PIGO, PIGS, PIGW, PIGX, PIK3C2G, PIK3R2, PIM2, PITX2, PKHD1, PKM, PKP1, PKP2, PLA2G4E, PLAU, PLCH2, PLD1, PLD4, PLD5, PLEKHG3, PLEKHG4B, PLEKHG6, PLEKHN1, PLK1, PLK4, PLXNA1, PMM2, PMS2, PMS2P1, PMS2P2, PMS2P5, PMS2P9, PNCK, PNMA1, PODNL1, POFUT1, POLA2, POLD2, POLDIP2, POLR1A, POLR1B, POMK, POP1, POP7, POPDC3, POTEI, POU2AF1, POU6F2, PPCS, PPEF1, PPFIA3, PPFIA4, PPAL4C, PPIF, PPIL1, PPM1H, PPP1R14C, PPP2R2C, PPP2R3A, PPP4R4, PRC1, PRDX4, PRICKLE3, PRIM2, PRIMA1, PRKDC, PRKX, PROC, PROM2, PRPF19, PRPF4, PRR11, PRSS21, PRSS22, PRSS3, PRSS50, PRTFDC1, PSMD2, PSMG1, PSMG3, PTDSS1, PTGES, PTGFRN, PTHLH, PTK6, PTPN20, PTPRF, PTPRH, PTPRN, PTTG1, PUS1, PXDNL, PYCR1, PYCRL, PYGB, RAB40B, RAB6B, RABEPK, RAC3, RACGAP1, RAD51, RAD51AP1, RAD51AP2, RAD54B, RAD54L, RAET1G, RAET1L, RAG1, RAI14, RALGPS2, RANBP1, RAP2B, RAPGEFL1, RARG, RASAL1, RASGEF1A, RASL11B, RASSF6, RASSF6, RBBP8, RBP1, RCC2, RCCD1, RCN2, REEP6, RFP3, RFPL1S, RGMA, RGS17, RGS4, RHBDD3, RHBDL2, RHCE, RHCG, RHNO1, RHOD, RIBC2, RIMBP3, RIMS2, RIOK1, RITA1, RMI2, RNASEH2A, RND2, RNF165, RNF187, RNF26, RNF27, ROBO1, RPEL1, RPIA, RPL39L, RPLP0P2, RPP25L, RPUUSD1, RRM1, RRM2, RTN4IP1, RTN4R, RUNX2, S100A16, S100A2, SAC3D1, SAMD1, SARS2, SASS6, SATB2, SBK1, SCAMP5, SCG2, SCG5, SCN5A, SCN8A, SCN9A, SCUBE3, SDK1, SDK2, SDR42E1, SEC14L2, SEC61A1, SEMA4A, SEMA4B, SEMA4F, SEPT3, SERINC2, SERPINB13, SERPINB3, SERPINB4, SERPINB5, SERPINB7, SERPINE2, SEZ6L2, SFN, SFRP2, SFRP4, SFXN1, SFXN4, SGPL1, SGPP2, SH3RF2, SHC4, SHCBP1, SHISA6, SHMT2, SHOX2, SIGMAR1, SIM2, SIPA1L3, SIRPG, SIX1, SKA1, SKP2, SLC16A1, SLC16A13, SLC16A9, SLC18B1, SLC1A4, SLC22A20, SLC24A2, SLC25A10, SLC25A15, SLC25A17, SLC27A4, SLC28A3, SLC2A5, SLC35F2, SLC35G1, SLC39A11, SLC47A2, SLC4A11, SLC4A3, SLC52A2, SLC5A6, SLC6A15, SLC6A8, SLC7A11, SLC7A8, SLC9A3R1, SLC9A4, SLC9A7, SMARCA4, SMARCB1, SMC2, SMO, SMOX, SMUG1, SNAP47, SNAPIN, SNCAIP, SNX31, SNX8, SORD, SOWAHB, SOWAHC, SOX15, SOX2, SOX21, SOX30, SOX4, SPAG4, SPAG5, SPAST, SPC24, SPC25, SPDL1, SPIB, SPINK5, SPOCK1, SPOCK3, SPRR1A, SPRR1B, SPRR2A, SPRR3, SPTBN2, SRD5A1, SRGAP3, SRPRB, SRXN1, ST14, ST6GAL2, ST6GALNAC2, STEAP3, STIL, STMN1, STON2, STXBP5L, SULF2, SURF2, SUV39H1, SV2A, SVOPL, SYT1, SYT12, SYT14, SYT7, TACO1, TADA1, TAF4B, TARS2, TBGR4, TBX18, TCF19, TCF20, TCF3, TDRKH, TENM2, TENM3, TENM4, TFAP4, TFB1M, TFR2, TFRC, TGM1, THAP7-AS1, THAP8, THBS2, THBS4, THEM6, THOC3, THSD7B, THY1, TIAM1, TICRR, TIGIT, TIMELESS, TIMM50, TIMM8A, TK1, TKFC, TLCD1, TLDC1, TLR10, TM4SF19, TMED3, TMEM102, TMEM116, TMEM132A, TMEM134, TMEM14A, TMEM161A, TMEM17, TMEM171, TMEM177, TMEM182, TMEM184A, TMEM185B, TMEM189, TMEM199, TMEM200C, TMEM201, TMEM203, TMEM206, TMEM216, TMEM223, TMEM241, TMEM41A, TMEM45A, TMEM54, TMEM63C, TMEM69, TMEM79, TMPRSS11D, TMPRSS11E, TMPRSS13, TMTC4, TNFRSF13C, TNFRSF17, TNFRSF21, TNFSF11, TNK1, TNS4, TOMM40, TONSL, TOP2A, TOPBP1, TOR3A, TP63, TP73, TPBG, TPD52, TPD52L1, TPI1, TPI1P2, TPTE2, TPX2, TRAF7, TRAF, TRAM1L1, TRAP1, TRIM16L, TRIM2, TRIM28, TRIM29, TRIM45, TRIM59, TRIM9, TRMT10C, TRMT12, TRMT6, TRMT61B, TROAP, TRPA1, TRPV3, TRPV4, TRPV6, TRUB2, TSFM, TSKU, TSPY26P, TSPYL5, TSR1, TSSC1, TSTA3, TTC26, TTC30A, TTC30B, TTC4, TTF2, TTI1, TTLL12, TTYH3, TUBA1C, TUBB3, TUBBP5, TUFT1, TWIST1, TXNDC5, TYMS, TYSND1, U2SURP, UBD, UBE2C, UBE2O, UBE2S, UGT2B17, UMPS, UNC80, UNG, UPK1B, UPK3BL, URB1, URB2, USH1C, USH1G, USP18, USP5, UTP14A, VANGL1, VGLL1, VOPP1, VPS25, VTCN1, VWA2, VWDE, WASF1, WDHD1, WDR12, WDR3, WDR54, WDR66, WDR77, WIPF3, WNK2, WNT3, WNT5A, XKR9, XPO5, XPR1, XRCC2, XXYL1, XYLB, XYLT2, YARS2, YBX2, YDJC, YEATS2, YIF1A, ZBED1, ZBED4, ZBTB12, ZBTB7C, ZBTB9, ZDHHC12, ZDHHC13, ZFHX4, ZFP64, ZIC2, ZIC5, ZMYND19, ZNF114, ZNF131, ZNF157, ZNF2, ZNF212, ZNF280C, ZNF385A, ZNF408, ZNF469, ZNF48, ZNF492, ZNF526, ZNF551, ZNF557, ZNF683, ZNF695, ZNF703, ZNF711, ZNF713, ZNF724P, ZNF74, ZNF749, ZNF774, ZNF786, ZNF860, ZNF890P, ZNHIT3, ZNRF3, ZSCAN5A, ZSWIM1, ZWILCH, ZYG11A

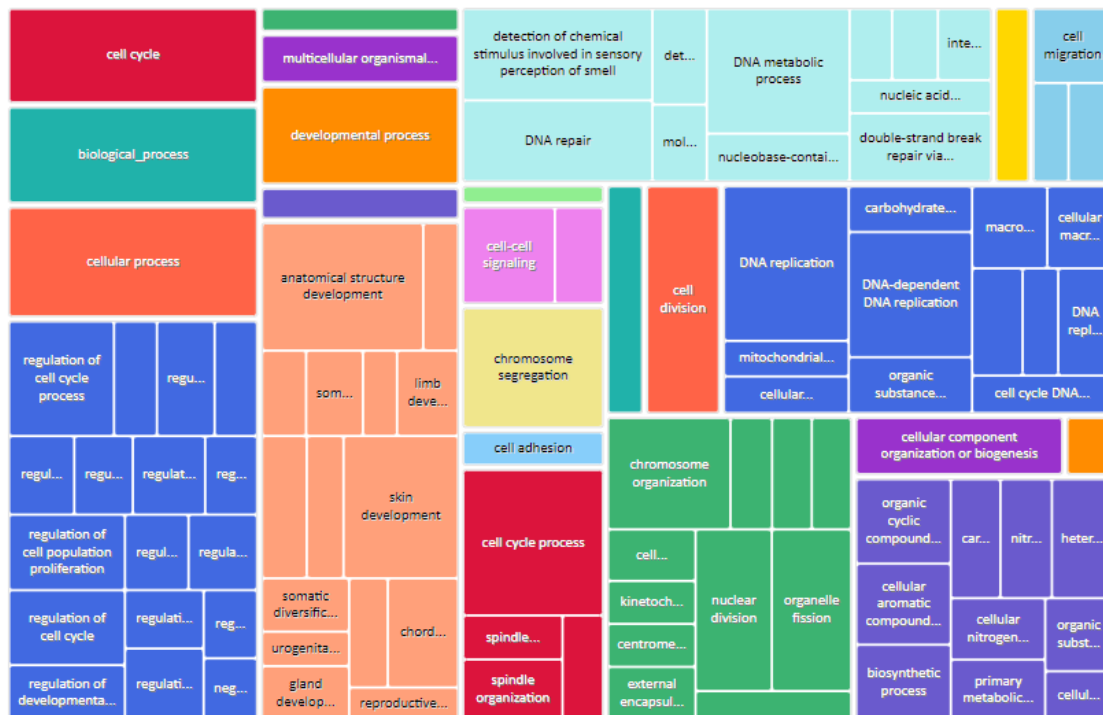
Tablica 23. Nastavak

Geni sa negativnom promjenom u razini ekspresije u eksperimentalnoj i validacijskoj kohorti
A2M, AATK, ABCA10, ABCA6, ABCA8, ABCA9, ABCB1, ABCC6, ABI3BP, ABLIM3, ABTB1, ACADL, ACE, ACKR1, ACKR4, ACRC, ACSL1, ACSM5, ACSS3, ACTC1, ACTG2, ACVRL1, ADAM33, ADAMTS1, ADAMTS8, ADAMTSL3, ADAMTSL4, ADARB1, ADGB, ADGRB3, ADGRE3, ADGRE5, ADGRF5, ADGRG3, ADGRL2, ADH1A, ADH1B, ADHFE1, ADIRF, ADRA1A, ADRB1, AFAP1L1, AFF1, AFF3, AGAP11, AGBL3, AGT, AGTPBP1, AGTR1, AGTR2, AJAP1, AK1, AKAP13, AKAP6, AKT3, ALAS2, ALB, ALDH1A2, ALDH3B1, ALOX5, ALOX5AP, ALPK3, ALPP, ALPPL2, ALS2CR12, AMY2A, AMY2B, ANG, ANGPT1, ANGPTL1, ANGPTL5, ANKFN1, ANKRD1, ANKRD20A2, ANKRD20A3, ANKRD29, ANKRD55, ANO2, ANO5, ANPEP, ANXA3, AOC3, AOC4P, AOX1, APBB1, APBB1IP, APOB, APOBR, APOD, APOL3, AQP1, AQP4, AQP9, ARG1, ARHGAP10, ARHGAP31, ARHGAP44, ARHGAP6, ARHGEF10, ARHGEF26, ARHGEF6, ARMCX1, ARRB1, ARRB2, ARRD2C, ART4, ASPA, ASRGL1, ATF3, ATG16L2, ATP13A4, ATP1A2, ATP1B2, ATP8B4, BCL2A1, BCL6, BCL6B, BDH2, BEND7, BEST1, BHMT2, BIN2, BMP5, BMPER, BMX, BPI, BTK, BTNL8, C10orf128, C10orf131, C10orf54, C10orf67, C11orf96, C14orf132, C19orf38, C1orf116, C1orf132, C1orf168, C1orf194, C1QTNF7, C2orf40, C4BPA, C4orf32, C5AR1, C5AR2, C5orf56, C6, C6orf118, C7, C8B, C8orf34-AS1, C9orf24, C9orf72, CA3, CA4, CAB39L, CACNA1D, CACNA2D2, CACNB2, CADM3, CALCOCO1, CALCRL, CAMK2N1, CAMP, CAPN9, CAPS2, CAPSL, CASKIN2, CASP12, CASQ2, CASS4, CAV1, CCBE1, CCDC173, CCDC180, CCDC30, CCDC39, CCDC68, CCDC69, CCDC78, CCDC81, CCDC84, CCDC85A, CCL14, CCL15, CCL2, CCL23, CCRL2, CD163, CD177, CD226, CD300E, CD300LF, CD302, CD33, CD36, CD37, CD93, CDH5, CDHR3, CDHR4, CDK15, CDKL2, CDOL1, CEACAM3, CEACAM3, CEACAM8, CEPBD, CEP126, CFAP221, CFAP45, CFAP46, CFAP52, CFAP58, CFAP70, CFD, CFL2, CFLAR, CFP, CGNL1, CHRDL1, CITED2, CLDN18, CLDN5, CLEC12A, CLEC14A, CLEC1A, CLEC3B, CLEC4E, CLEC7A, CLIC2, CLIC3, CLIC5, CLK4, CMTM2, CNN1, CNRIP1, CNTFR, CNTN6, COBL, COL13A1, COL4A3, COL4A4, COL6A5, COL6A6, COX7A1, CPA1, CPA3, CPAMD8, CPB1, CPED1, CPM, CR1, CREB5, CREBRF, CRIP2, CSF3R, CSGALNACT2, CST7, CTSG, CXCL5, CXCR1, CXCR2, CYBRD1, CYP1A1, CYP27A1, CYP3A5, CYP3A7, CYP4B1, CYP4V2, CYTIP, DAB2, DAPK2, DAW1, DCN, DCSTAMP, DDX3Y, DDX60L, DEFA1, DEFA1B, DEFA3, DENND2A, DENND3, DGKE, DISP1, DKK3, DLL4, DNAAF1, DNAAF3, DNAH10, DNAH12, DNAH3, DNAH6, DNAH7, DNAH9, DNAI2, DNAL1, DNASE2B, DNM3, DOC2A, DOCK4, DOCK9, DOK3, DPEP2, DPP6, DPT, DPY19L2P2, DTNA, DUOX1, DUSP1, DUSP8, DYNLRB2, DYSF, ECM2, ECSCR, EDNRB, EFCAB13, EFCC1, EFHB, EGFL7, EMCN, EMP2, ENG, ENPEP, EPAS1, EPB41L2, EPB41L3, EPB41L5, ERG, ERICH2, ERICH3, ESAM, ESYT3, ETS1, F10, F11, F2RL3, F8, FABP3, FABP4, FAM105A, FAM107A, FAM150B, FAM153C, FAM162B, FAM166B, FAM167A, FAM179A, FAM184A, FAM189A2, FAM47E-STBD1, FAM65A, FAM65B, FAM65C, FAT3, FAT4, FAXDC2, FBLN1, FBLN5, FBP1, FCAR, FCER1G, FCGR3B, FCN1, FCN3, FERMT2, FFAR4, FGD5, FGF10, FGF14, FGF2, FGF7, FGFR4, FGR, FHL1, FHL5, FIBIN, FIGF, FKBP5, FLG2, FLI1, FLNC, FLT1, FLT4, FLVCR2, FMN1, FMO2, FOLR1, FOXA2, FOXF1, FPR1, FPR2, FRMD3, FRMPD2, FUT1, FXYP1, FXYP6, FZD4, G0S2, GAB2, GABARAPL1, GADD45B, GADD45G, GALNT15, GALNT16, GALNT18, GATA2, GCOM1, GPD3, GFRA1, GGTL1, GGTL2, GIMAP1-GIMAP5, GIMAP4, GIMAP5, GIMAP6, GIMAP7, GKN2, GLDN, GLIPR2, GLRX, GMFG, GNA14, GNG11, GNLY, GNRH1, GP6, GPBAR1, GPC5, GPD1, GPD1L, GPIHBP1, GPM6A, GPM6B, GPR146, GPR4, GPRC5A, GPSM3, GPX3, GRAMD2, GRAP, GRASP, GRIA1, GRK5, GSTM5, GUCY1A2, GYPC, GYPE, HECW2, HEMGN, HEYL, HIF3A, HIGD1B, HLA-E, HLA-F-AS1, HMGCLL1, HNF1B, HOPX, HOXA2, HOXA5, HPCAL1, HPGDS, HPSE2, HPX, HRH2, HSD11B1, HSD17B6, HSD17B7P2, HSPA12B, HSPA1A, HSPA1B, HSPB2, HSPB6, HSPB7, HTR1F, HYAL1, HYAL2, HYPK, ICAM1, ICAM2, ID4, IFITM2, IGFBP6, IGIP, IGF10, IL10, IL18R1, IL18RAP, IL1R1, IL1RL1, IL3RA, IL4R, IL5RA, IL6ST, IL7R, INMT, INMT-FAM188B, IRAK3, ITGA1, ITGA10, ITGA2B, ITGA8, ITGAM, ITGAX, ITGB3, ITIH3, ITIH5, ITM2A, ITPR1, JADE1, JAM2, JAM3, JDP2, JPH2, KANK2, KANK3, KANSL1L, KAT2B, KCNA5, KCNB1, KCNK3, KCNMB1, KCNN2, KCNT2, KCTD12, KDR, KHDRBS2, KIAA0040, KIAA0408, KIAA1462, KIR2DL1, KIR2DL3, KIR2DS4, KIR3DL1, KL, KLF15, KLF9, KLHDC1, KLHL41, KLKB1, KLRD1, KLRF1, KNDC1, KRT7, LAMA2, LAMB2, LAMC3, LAT2, LATS2, LCP2, LDB2, LDB3, LDLRAD1, LGI4, LHFP, LIFR, LILRA1, LILRA2, LILRA6, LILRA6, LILRB2, LILRB3, LIMCH1, LIMS2, LIRN7A, LINC00632, LINC01140, LINC01272, LIPN, LMCD1, LMO2, LMO7, LONRF2, LONRF3, LOXHD1, LPL, LRCH2, LRSP2, LRP2BP, LRRC25, LRRC32, LRRC43, LRRC46, LRRN3, LRRN4, LSAMP, LSMEM1, LTK, LYVE1, MAATS1, MACF1, MACROD2, MAFF, MAFK, MAGI1, MAL, MAMDC2, MAMLD1, MAOA, MAOB, MAP3K19, MAP6, MAPK15, MAPK4, MARC2, MARCH10, MARCO, MASP1, MCAM, MCEMP1, MEFV, MEOX2, MFAP4, MGAM, MGARP, MGLL, MGP, MICU3, MLKL, MMP8, MMRN1, MMRN2, MNDA, MOCS1, MORN5, MPO, MPP1, MRC1, MROH7-TTC4, MROH9, MS4A14, MS4A15, MS4A3, MS4A8, MSR1, MSRB3, MT1A, MT1E, MT1M, MT2A, MTURN, MUSK, MUSTN1, MXD1, MYCBPAP, MYCT1, MYH10, MYH7, MYO15B, MYO1F, MYO7B, MYOC, MYOCD, MYOM1, MYOZ1, MYRF, MYZAP, N4BP2L1, N4BP2L2, NAALADL1, NABP1, NCF1, NCF2, NCF4, NCKAP5, NCR1, NDNF, NEAT1, NECAB1, NEDD9, NEGR1, NEK10, NEK5, NES, NEXN, NFE2, NKD2, NKX2-1, NLRC4, NLRP3, NME5, NME8, NOSTRIN, NOTCH4, NPNT, NPR1, NR3C2, NR5A2, NRPI, NTNG1, NTNG2, NTRK3, NXF3, OGN, OLR1, OR7E47P, OSCAR, OSR1, OVCH1, P2RY14, PABPC5, PADI4, PAG1, PALD1, PALM, PALM2, PALM2-AKAP2, PALM3, PAPSS2, PARD3B, PCDH12, PCDH15, PDE1C, PDE2A, PDE3B, PDE4D, PDE5A, PDE8B, PDGFRA, PDK4, PDLIM2, PEA1, PEAR1, PEBP4, PECAM1, PEG3, PGM5, PGR, PHACTR1, PHACTR3, PHEX, PHYHD1, PI16, PID1, PILRA, PIP5K1B, PKIG, PKNOX2, PLA1A, PLA2G2A, PLA2G4C, PLAC9, PLCXD3, PLEKHH2, PLEKHO2, PLIN2, PLLP, PLN, PLP1, PLSCR4, PLXNA2, PNMA2, PNPLA2, PODXL, PPBP, PPM1K, PPP1R14A, PPP1R15A, PPP1R32, PRAM1, PRDM16, PRDM5, PRDM6, PRELP, PREX1, PREX2, PRG4, PRICKLE2, PRKCB, PRKCE, PRKCZ, PRKG1, PROK2, PRR15L, PRRG1, PRUNE2, PRX, PTGDS, PTGIR, PTH1R, PTPN21, PTPN22, PTPRB, PTPRC, PTPRM, PTPRQ, PTX3, PXDC1, PZP, RAB17, RAB37, RAB8B, RADIL, RAI2, RALGAP2, RAMP2, RAMP3, RANBP3L, RAPGEF5, RARA, RARRES2, RASGRP2, RASGRP4, RASIP1, RASL12, RASSF1, RASSF2, RBMS2, RBP4, RBPM5, RCAN2, RCSD1, RDH5, REG1A, RERG, RETN, RFX2, RGCC, RGL4, RGS11, RGS13, RGS18, RGS2, RGS22, RGS9, RHOB, RHO, RNASE1, RNASE2, RNASE4, RND1, RNF122, RNF125, RNF146, RNF180, RNPC3, ROBO2, ROBO4, ROPN1L, ROR1, RP1, RPGR, RPS6KA2, RRAD, RRAS, RSPH1, RSPH10B2, RSPO3, RSPO4, RSRP1, RTKN2, RUNX1T1, S100A12, S100A4, S1PR1, SAMHD1, SAMS1, SAXO2, SCARA5, SCARF1, SCEL, SCHIP1, SCML1, SCN3B, SCN4B, SCN7A, SCUBE1, SDPR, SDR16C5, SEC14L6, SELENBP1, SELP, SEMA3B, SEMA3G, SEMA6A, SEMA6B, SEPT4, SERPING1, SERTM1, SESN1, SFTA2, SFTA3, SFTPA1, SFTPA2, SFTPB, SFTPD, SGIP1, SH2D3C, SH3BP5, SH3D19, SHROOM1, SHROOM4,

Tablica 23. Nastavak

SIGLEC11, SIGLEC14, SIGLEC17P, SIGLEC5, SIGLEC9, SIRPB1, SIRPB2, SLA, SLC11A1, SLC14A1, SLC16A12, SLC24A4, SLC25A37, SLC2A3, SLC2A4, SLC34A2, SLC39A8, SLC4A1, SLC5A4, SLC8A3, SLC9A3R2, SLC01C1, SLC02A1, SLC04A1, SLFN14, SLIT2, SLIT3, SMAD9, SMG1P7, SNHG7, SNRK, SNX22, SOD3, SOGA3, SORBS1, SORBS2, SORT1, SOX17, SOX18, SPAG8, SPARCL1, SPATA13, SPATA6, SPATA6L, SPI1, SPNS2, SPOCK2, SPTBN1, SSH2, ST6GALNAC3, ST6GALNAC5, ST6GALNAC6, STAC, STARD13, STARD8, STAT4, STEAP4, STK33, STON1, STON1-GTF2A1L, STX2, STXBP1, STXBP6, SULT1A1, SULT1A2, SULT1B1, SULT1C4, SVEP1, SYNE1, SYNM, SYNPO2, SYT15, SYTL3, TAL1, TARID, TBX2, TBX3, TBX4, TBX5, TBXA2R, TBXAS1, TCEAL2, TCF21, TDDR10, TEC, TEK, TEK11, TEX14, TFEC, TFPI, TG, TGFBR2, TGFBR3, TGM2, THBS1, THSD1, TIMP3, TIMP4, TLR2, TLR4, TM6SF1, TMC2, TMEM100, TMEM125, TMEM150B, TMEM170B, TMEM2, TMEM212, TMEM220, TMEM236, TMEM47, TMEM71, TMOD1, TMPRSS2, TNFAIP8L3, TNFRSF10C, TNFRSF10D, TNFSF12-TNFSF13, TNFSF14, TNF, TNF1, TNXA, TNXB, TOX2, TPK1, TPM2, TPPP3, TPSAB1, TPSB2, TREM1, TRHDE, TRIM58, TRIP10, TRPC3, TSC22D3, TSNAXIP1, TSPAN12, TSPAN19, TSPYL2, TTL6, TTL7, TUBB1, TXNIP, UNC13D, USHBP1, USP53, VAMP5, VEGFC, VEPH1, VGLL3, VIM, VIPR1, VNN2, VNN3, VSIG2, VSTM4, VWA3A, VWA3B, VWF, WAS, WBSCR17, WDR38, WFDC1, WFS1, WHAMMP3, WISP2, WWC2, XPNPEP2, YPEL2, YPEL3, YPEL4, ZBBX, ZBTB16, ZC3H12B, ZEB2, ZFP92, ZFPM2, ZIM2, ZMAT1, ZMYND10, ZMYND15, ZNF331, ZNF366, ZNF385B, ZNF385D, ZNF442, ZNF483, ZNF493, ZNF578, ZNF728, ZNF83, ZNF876P

Kako bi provjerili ulogu gena čija je ekspresija promijenjena u tumorskim uzorcima u obje skupine ponovili smo analizu genske ontologije za kategoriju Biološki procesi. Rezultati genske ontologije pokazuju obogaćenje za 223 pojma genske ontologije za gene sa pozitivnom promjenom razine ekspresije u tumorskim uzorcima. Rezultati ukazuju da su geni sa pozitivnom promijenom razine ekspresije u tumorskim uzorcima uključeni u procese regulacije staničnog ciklusa (GO:0007049, GO:0022402), diobe stanice (GO:0051301, GO:0000280), metaboličke procese koji uključuju DNA (GO:0006259) i procese diferencijacije epitelnih stanica (GO:0030855) (Slika 45.).



Slika 45. Rezultati GO analize za gene sa pozitivnom promjenom u razini ekspresije u obje testirane kohorte

gena sa najvećom negativnom promjenom u razini ekspresije u eksperimentalnoj kohorti, smanjena ekspresija je potvrđena za 16 gena (80%). Smanjena ekspresija CA1, CELA3A, REG1B i PNLIP nije potvrđena u validacijskoj kohorti.

Nadalje, rezultati diferencijalne ekspresije na validacijskoj kohorti pokazuju da se svih 40 identificiranih visokoumreženih gena u eksperimentalnoj kohorti, pojačano eksprimiraju i u validacijskoj kohorti (Tablica 23.).

Tablica 24. Validacija ekspresije visokoumreženih gena identificiranih u eksperimentalnoj kohorti. EK-eksperimentalna kohorta, VK-validacijska kohorta

Geni	Log2FC (EK)	Prilagodena p vrijednost (EK)	Log2FC (VK)	Prilagodena p vrijednost (VK)
AURKB	4,210	$2,02 \times 10^{-7}$	3,7469	$1,34 \times 10^{-289}$
CDCA8	2,680	$2,03 \times 10^{-3}$	3,9829	$1,0 \times 10^{-150}$
CHEK1	3,0230	$1,93 \times 10^{-5}$	3,2078	$1,0 \times 10^{-150}$
DTL	2,5119	$9,02 \times 10^{-5}$	3,7584	$2,31 \times 10^{-299}$
KIF20A	3,9805	$4,90 \times 10^{-9}$	4,0517	$1,78 \times 10^{-252}$
KIF23	3,3977	$5,16 \times 10^{-13}$	3,2107	$1,0 \times 10^{-150}$
MAD2L1	3,0968	$5,59 \times 10^{-7}$	4,2112	$1,0 \times 10^{-150}$
MCM2	4,4010	$2,45 \times 10^{-12}$	4,1558	$1,0 \times 10^{-150}$
MKI67	2,7432	$3,42 \times 10^{-7}$	3,5252	$7,16 \times 10^{-271}$
PLK4	2,6555	$8,81 \times 10^{-7}$	2,9232	$1,0 \times 10^{-150}$
RAD51	2,4204	$3,37 \times 10^{-3}$	3,8205	$1,0 \times 10^{-150}$
ASPM	3,1130	$9,66 \times 10^{-7}$	4,0156	$1,0 \times 10^{-150}$
BIRC5	4,1465	$1,80 \times 10^{-8}$	4,7484	$1,0 \times 10^{-150}$
BUB1B	3,9146	$2,85 \times 10^{-11}$	3,7539	$1,30 \times 10^{-298}$
CCNA2	2,8514	$1,79 \times 10^{-5}$	3,3023	$1,0 \times 10^{-150}$
CDC20	5,2276	$1,27 \times 10^{-13}$	4,5972	$1,0 \times 10^{-150}$
CDC45	6,6200	$5,91 \times 10^{-12}$	4,3992	$1,0 \times 10^{-150}$
CDC6	2,7690	$2,85 \times 10^{-5}$	3,8200	$1,0 \times 10^{-150}$
CDCA5	3,4299	$6,93 \times 10^{-4}$	3,8856	$1,0 \times 10^{-150}$
CDT1	2,3076	$2,69 \times 10^{-2}$	3,5159	$2,24 \times 10^{-273}$

Tablica 24. Nastavak

CENPE	2,3291	$8,95 \times 10^{-5}$	3,3456	$1,0 \times 10^{-150}$
CENPF	4,3412	$7,60 \times 10^{-14}$	4,6327	$1,0 \times 10^{-150}$
ESPL1	3,8580	$8,61 \times 10^{-9}$	2,9904	$4,89 \times 10^{-251}$
FEN1	3,0117	$1,20 \times 10^{-5}$	2,8765	$1,0 \times 10^{-150}$
KIF11	2,3226	$2,99 \times 10^{-5}$	3,9931	$1,0 \times 10^{-150}$
KIF15	2,6985	$7,95 \times 10^{-7}$	3,3893	$1,0 \times 10^{-150}$
KIF2C	2,8063	$5,30 \times 10^{-8}$	4,1223	$1,0 \times 10^{-150}$
MELK	2,4082	$9,17 \times 10^{-4}$	4,5199	$1,0 \times 10^{-150}$
NCAPG	2,8047	$4,30 \times 10^{-7}$	3,4811	$5,86 \times 10^{-306}$
NDC80	3,2180	$1,60 \times 10^{-4}$	3,8466	$1,0 \times 10^{-150}$
RRM2	3,3062	$5,94 \times 10^{-7}$	4,0634	$7,22 \times 10^{-250}$
TOP2A	2,9909	$1,47 \times 10^{-6}$	4,6669	$1,0 \times 10^{-150}$
TTK	2,9952	$2,65 \times 10^{-8}$	4,7322	$1,0 \times 10^{-150}$
UBE2C	2,5404	$5,62 \times 10^{-5}$	4,5825	$1,0 \times 10^{-150}$
CCNB1	2,4968	$5,29 \times 10^{-7}$	4,1672	$1,0 \times 10^{-150}$
CCNB2	2,5522	$7,04 \times 10^{-7}$	4,1882	$1,0 \times 10^{-150}$
CDK1	2,6535	$4,28 \times 10^{-5}$	3,9085	$1,0 \times 10^{-150}$
KIF4A	4,0891	$3,84 \times 10^{-14}$	4,5248	$1,0 \times 10^{-150}$
PLK1	2,2889	$1,26 \times 10^{-2}$	3,5367	$1,0 \times 10^{-150}$
TPX2	3,2062	$5,79 \times 10^{-10}$	4,7961	$1,0 \times 10^{-150}$

Također, smjer promjene ekspresije potvrđen je i za 8 od 10 *seed* gena identificiranih klastera. *Seed* geni čija je pozitivna promjena ekspresije potvrđena u obje kohorte su TRIP13, DQX1, POU2AF1, NCAPD2 i ADAMTS18, dok je negativna promjena ekspresije potvrđena za S1PR1, GNG11 i COL6A6 gene. Gen S100A9 je smanjeno eksprimiran u eksperimentalnoj kohorti dok je u validacijskoj kohorti zabilježena njegova povećana ekspresija. Gen SLX4 je pojačano eksprimiran u eksperimentalnoj kohorti ali nije diferencijalno eksprimiran u validacijskoj kohorti (Tablica 25.).

Tablica 25. Validacija ekspresije *seed* gena u obje testirane kohorte. EK- eksperimentalna kohorta, VK-validacijska kohorta.

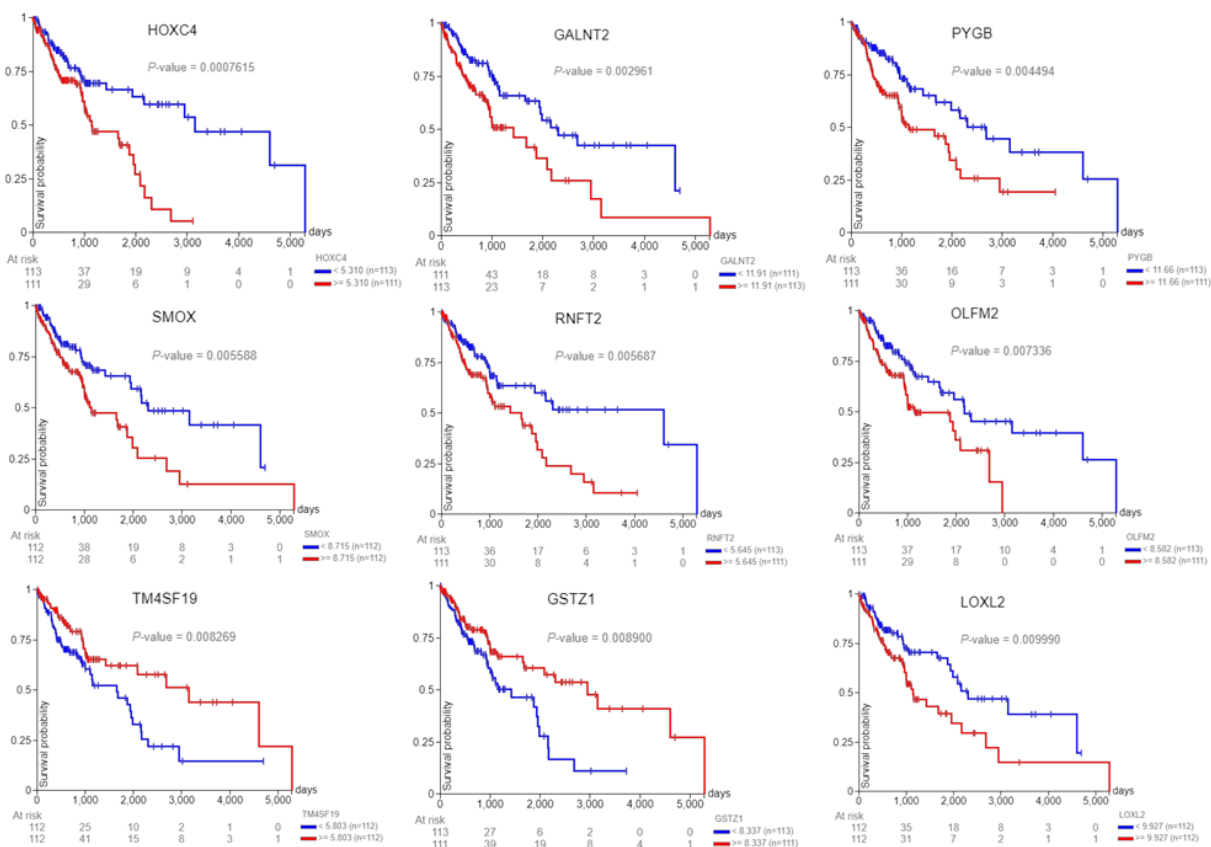
MCODE Klaster	Seed gen	$\log_2FC$ (EK)	Prilagođena p vrijednost (EK)	$\log_2FC$ (VK)	Prilagođena p vrijednost (VK)
Klaster 1	TRIP13	3,48	$6,17 \times 10^{-11}$	4,55	$<1 \times 10^{-150}$
Klaster 2	S1PR1	-3,02	$5,29 \times 10^{-5}$	-3,31	$<1 \times 10^{-150}$
Klaster 3	GNG11	-2,42	$2,38 \times 10^{-3}$	-2,23	$<1 \times 10^{-150}$
Klaster 4	DQX1	4,09	$1,10 \times 10^{-2}$	4,61	$<1 \times 10^{-150}$
Klaster 5	S100A9	-4,00	$2,36 \times 10^{-4}$	1,66	$1,60 \times 10^{-24}$
Klaster 6	COL6A6	-2,53	$1,29 \times 10^{-2}$	-3,31	$1,22 \times 10^{-138}$
Klaster 7	POU2AF1	2,25	$4,05 \times 10^{-2}$	2,29	$5,22 \times 10^{-79}$
Klaster 8	NCAPD2	2,15	$4,17 \times 10^{-9}$	2,68	$<1 \times 10^{-150}$
Klaster 9	SLX4	1,99	$1,56 \times 10^{-2}$	0,18	$2,24 \times 10^{-4}$
Klaster 10	ADAMTS18	3,18	$1,18 \times 10^{-2}$	4,27	$1,76 \times 10^{-118}$
Klaster 11	x	x	x	x	x

Na validacijskoj kohorti smo provjerili i ekspresiju gena identificiranih kao visokoumreženi geni u *sienna3* klasteru koji korelira sa procesom metastaziranja. Od pronađenih 12 visokoumreženih gena u *sienna3* klasteru u eksperimentalnoj kohorti, geni HLA-E, FYN, GIMAP4, GIMAP5, GIMAP7, GIMAP8, KLRF1 i UNC13D su smanjeno eksprimirani u obje kohorte, dok su geni PRF1 i FGL2 povišeno eksprimirani u obje kohorte. Gen STAT5B je smanjeno eksprimiran u validacijskoj kohorti ali je njegova ekspresija u eksperimentalnoj kohorti tik ispod postavljene granične vrijednosti ($\log_2FC=-0,99$), dok je ekspresija HLA-DPB1 gena u validacijskoj kohorti gotovo nepromijenjena u odnosu na zdravo tkivo ($\log_2FC=-0,13$).

4.1.9. Rezultati analize utjecaja razine ekspresije gena na preživljenje

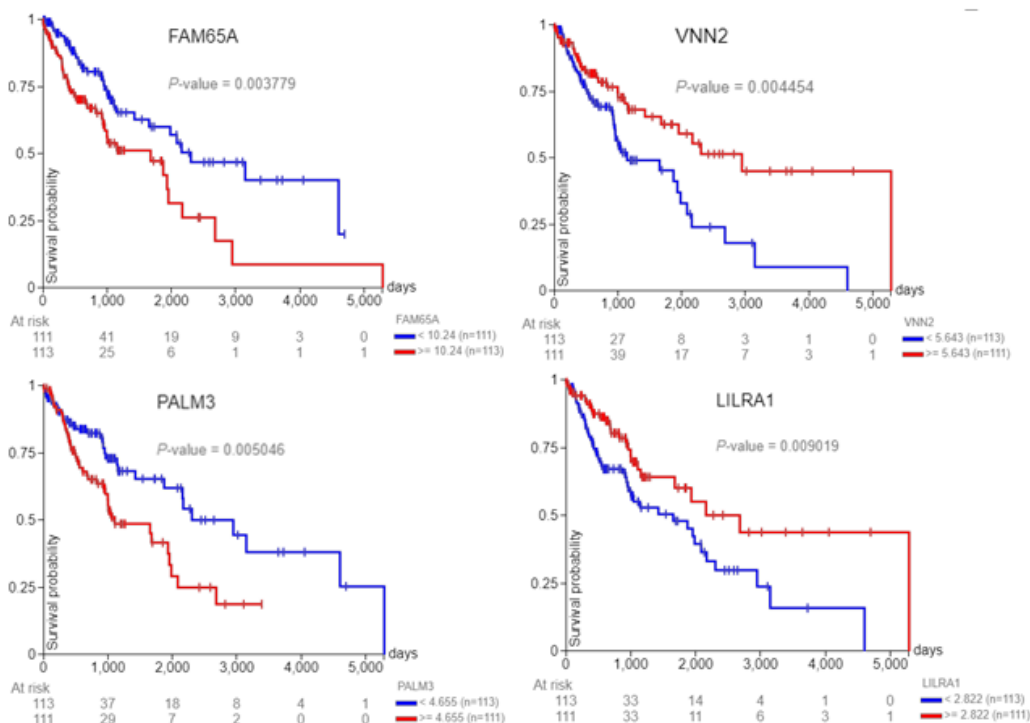
Kako bi provjerili da li su geni sa promijenjenom razinom ekspresije u tumorskim uzorcima povezani sa ukupnim preživljenjem pacijenata oboljelih od tumora pluća skvamoznih stanica, napravljena je analiza sveukupnog preživljenja. Analiza je napravljena za sve gene sa promijenjenom ekspresijom u obje kohorte, njih ukupno 2466, pomoću *Xena* mrežnog alata. Obzirom da je eksperimentalna kohorta relativno mala za ovakvu analizu, analiza je napravljena na validacijskoj TCGA kohorti.

Rezultati ove analize pokazali su da je razina ekspresije 89 gena povezana sa sveukupnim preživljenjem pacijenata, od kojih 57 gena (65 %) ima pozitivno promijenjenu razinu ekspresije, a 31 gen (35 %) ima negativno promijenjenu razinu ekspresije u obje testirane kohorte. Među genima sa pozitivno promijenjenom ekspresijom u tumorskim uzorcima, najznačajniji utjecaj na preživljenje pacijenata ($p < 0,01$) imaju geni HOXC4, GALNT2, PYGB, SMOX, RNFT2, OLFM2, TM4SF19, GSTZ1 i LOXL2. Od navedenih gena, sa lošijom prognozom povezuje se povišena razina ekspresije gena HOXC4 ($p = 7,61 \times 10^{-4}$), GALNT2 ($p = 2,96 \times 10^{-3}$), PYGB ($p = 4,49 \times 10^{-3}$), SMOX ($p = 5,58 \times 10^{-3}$), RNFT2 ($p = 5,68 \times 10^{-3}$), OLFM2 ($p = 7,33 \times 10^{-3}$) i LOXL2 ($p = 9,99 \times 10^{-3}$), te smanjena razina ekspresije gena TM4SF19 ($p = 8,26 \times 10^{-3}$) i GSTZ1 ($p = 8,89 \times 10^{-3}$) (Slika 46.). Među genima koji su smanjeno eksprimirani u obje ispitivane kohorte, najznačajnija povezanost sa preživljenjem pacijenata ($p < 0,01$) pronađena je za gene FAM65A ($p = 3,77 \times 10^{-3}$), VNN2 ($p = 4,44 \times 10^{-3}$), PALM3 ($p = 5,04 \times 10^{-3}$) i LILRA1 ($p = 9,01 \times 10^{-3}$). Od navedenih gena, povećana ekspresija gena VNN2 i LILRA1 povezuje se sa boljom prognozom, dok se povećana ekspresija gena FAM65A i PALM3 povezuje sa lošijom prognozom za pacijente oboljele od tumora pluća skvamoznih stanica (Slika 47.).



Slika 47. Utjecaj ekspresije gena sa povišenom razinom ekspresije u ispitivanim kohortama na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica. Crvene linije predstavljaju ekspresiju gena veću, a plave linije ekspresiju gena manju od medijalne ekspresije gena. P vrijednost izračunata je pomoću log-rank testa, a Kaplan-Meier krivulje vizualizirane su pomoću Xena mrežnog alata.

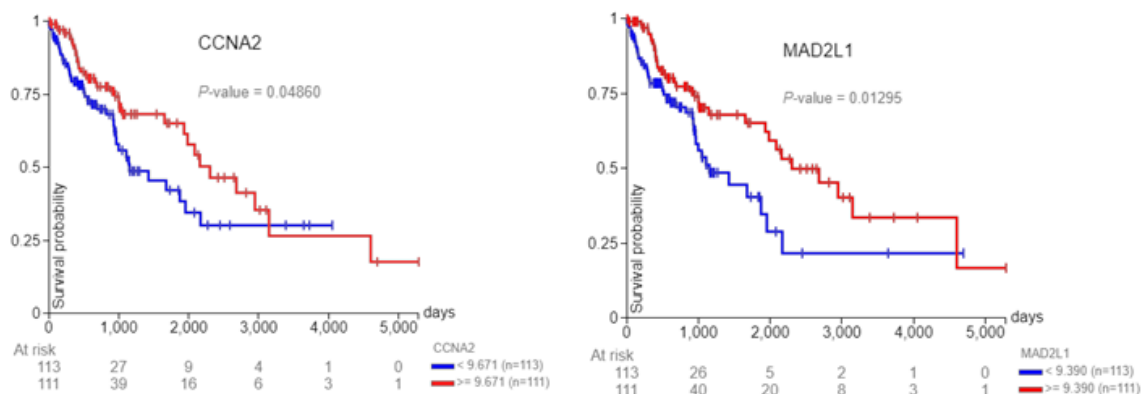
Povezanost ekspresije sa preživljenjem pacijenata nije pronađena za niti jedan od 20 najviše pozitivno promijenjenih gena u eksperimentalnoj kohorti, odnosno 8 gena sa potvrđenom razinom ekspresije u obje skupine ($p > 0,05$). Također, povezanost sa preživljenjem pacijenata nije pronađena niti za gene sa najvećom negativnom promjenom u ekspresiji u obje testirane kohorte ($p > 0,05$).



Slika 48. Utjecaj ekspresije gena sa smanjenom razinom ekspresije u ispitivanim kohortama na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica. Crvene linije predstavljaju ekspresiju gena veću, a plave linije ekspresiju gena manju od medijalne ekspresije gena. P vrijednost izračunata je pomoću log-rank testa, a Kaplan-Meier krivulje vizualizirane su pomoću Xena mrežnog alata.

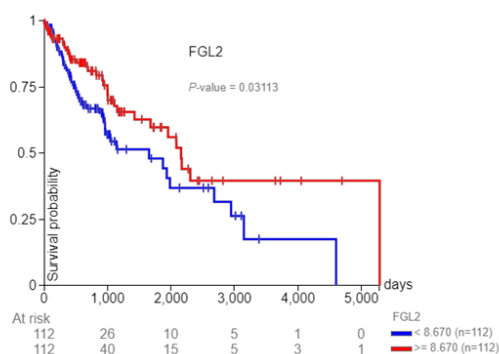
Od 40 gena koji su identificirani kao visokoumreženi geni unutar mreže proteinskih interakcija (PPI mreže), povezanost razine ekspresije sa preživljenjem pacijenata pokazala se za gene CCNA2 ($p=0,048$) i MAD2L1 ($p=0,012$). Povišena ekspresija oba gena ukazuje na povoljniju prognozu za pacijente oboljele od tumora pluća skvamoznih stanica (Slika 49.)

Nadalje, nije pronađena povezanost razine ekspresije gena sa sveukupnim preživljenjem za one gene koji su identificirani kao *seed* geni unutar MCODE klastera ($p>0,05$).



Slika 49. Utjecaj ekspresije visokoumreženih gena u eksperimentalnoj kohorti na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica. Crvene linije predstavljaju ekspresiju gena veću, a plave linije ekspresiju gena manju od medijalne ekspresije gena. P vrijednost izračunata je pomoću log-rank testa, a Kaplan-Meier krivulje vizualizirane su pomoću Xena mrežnog alata.

Također, provjeren je utjecaj razine ekspresije na sveukupno preživljenje za svih 115 gena koji su se grupirali unutar *sienna3* klastera koji je pokazao korelaciju sa procesom metastaziranja. Povezanost razine ekspresije sa sveukupnim preživljenjem pacijenata potvrđena je za gen FGL2 ($p=0,031$), koji je identificiran i kao visokoumreženi gen unutar *sienna3* klastera (Slika 50.).



Slika 50. Utjecaj ekspresije visokoumreženog FGL2 gena u *sienna3* klasteru na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica. Crvene linije predstavljaju ekspresiju gena veću, a plave linije ekspresiju gena manju od medijalne ekspresije gena. P vrijednost izračunata je pomoću log-rank testa, a Kaplan-Meier krivulje vizualizirane su pomoću Xena mrežnog alata.

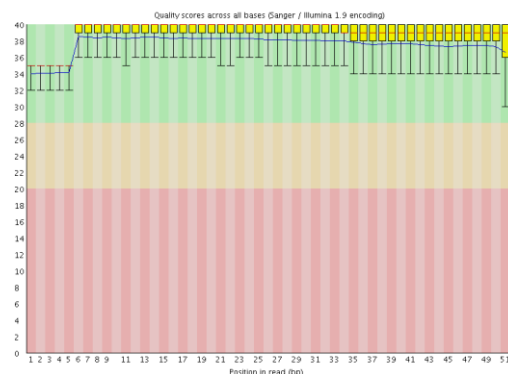
4.2. Rezultati analize malih RNA molekula

Kako bi utvrdili razliku u ekspresijskom profilu miRNA u tumoru pluća skvamoznih stanica, u odnosu na zdrave kontrole, pripremili smo knjižnice za sekvenciranje malih molekula RNA metodom sekvenciranja slijedeće generacije. Knjižnice za analizu su pripremljene za 16 FFPE uzoraka pacijenata oboljelih od tumora pluća skvamoznih stanica i četiri FFPE uzorka zdravih kontrola. Knjižnice su sekvencirane na platformi Illumina HiSeq2500. Dobivenim sekvencama određena je kvaliteta sekvenciranja, potom su im uklonjene sekvence adaptera nakon čega su poravnate na referentni genom GRCh38_p10. Identificirane su miRNA u svakom uzorku pomoću *miRBase* baze podataka, nakon čega je napravljena analiza diferencijalne ekspresije miRNA u tumorskom u odnosu na zdravo tkivo. Potom su određeni ciljni geni miRNA sa promijenjenom razinom ekspresije te su identificirani geni čija je promjena u razini ekspresije suprotna od razine ekspresije miRNA.

4.2.1. Rezultati provjere kvalitete neobrađenih sekvenci i uklanjanje sekvenci adaptera

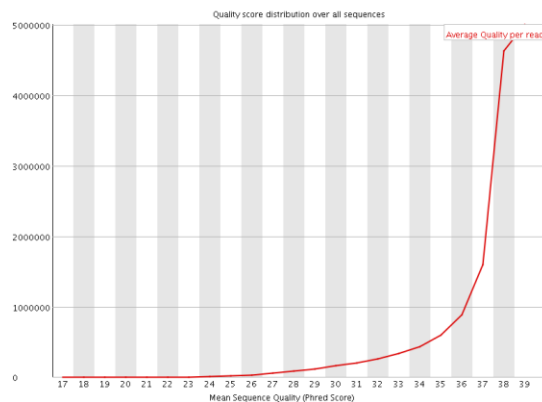
Kao i za transkriptomске podatke, i za miRNA podatke sekvenciranja potrebno je prvo provjeriti kvalitetu očitanih sekvenci dobivenih sekvenciranjem. U našim analizama za procjenu kvalitete neobrađenih sekvenci koristili smo mrežni oblak *BaseSpace* i aplikaciju *FASTQC*.

Rezultati modula *Kvaliteta očitanih sekvenci* pokazuju visoke Phred vrijednosti kvalitete za sve pripremljene i sekvencirane knjižnice. Na Slici 51. prikazan je primjer rezultata ovog modula za očitane sekvence jednog uzorka. Kao i kod transkriptomskih podataka, i na ovom primjeru vidljivo je da su vrijednosti kvalitete nešto niže na početku i na kraju očitane sekvence, što je očekivani rezultat na većini platformi za sekvenciranje, uključujući i platformu HiSeq2500 na kojoj su pripremljene knjižnice sekvencirane. Usprkos tome, vidljivo je da su vrijednosti kvalitete očitane sekvence visoke za sve baze u sekvenci, odnosno da sve baze u očitanoj sekvenci imaju Phred vrijednost kvalitete >Q30.



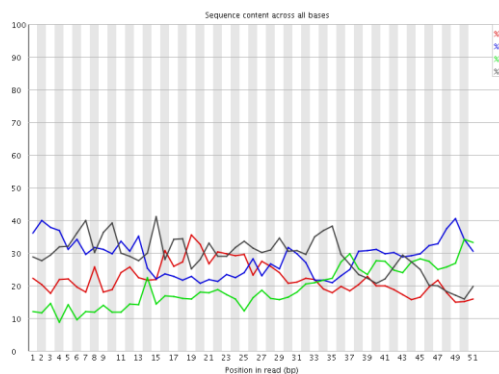
Slika 51. Primjer rezultata u modulu *Kvaliteta očitanih sekvenca po bazama*. Pozadina grafa označava visoku Phred vrijednost kvalitete očitanih sekvenca, žuta boja označava prihvatljive vrijednosti kvalitete, crvena boja označava niske vrijednosti kvalitete očitanih sekvenca.

Rezultati modula *Sažetak vrijednosti kvalitete po očitanoj sekvenci* potvrđuju dobre rezultate modula *Kvaliteta očitanih sekvenca po bazama* za sve uzorke u analizi. Na Slici 52. prikazan je primjer rezultata ovog modula za jednu očitane sekvence u uzorku. Na ovom primjeru vidljivo je da većina sekvenca u uzorku ima Phred vrijednosti kvalitete >35 i da ne postoji podskupina sekvenca sa malim vrijednostima kvalitete.



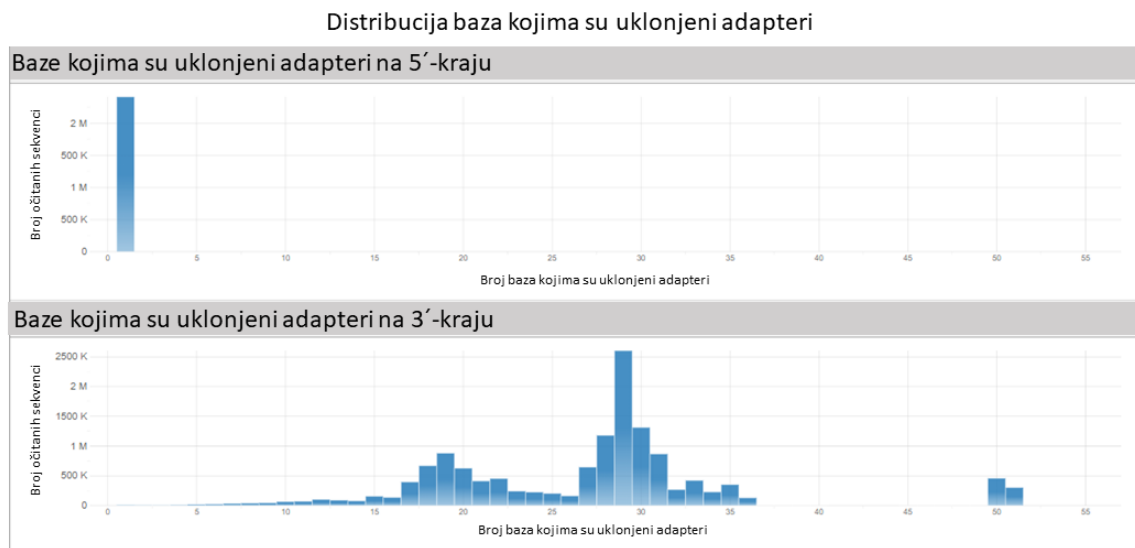
Slika 52. Primjer rezultata za modul *Sažetak vrijednosti kvalitete po očitanoj sekvenci*.

Rezultati modula *Provjera očitanih sekvenca po bazama* ukazuju da baze u sekvenci nisu podjednako zastupljene. U ovom modulu očekivana je podjednaka zastupljenost svake baze u sekvenci. Kao što je vidljivo na primjeru rezultata na Slici 53. u prvih 10-ak baza očekivana je nejednaka zastupljenost baza. Također, vidljivo je da je i u ostatku očitane sekvence zastupljenost svake od četiri baza nije podjednaka. Obzirom da se radi o kratkim sekvencama, duljine 50 nukleotida, kojima nisu uklonjeni adapteri prije provjere kvalitete, ovaj rezultat nije iznenađujući.



Slika 53. Primjer rezultata za modul *Provjera očitanih sekvenci po bazama*

Prije analize diferencijalne ekspresije neobrađenim sekvencama su uklonjene sekvence adaptera. Na Slici 54. je primjer rezultata za jedan uzorak u analizi. Kao što je vidljivo na ovom primjeru, na 5'-kraju uklonjeno je više od 2 milijuna baza na početku očitane sekvence, dok je najveći broj očitanih sekvenci kojima su uklonjeni adapteri na 3'-kraju bio duljine 29 nukleotida.



Slika 54. Primjer rezultata distribucija baza kojima su uklonjeni adapteri. Na y-osi izražen je broj očitanih sekvenci, dok x-os predstavlja poziciju u očitanoj sekvenci na kojoj su adapteri uklonjeni

Distribucija duljine očitanih sekvenci prije i nakon uklanjanja prikazana je na primjeru jednog uzorka na Slici 55. Kao što je vidljivo na primjeru, prije uklanjanja sekvenci adaptera očitane sekvence bile su duljine 51 nukleotid. Nakon uklanjanja sekvenci adaptera vidljivo je da je najveći broj očitanih sekvenci duljine 22 nukleotida, zatim po broju očitanih sekvenci slijede sekvence duljine 20-25 nukleotida te 29-34 nukleotida.



Slika 55. Primjer rezultata distribucije duljine očitanih sekvenci prije i nakon uklonjenih sekvenci adaptera. Na y-osi izražen je broj očitanih sekvenci, dok x-os predstavlja duljinu očitane sekvence izraženu u broju nukleotida

U Tablici 26. prikazan je broj očitanih sekvenci i baza prije i nakon uklanjanja sekvenci adaptera za sve uzorke u analizi.

Tablica 26. Prikaz broja očitanih sekvenci i baza prije i nakon uklanjanja sekvenci adaptera. Brojevi u zagradama označavaju postotak očitanih sekvenci/baza nakon uklanjanja adaptera u odnosu na ukupan broj očitanih sekvenci/baza

Uzorak	Broj očitanih sekvenci prije uklanjanja adaptera	Broj očitanih sekvenci nakon uklanjanja adaptera	Broj baza prije uklanjanja adaptera	Broj baza nakon uklanjanja adaptera
2	14 008 519	13 246 214 (94,56%)	714 434 469	335 642 933 (46,98%)
28	18 411 640	17 502 934 (95,06%)	938 993 640	507 434 972 (54,04%)
44	13 440 817	8 997 405 (66,94%)	685 481 667	220 506 299 (32,17%)
55	14 495 339	13 667 803 (94,29%)	739 262 289	366 689 204 (49,6%)
3260	16 166 997	14 985 865 (92,69%)	824 516 847	415 623 228 (50,41%)
4613	13 005 463	8 485 045 (65,24%)	663 278 613	194 579 646 (29,34%)
4737	10 997 518	6 203 356 (56,41%)	560 873 418	154 976 642 (27,63%)
7202	13 541 485	8 985 418 (66,35%)	690 615 735	211 964 888 (30,69%)
8323	14 176 785	13 583 318 (95,81%)	723 016 035	408 916 334 (56,56%)
12545	16 286 908	10 960 861 (67,3%)	830 632 308	253 549 414 (30,52%)
12698	10 388 789	5 262 658 (50,66%)	529 828 239	127 555 332 (24,07%)
16734	14 887 588	11 418 030 (76,69%)	759 266 988	290 492 688 (38,26%)
17664	16 756 773	9 284 186 (55,41%)	854 595 423	215 512 146 (25,22%)

Tablica 27. Nastavak

18356	15 928 030	2 996 842 (18,81%)	812 329 530	47 557 902 (5,85%)
19694	13 429 397	7 569 066 (56,36%)	684 899 247	157 007 265 (25,55%)
20432	11 632 065	8 478 665 (72,89%)	593 235 315	216 597 925 (36,51%)
21671	12 429 821	9 519 058 (76,58%)	633 920 871	239 097 003 (37,72%)
25486	13 454 297	11 599 477 (86,21%)	686 169 147	299 234 316 (43,61%)
25609	16 296 767	14 998 289 (92,03%)	831 135 117	399 482 907 (48,06%)
25850	15 743 581	9 622 319 (61,12%)	802 922 631	236 545 365 (29,46%)

4.2.2. Rezultati poravnavanja neobrađenih sekvenci na referentni genom *GRCh38_p10* i detekcije miRNA vrsta

Kako bi mogli utvrditi kojim miRNA očitane sekvence pripadaju, potrebno ih je poravnati na referentni genom. U ovu svrhu korišten je mrežni alat *sRNAbench*. Za poravnavanje je odabran referentni genom čovjeka verzije *GRCh38_p10*, a *miRBase* baza podataka za detekciju miRNA. Sažetak rezultata poravnavanja prikazan je u Tablici 27. Broj očitanih sekvenci koje su se poravnale na referentni genom, nakon uklanjanja sekvenci adaptera, je 2-15 milijuna po uzorku. Kao što je vidljivo iz tablice, na referentni genom uspješno je poravnato >75 % sekvenci kojima su uklonjeni adapteri u uzorcima. Udio sekvenci poravnatih na miRNA u uzorcima je različit. Najmanje sekvenci poravnatih na miRNA sekvence imaju uzorci 18356 (0,03 %) i 12698 (8,67 %), zbog čega su uklonjeni iz daljnjih analiza. Udio detektiranih zrelih miRNA u većini preostalih uzoraka iznosi od 30-40 %.

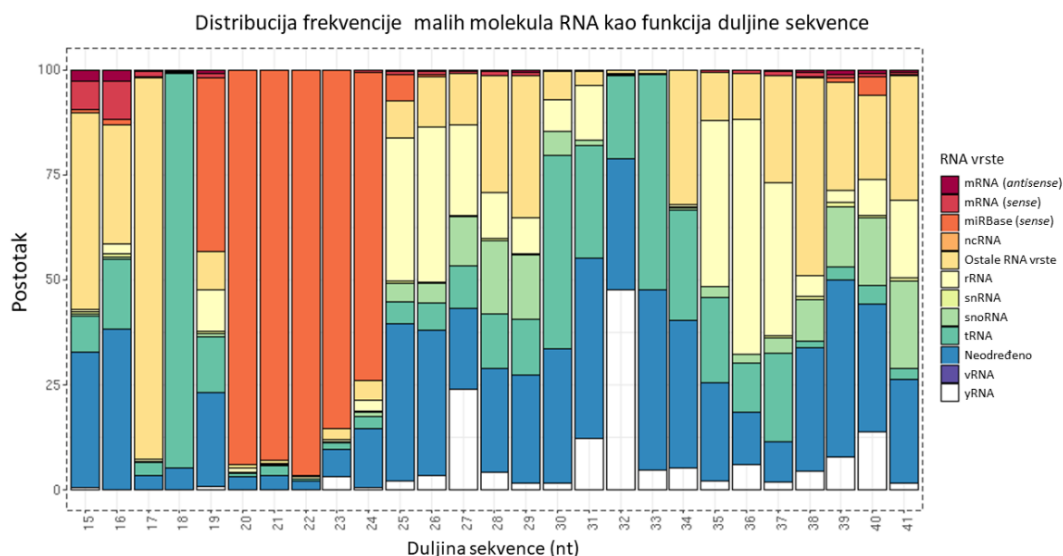
Tablica 27. Sažetak rezultata poravnavanja očitanih sekvenci na referentni genom. Uzorci označeni narančastom bojom uklonjeni su iz daljnjih analiza

Uzorak	Ukupan broj očitanih sekvenci	Sekvence poravnate na genom	Sekvence poravnate na miRNA (miRBase v22)	Detektirane zrele miRNA (% od n u bazi)
4613	13,005,463	7,371,930 (89,86%)	3,860,368 (56,26%)	1006 (37,88%)
25609	16,296,767	11,520,872 (80,27%)	3,930,193 (37,36%)	988 (37,2%)
8323	14,176,785	9,494,740 (75,52%)	1,402,396 (16,74%)	709 (26,69%)
12545	16,286,908	8,364,516 (78,53%)	2,476,238 (33,14%)	851 (32,04%)
16734	14,887,588	9,339,397 (88,0%)	4,744,581 (56,67%)	1051 (39,57%)

Tablica 27. Nastavak

2	14,008,519	11,237,213 (89,06%)	5,666,249 (54,72%)	1018 (38,33%)
25850	15,743,581	7,845,780 (86,86%)	2,484,592 (37,37%)	877 (33,02%)
44	13,440,817	7,817,429 (91,13%)	3,231,155 (47,08%)	869 (32,72%)
4737	10,997,518	5,047,503 (87,89%)	778,354 (21,58%)	709 (26,69%)
18356	15,928,030	2,798,112 (93,74%)	881 (0,03%)	73 (2,75%)
17664	16,756,773	7,896,540 (89,9%)	1,475,364 (25,48%)	782 (29,44%)
28	18,411,640	15,333,389 (90,13%)	4,054,225 (27,41%)	644 (24,25%)
55	14,495,339	11,080,379 (86,8%)	4,931,062 (49,05%)	971 (36,56%)
3260	16,166,997	11,243,105 (77,78%)	2,994,008 (28,22%)	870 (32,76%)
25486	13,454,297	9,249,819 (89,01%)	3,446,200 (44,28%)	898 (33,81%)
7202	13,541,485	7,340,102 (87,46%)	2,198,458 (36,7%)	882 (33,21%)
19694	13,429,397	6,176,268 (85,24%)	3,100,894 (55,67%)	893 (33,62%)
21671	12,429,821	7,735,680 (89,85%)	3,981,371 (58,19%)	941 (35,43%)
12698	10,388,789	3,715,178 (87,44%)	206,468 (8,67%)	472 (17,77%)
20432	11,632,065	6,349,241 (86,65%)	1,096,256 (23,21%)	713 (26,84%)

Kao što je vidljivo iz Tablice 27. udio sekvenci koje su poravnate na miRNA je u većini uzoraka manji od 50 %. Sekvence koje se nisu poravnale na miRNA, poravnate su i prepoznate kao sekvence mRNA (*sense / antisense*), nekodirajuće RNA, ribosomalne RNA, male jezgrene RNA, transportne RNA, virusne RNA ili yRNA. Na Slici 56. prikazan je primjer distribucije poravnatih malih molekula RNA u jednom uzorku. Kao što je vidljivo sa slike, većina sekvenci duljine 19-24 nukleotida identificirano je kao miRNA vrste.

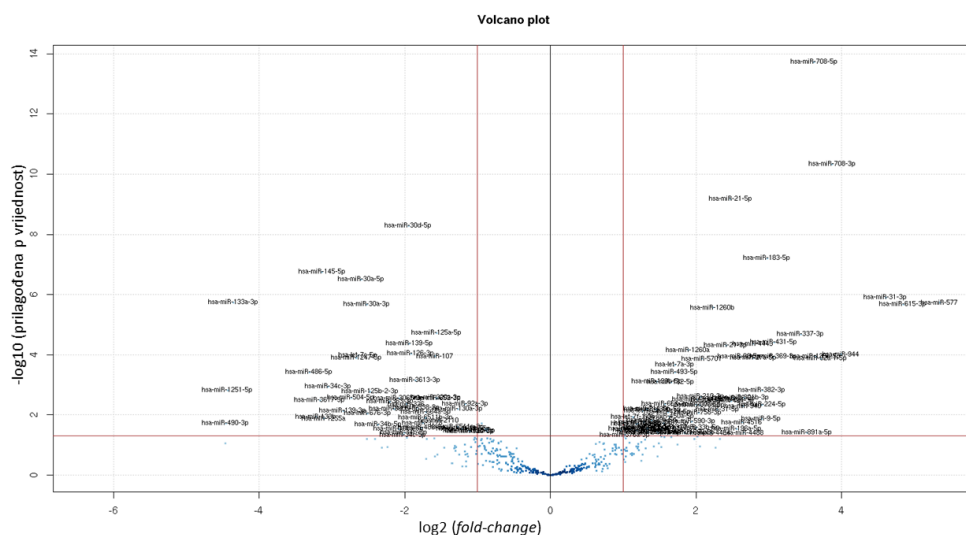


Slika 56. Distribucija malih molekula RNA na primjeru poravnanja sekvenci iz jednog uzorka. ncRNA-nekodirajuće RNA, rRNA- ribosomalna RNA, sn(o)RNA- male jezgrene RNA, tRNA-transportna RNA, vRNA- virusne RNA

4.2.3. Rezultati analize diferencijalne ekspresije miRNA

Kako bi ispitali profil ekspresije miRNA koji doprinosi fenotipu tumora pluća skvamoznih stanica, u odnosu na zdrave kontrole, napravljena je analiza diferencijalne ekspresije. Za ovu analizu korištena je matrica kvantificiranih očitavanja na razini miRNA i mrežni alat *sRNAdet*. miRNA su okarakterizirane kao diferencijalno eksprimirane na temelju postavljene granične vrijednosti $|\log_2 FC| \geq 1$ i prilagođene statističke značajnosti $p < 0,05$, prema metodi Benjamin-Hochberg [4].

Rezultati analize diferencijalne ekspresije miRNA pokazali su da 122 miRNA u uzorcima tumora pluća skvamoznih stanica ima promijenjenu razinu ekspresije u odnosu na zdrave kontrole. Od 122 miRNA, 75 miRNA ima pozitivnu, a 47 miRNA ima negativno promijenjenu razinu ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole. Na Slici 57. prikazani su rezultati diferencijalne ekspresije u obliku *volcano* plot. Kao što je vidljivo sa slike, statistički najznačajnije promijene detektirane su za hsa-miR-708-5p ($p = 1,89 \times 10^{-14}$), hsa-miR-708-3p ($p = 4,70 \times 10^{-11}$), hsa-miR-21-5p ($p = 6,81 \times 10^{-10}$), hsa-miR-30d-5p ($p = 5,19 \times 10^{-9}$) i hsa-miR-183-5p ($p = 6,15 \times 10^{-8}$).



Slika 57. Rezultati diferencijalne ekspresije miRNA, prikazani u obliku *volcano* grafa. Crvene vertikalne linije prikazuju granične vrijednosti promjene ekspresije (*fold-change*), a horizontalne granične vrijednosti prilagođene p vrijednosti na $-\log_{10}$ skali. Plave točke na grafu označavaju miRNA-e čija promjena u razini ekspresije nije značajna bilo zbog većih vrijednosti od postavljene granične vrijednosti za p prilagođenu vrijednost ili zbog manje promjene u razini ekspresije od postavljenih graničnih vrijednosti

20 miRNA-i, 10 sa najvećom pozitivnom i 10 sa najvećom negativnom promijenom u razini ekspresije prikazano je u Tablici 28.

Tablica 27. Prikaz 10 najviše promijenjenih miRNA sa pozitivnom (A) i negativnom (B) promijenom u razini ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole. Narančastom obojom označene su miRNA sa pozitivnom promjenom u razini ekspresije, dok su miRNA sa negativnom promjenom u razini ekspresije označene plavom bojom.

A)

miRNA	log2FC	Prilagođena p vrijednost
hsa-miR-577	5,34	$1,82 \times 10^{-6}$
hsa-miR-615-3p	4,84	$2,16 \times 10^{-6}$
hsa-miR-31-3p	4,60	$1,29 \times 10^{-6}$
hsa-miR-944	3,99	$9,54 \times 10^{-5}$
hsa-miR-708-3p	3,87	$4,70 \times 10^{-11}$
hsa-miR-92a-1-5p	3,71	$1,34 \times 10^{-4}$
hsa-miR-135b-5p	3,64	$1,14 \times 10^{-4}$
hsa-miR-708-5p	3,62	$1,89 \times 10^{-14}$
hsa-miR-891a-5p	3,52	$3,82 \times 10^{-2}$
hsa-miR-337-3p	3,43	$2,14 \times 10^{-5}$

B)

miRNA	log2FC	Prilagođena p vrijednost
hsa-miR-490-3p	-4,45	$1,89 \times 10^{-2}$
hsa-miR-1251-5p	-4,42	$1,53 \times 10^{-3}$
hsa-miR-133a-3p	-4,35	$1,82 \times 10^{-6}$
hsa-miR-486-5p	-3,31	$3,82 \times 10^{-4}$
hsa-miR-133b	-3,21	$1,14 \times 10^{-2}$
hsa-miR-3617-5p	-3,16	$3,23 \times 10^{-3}$
hsa-miR-145-5p	-3,13	$1,75 \times 10^{-7}$
hsa-miR-1255a	-3,10	$1,30 \times 10^{-2}$
hsa-miR-34c-3p	-3,04	$1,13 \times 10^{-3}$
hsa-miR-139-3p	-2,84	$7,21 \times 10^{-3}$

4.2.4. Rezultati predikcija ciljnih gena miRNA sa promijenjenom razinom ekspresije i interakcije mRNA-miRNA

Kako bi razumijeli funkciju miRNA sa promijenjenom ekspresijom u tumorskim uzorcima, u odnosu na zdrave kontrole, tražili smo gene na koje miRNA mogu utjecati (engl. *target genes*) pomoću multiMir paketa u R programu. Ako je interakcija miRNA-gen validirana eksperimentalno u bazama podataka, ili je interakcija miRNA-gen predviđena računalno u više od tri baze podataka integriranih u multiMir paketu, smatrali smo da interakcija zaista postoji. Pretraga validiranih interakcija miRNA-gen temeljila se na tri baze podataka: *miRecords*, *miRTarBase* i *TarBase*. Pretraga predviđenih interakcija miRNA-gen temeljila se na 8 baza podataka: DIANA-microT, ElMMO, MicroCosm, miRanda, miRDB, PicTar, PITA i TargetScan.

Pretragom navedenih baza podataka, za miRNA sa pozitivnom promijenom u razini ekspresije, pronašli smo 14050 *target* gena, a za miRNA sa negativnom promijenom u razini ekspresije 12499 *target* gena. Obzirom da se radi o velikom broju identificiranih gena, prikazani su rezultati samo za dvije miRNA sa najvećom promijenom u razini ekspresije, hsa-miR-577 i hsa-miR-139-3p (Tablica 29.)

Tablica 29. Pronađeni *target* geni za has-miR-577 i has-miR-139-3p

miRNA	Predviđeni <i>target</i> geni
hsa-miR-139-3p (217 <i>target</i> gena)	ADAMTS1, AGPAT3, AHS2P, AKT1S1, ALG14, ALOX15, ASNS, ATF6, ATG2A, ATP5F1A, B2M, B4GALT5, BARHL1, BCAR1, BCR, BMP1A, BPTF, BRD4, BTBD7, C14orf180, C1orf198, CCAR2, CCDC117, CCDC36, CCDC71L, CDC37, CDK16, CECR2, CELF1, CERCAM, CHD7, CHEK1, CNBP1, CREBBP, CSF3, CTDSP1, DAD1, DARS1, DBNDD1, DDX19A, DNAL1, DYRK1A, E2F3, EEF1D, EHD3, ELAVL1, EN1, FAM102B, FAM162A, FAM184B, FAM83H, FN3K, FOXC1, GDF5OS, GET4, GFRA1, GIGYF1, GZMM, H2AC12, H2BC15, HAND2, HIPK1, HLA-A, HMGB2, HNRNPC, HOXB6, HOXB9, IER5, IGF2BP1, JUND, KANK2, KLHL21, KMT2D, KRT80, LARP4B, LDLR, LIMD2, LNX2, LRPPRC, LRRC45, LRTOMT, LYPLA2, MAP2K3, MAP4, MAPK1, MBD3, MCU, MIDN, MIEF1, MIF, MLLT6, MMP11, MRPL9, MTDH, MTHFD1, MYO10, NCOR2, NFIX, NLGN4X, NOB1, NPTX1, NUDT21, OGA, ORAI2, PALD1, PARD6G-AS1, PAX5, PDGFRA, PGAM5, PIWIL2, PKIA, PLBD2, PNPO, PPP3R1, PRKCSH, PSEN1, PSIP1, PSMD14, PTPN6, PVR, PXN, QPCTL, RAB11B, RABL6, RASD1, RBL2, RBM8A, RBMS3, REEP3, RND2, RNF149, RNF19A, RNF216, RPA1, SDC3, SDCBP, SETD5, SIN3CAF, SLC16A7, SLC1A5, SLC38A4, SLC43A2, SLC5A6, SLC7A5, SMC4, SMIM7, SMTN, SNRNP27, SORBS2, SORL1, SOX4, SP1, SPCS3, SPEN, SPPL3, SPSB1, SREBF2, SSBP2, SUN2, SUSD1, SYT7, TAOK1, TBX20, TERF2, TEX22, THOP1, THY1, TLE3, TM9SF3, TMEM170A, TMEM33, TP53INP1, TRIM71, TUBD1, UBE2G1, UCP2, URGCP-MRPS24, USP22, VPS33A, VWA1, WDR13, WDR45B, WNK1, YAE1, YY1, ZBTB7A, ZDHHC24, ZIC2, ZMIZ2, ZNF217, ZNF486, ZNF527, ZNF550, ZNF570, ZNF580, ZNF585B, ZNF589, ZNF711, ZSWIM7, ZYG11A, AMDHD2, COX8A, BAZ2B

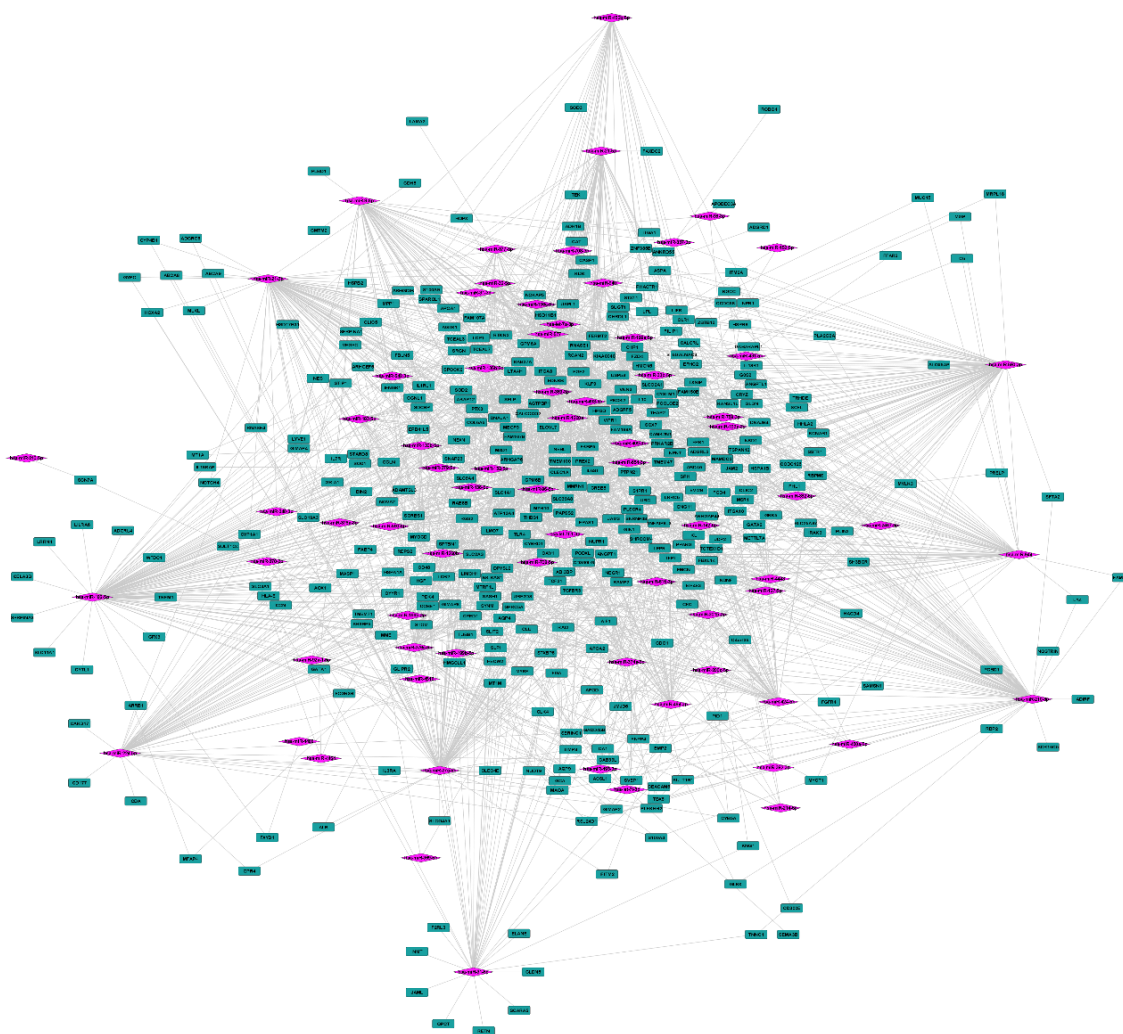
Tablica 29. Nastavak

hsa-miR-577 (417 <i>target</i> gena)	ARPP19, CCNE2, CLIC4, DCTN5, IER3IP1, IGF2BP1, LMNB2, MAPK6, MFSD9, MTRF1, PARP1, PMAIP1, PXK, RAN, SF1, SNRPB, TNRC6A, TSNAX, VGLL4, WEE1, ZNF623, ZNFX1, SRP54, AAK1, ABHD2, ACIN1, ACSL4, ACVR2B, ADAL, ADD3, AIP, AK4, AKAP11, AKAP12, ALPK3, ANKRD17, ANKRD6, ANO6, ANXA4, AP1B1, AP3M2, ARHGAP39, ARID3A, ARID5B, ARL15, ASB1, ASH1L, ATG12, ATL2, ATMIN, ATN1, ATP2A2, ATP2B4, ATP5F1C, ATP6V1G1, ATP8B2, ATRNL1, AZIN1, BAG2, BCL2, BCL2L11, BLZF1, BRIX1, BRWD3, BVES, C12orf10, C12orf57, C17orf75, C4orf46, CAMK2N1, CAMLG, CAPZA1, CASC4, CASP3, CCDC82, CCN2, CCND2, CCNT2, CDC20, CDCA4, CDK4, CDS2, CENPN, CERT1, CGAS, CHD9, CIP2A, CKAP2, CLCN3, CLEC12B, CLEC14A, CLEC2D, CLIP4, CMPK1, CNEP1R1, CNN3, COX10, CPEB4, CREBRF, CSDE1, CSTB, CTDSP12, DCK, DDX21, DDX52, DGKH, DHX40, DIAPH3, DNAJB4, DNAJB9, DNAJC16, DOCK7, DYRK2, E2F3, E2F5, EBPL, EDA2R, EFR3A, EIF1AX, EIF2S1, EIF4G2, ELF3, ELK4, ELL2, ENPP2, ENPP4, ENTPD7, EPDR1, EPHA7, EPN2, ERO1A, ETFDH, EVI5, FAM135A, FAM217B, FAM3C, FBXO8, FBXW2, FDX1, FEM1C, FGF2, FGF21, FKBP14, FKBP1A, FKTN, FNDC3B, FOXP1, FRS2, G3BP1, G6PC, GALNT1, GALNT3, GCC2, GCSAM, GDAP2, GLCC11, GNPTAB, GOLPH3, GRB10, GTPBP2, GXYLT2, H2AC6, H2BC12, HBS1L, HELZ, HEXA, HIC2, HIPK1, HNRNPA1, HNRNPA2B1, HNRNPH3, HNRNPU, HOXA3, HP1BP3, HPRT1, HSDL2, HSPA13, HSPB2, ID4, IFI27L1, IFNAR1, IGF1R, IGSF1, INTS7, IPO7, KANSL1, KAT6B, KBTBD2, KBTBD6, KCTD15, KDM7A, KHDCC4, KLHDC10, KLHL15, KMT2C, KMT5C, KPNA1, KPNA6, LANCL1, LARP7, LATS1, LDHA, LDHA, LLPH, LRRC58, LRRC8D, LRRFIP2, LSM14B, LUC7L2, MAFG, MAP3K2, MBD5, MCL1, MDH2, MDM2, MED23, METAP1D, MFSD14A, MIER3, MLEC, MMADHC, MRPL17, MRPS10, MSANTD2, MSANTD4, MSH6, MTAP, MYC, MYH9, N4BP2L2, NDUFAF3, NDUFS3, NDUFS3, NEDD1, NFIB, NHLRC2, NHSL2, NLRP1, NNT, NPAT, NSD1, NSD2, NUDT21, NUP58, OGDH, OGT, OTUD3, OTUD4, PANK3, PARD6B, PARP2, PARBP, PATL1, PCDHGA8, PCGF5, PCNP, PDS5B, PGBD2, PGBD4, PGRMC1, PHF21A, PIK3CA, PLAGL2, PLCB1, PLP2, PLS1, POC1B-GALNT4, POGK, POGZ, PPP1CB, PPP1R12A, PRKACB, PROK2, PSMD12, PSME4, PSMG1, PTK2B, PTN, PTP4A1, PURB, PWWP3A, RAB14, RAB25, RAB2A, RABGAP1, RANBP6, RBM12, RBM26, RBM48, RC3H2, RCC1, RELCH, RFT1, RFX7, RHOA, RIMBP2, RNF122, RNF138, RNF38, RORA, RPP14, RPRD2, RPS6KA5, RSBN1, RTN4, SART3, SCAF4, SDC2, SEC62, SECISBP2L, SEMA7A, SEMA7A, SERTAD3, SESN3, SETD7, SFT2D2, SH3GLB1, SH3TC2, SIPA1L1, SLC10A7, SLC24A4, SLC25A3, SLC25A36, SLC5A6, SLFN11, SMG1, SMG7, SMU1, SNRK, SNX16, SNX5, SOCS4, SOD2, SOD2, SORBS1, SRSF1, SRSF11, SSBP2, SSR3, STK24, STYX, SYNE3, SZRD1, TAF1D, TAF4, TAF5L, TAL1, TBPL1, TEAD1, TEAD2, TM9SF3, TMEM101, TMEM133, TMEM59, TMEM64, TMEM98, TMEM9B, TNKS2, TNRC6A, TOR1AIP1, TP53INP1, TSEN15, TUBG1, TWF1, TWSG1, UBE2D3, UBE2G1, UBN2, UBQLN1, UMAD1, UQCC1, USO1, USP37, VCIPI1, VDAC2, VEGFC, WAC, WDR26, WDYHV1, WSB2, XPO4, YWHAH, ZBTB11, ZBTB18, ZBTB20, ZBTB22, ZBTB22, ZBTB22, ZBTB22, ZBTB22, ZBTB37, ZBTB4, ZCCHC2, ZCCHC24, ZFYVE21, ZNF211, ZNF217, ZNF256, ZNF396, ZNF431, ZNF451, ZNF587B, ZNF592, ZNF598, ZNF614, ZNF623, ZNF624, ZNF648, ZNF791, ZNF805, ZNF92, ZNRF2, ZSWIM6, ZWINT, SAMSN1, PDXK, C7orf13
---	---

Kako bi bolje razumijeli interakcije između gena i miRNA sa promijenjenom ekspresijom u eksperimentalnoj kohorti, odabrali smo one gene sa smanjenom ekspresijom koji se nalaze i među identificiranim *target* genima miRNA sa povišenom ekspresijom u tumorima. Također, odabrani su i geni sa pozitivnom promjenom u razini ekspresije, ukoliko se nalaze među identificiranim *target* genima miRNA koje imaju smanjenu ekspresiju. Napravljene su dvije mreže interakcija

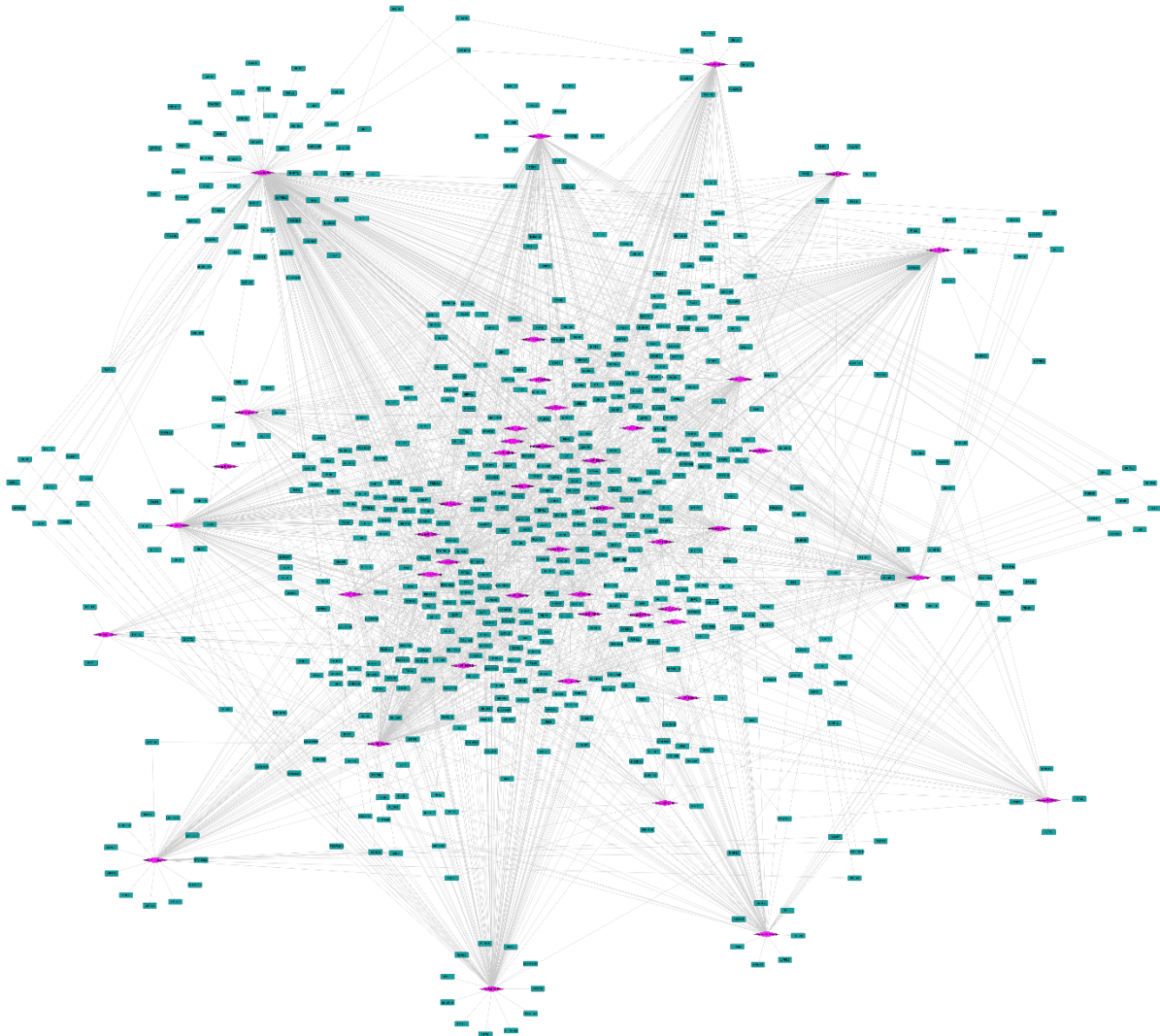
mRNA-miRNA; prva mreža uključuje miRNA sa pozitivnom promjenom ekspresije i mRNA sa negativnom promjenom ekspresije, dok druga mreža uključuje miRNA sa negativnom promjenom ekspresije i mRNA sa pozitivnom promjenom ekspresije.

Kao što je vidljivo sa Slike 58. mreža interakcija miRNA sa pozitivnom i mRNA sa negativnom promjenom razine ekspresije sastoji se od 72 miRNA i 363 *target* gena povezanih sa 1839 interakcija. Za tri miRNA (hsa-miR-891a-5p, hsa-miR-147b-3p i hsa-miR-5701) nisu pronađene interakcije sa mRNA.



Slika 58. Prikaz interaktoma miRNA sa pozitivnom i mRNA sa negativnom promijenom ekspresije u uzorcima skvamoznog tumora pluća. miRNA označene su ljubičastom bojom, a ciljni geni zelenom bojom. Sive linije ukazuju na interakcije gena i miRNA.

Mreža interakcija miRNA sa negativnom i mRNA sa pozitivnom promjenom razine ekspresije sastoji se od 46 miRNA i 618 *target* gena povezanih sa 2003 interakcija (Slika 59.). Za hsa-miR-3178 nisu pronađene interakcije sa mRNA.



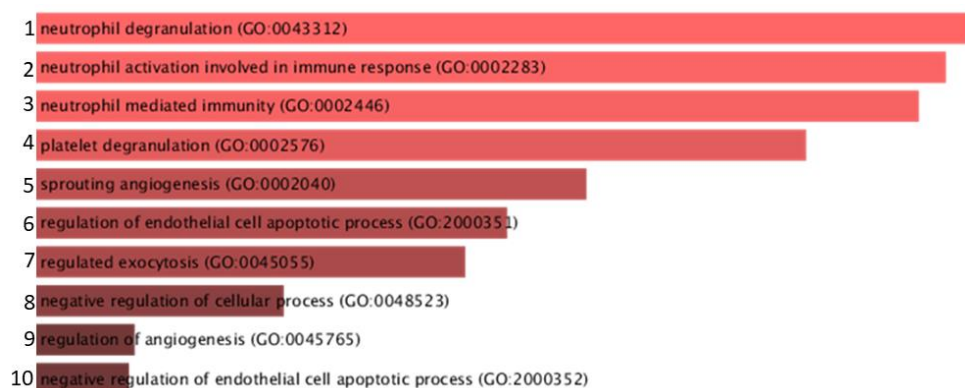
Slika 59. Prikaz interaktoma miRNA sa negativnom i mRNA sa pozitivnom promijenom ekspresije u uzorcima skvamoznog tumora pluća. miRNA označene su ljubičastom bojom, a ciljni geni zelenom bojom. Sive linije ukazuju na interakcije gena i miRNA.

Za miRNA sa najvećom promjenom u razini ekspresije, hsa-miR-577, predviđena je interakcija sa 16 gena sa smanjenom ekspresijom, poput CAV1, KLF9 i VEGFC. Za miR-490-3p, miRNA sa najvećom negativnom promjenom ekspresije, predviđena je interakcija sa 17 gena sa povišenom ekspresijom, uključujući SAPCD2, GCNT3, RAP2B, ZYG11A i dva visokoumrežena gena iz PPI

mreže, CDK1 i RAD51. Među miRNA sa pozitivnom promjenom ekspresije, hsa-miR-182-5p i hsa-miR-590-3p imaju najveći broj interakcija sa genima sa smanjenom ekspresijom (94 i 50 *target* gena). Među miRNA sa negativnom promjenom ekspresije, hsa-miR-107 i hsa-let-7c-5p imaju najveći broj interakcija sa genima sa povećanom ekspresijom (351 i 175 *target* gena).

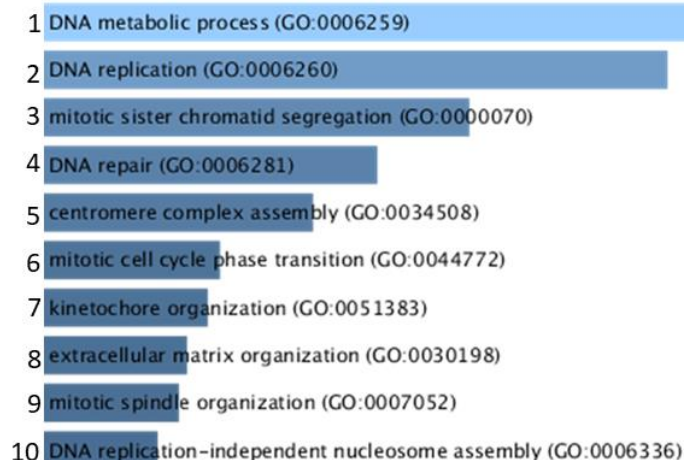
Velik broj interakcija navedenih 4 miRNA (hsa-miR-182-5p, hsa-miR-590-3p, hsa-miR-107 i hsa-let-7c-5p) sa *target* genima može ukazivati na njihovu potencijalno važnu ulogu u tumoru pluća skvamoznih stanica.

Radi boljeg razumijevanja uloge gena koji su promijenjeno eksprimirani u suprotnom smjeru od miRNA, napravili smo analizu genske ontologije, za kategoriju Biološki procesi, pomoću mrežnog alata *Enrichr*. Kao što je vidljivo sa Slike 60., geni sa negativno promijenjenom razinom ekspresije, koji su ciljni geni miRNA sa povišenom razinom ekspresije, imaju ulogu u regulaciji imunološkog odgovora. Primjerice, sudjeluju u procesu degranulacije i aktivacije neutrofila, ali i u regulaciji angiogeneze i apoptoze endotelних stanica.



Slika 60. Analiza genske ontologije za ciljne gene miRNA sa pozitivnom promjenom u razini ekspresije koji su promijenjeno eksprimirani u suprotnom smjeru. Analiza je napravljena za kategoriju Biološki procesi. Obogaćeni pojmovi prikazani su prema p-vrijednostima. Prijevod GO pojmova: 1. Degranulacija neutrofila; 2. Aktivacija neutrofila uključenih u imunološki odgovor; 3. Imunološki odgovor posredovan neutrofilima; 4. Degranulacija trombocita; 5. Početak procesa angiogeneze; 6. Regulacija apoptoze endotelnih stanica; 7. Regulacija egzocitoze; 8. Negativna regulacija staničnih procesa; 9. Regulacija angiogeneze; 10. Negativna regulacija apoptoze endotelnih stanica

Kao što je vidljivo sa Slike 61., geni sa pozitivno promijenjenom razinom ekspresije, koji su ciljni geni miRNA sa negativno promijenjenom razinom ekspresije, imaju ulogu u replikaciji DNA, DNA popravku i procesima povezanim sa regulacijom staničnog ciklusa.



Slika 61. Analiza genske ontologije za ciljne gene miRNA sa negativnom promjenom u razini ekspresije koji su promijenjeno eksprimirani u suprotnom smjeru. Analiza je napravljena za kategoriju Biološki procesi. Obogaćeni pojmovi prikazani su prema p-vrijednostima. Prijevod GO pojmova: 1. Metabolički procesi DNA; 2. Replikacija DNA; 3. Razdvajanje sestrinskih kromatida; 4. Popravak DNA; 5. Sastavljanje centromernog kompleksa; 6. Tranzicija u različite faze mitoze; 7. Organizacija kinetohora; 8. Organizacija vanstaničnog matriksa; 9. Organizacija diobenog vretena; 10. Sastavljanje nukleosoma neovisno o replikaciji DNA

4.3. Rezultati analize proteoma

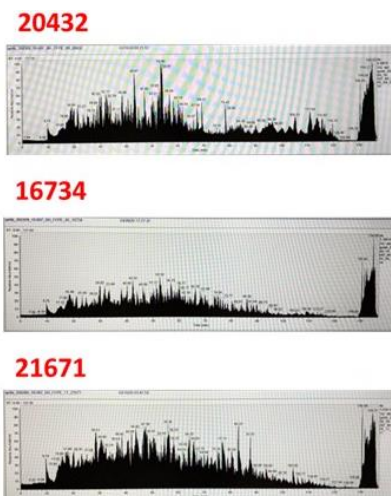
U svrhu profiliranja ekspresije proteina u tumoru pluća skvamoznih stanica i zdravih kontroli, napravljena je analiza proteoma, na razini pilot studije, pomoću spektrometrije masa. U proteomske analize uključena su tri uzorka tumora pluća skvamoznih stanica i tri zdrave kontrole. Iz svakog uzorka izolirani su ukupni proteini u tri biološke replike. Proteinima je određena koncentracija te su pripremljeni za analizu na spektrometru masa. Spektri omjera mase i naboja korišteni su za identifikaciju proteina u uzorcima. Potom su identificirani proteini čija je razina ekspresije promijenjena u tumorskim, u odnosu na zdravo tkivo.

4.3.1. Rezultati identifikacije proteina i provjere kvalitete podataka

Spektri omjera mase i naboja peptida (m/z omjeri), dobiveni analizom na spektrometru masa, korišteni su za identifikaciju proteina pomoću *MaxQuant* programa i referentnog ljudskog proteoma iz Uniprot baze podataka. *MaxQuant* program, osim identifikacije proteina, omogućava i njihovu kvantifikaciju u uzorcima izraženu pomoću *LFQ* vrijednosti intenziteta. *LFQ* (engl. *label-free quantification*) intenziteti su normalizirane vrijednosti količine pojedinog proteina u uzorku. Ove vrijednosti korištene su za provjeru kvalitete podataka dobivenih spektrometrijom masa.

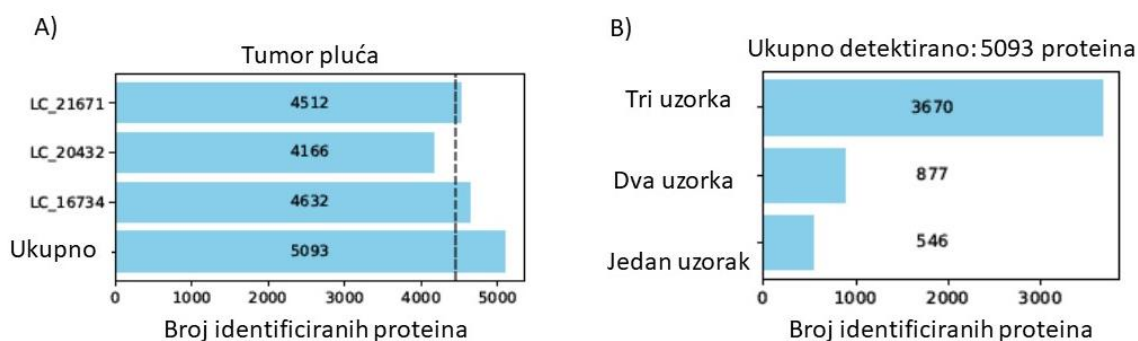
Provjera kvalitete podataka uključivala je provjeru distribucije *LFQ* vrijednosti intenziteta, korelaciju *LFQ* vrijednosti između bioloških replika pojedinih uzoraka te PCA analizu (engl. *Principal component analysis*). Provjera kvalitete podataka napravljena je pomoću programa *Perseus*.

Za identifikaciju proteina pomoću *MaxQuant* programa, korišteni su omjeri mase i naboja peptida, koji su dobiveni analizom na spektrometru masa. Primjer spektara dobivenih spektrometrijom masa prikazani su na Slici 62.



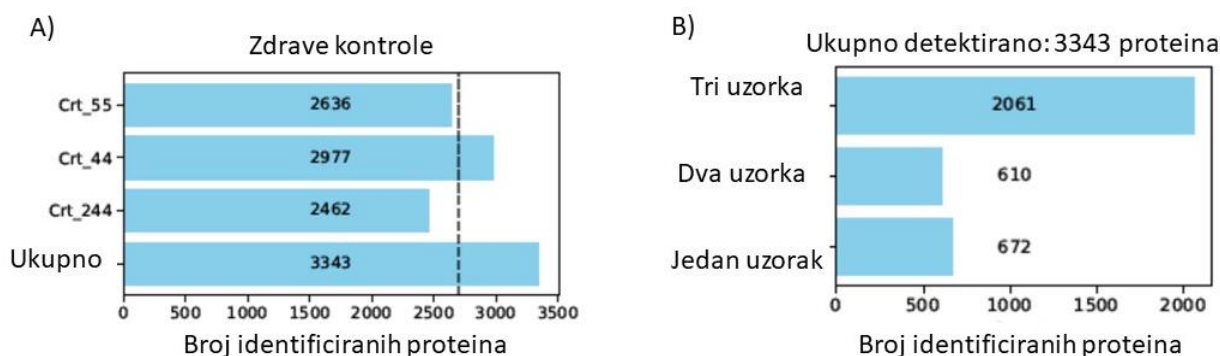
Slika 62. Primjer spektara mase i naboja dobivenih analizom na spektrometru masa.

U tumorskim tkivima identificirano je ukupno 5093 proteina, od kojih je 4632 identificirano u uzorku 16734; 4166 proteina u uzorku 20432 i 4512 proteina u uzorku 21671 (Slika 63. A)). Od ukupno 5093 identificiranih proteina, 3670 proteina identificirano je u sva tri tumorska uzorka (72 %), 877 proteina detektirano je u dva tumorska uzorka (17,2 %), dok je 546 proteina detektirano u samo jednom tumorskom uzorku (10,7 %) (Slika 63. B))



Slika 63. Prikaz ukupnog broja identificiranih proteina (A) i broja identificiranih proteina u uzorcima (B) spektrometrijom masa u tumorskom tkivu.

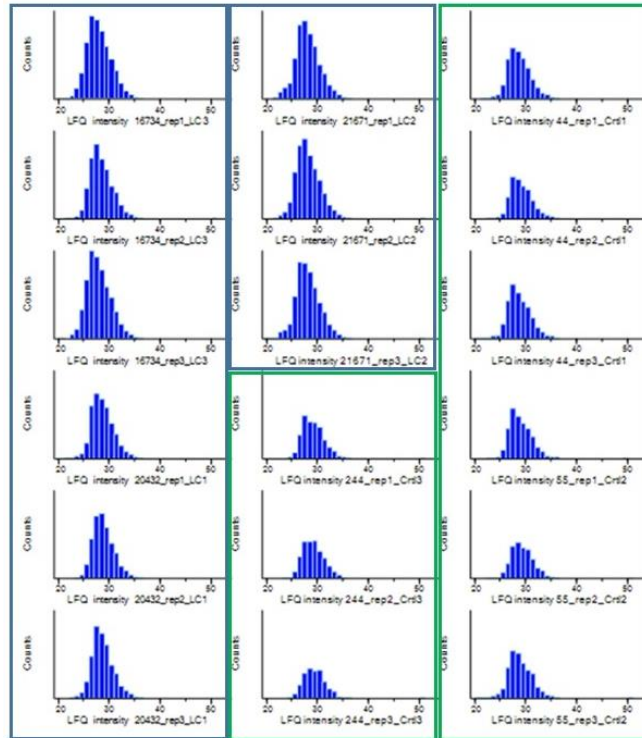
U zdravim tkivima identificirano je ukupno 3343 proteina, od kojih je 2636 identificirano u uzorku 55; 2977 proteina u uzorku 44 i 2462 proteina u uzorku 244 (Slika 64. A)). Od ukupno 3343 identificirana proteina, 2061 protein identificiran je u sva tri uzorka (61 %), 610 proteina identificirano je u dva uzorka zdravih kontrola (18,2 %), dok je 672 proteina identificirano u samo jednom uzorku zdravih kontrola (20 %) (Slika 64 B))



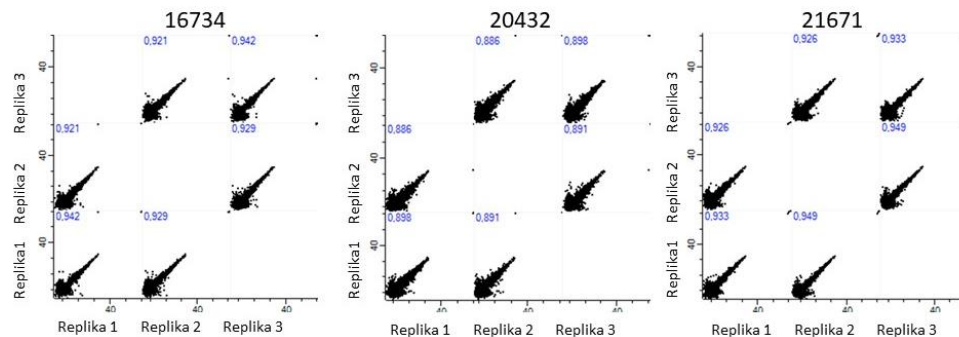
Slika 64. Prikaz ukupnog broja identificiranih proteina (A) i broja identificiranih proteina u uzorcima (B) spektrometrijom masa u zdravom tkivu.

U svrhu provjere kvalitete podataka dobivenih analizom na spektrometru masa, prvo je provjerena distribucija *LFQ* vrijednosti intenziteta za svaki uzorak u analizi. Kod podataka dobre kvalitete očekuje se normalna distribucija *LFQ* vrijednosti intenziteta u svim analiziranim. Kao što je vidljivo sa Slike 65. distribucije *LFQ* intenziteta analiziranih uzoraka približno prate očekivanu normalnu distribuciju, što ukazuje da su rezultati dovoljno dobri za daljnje analize.

Prije daljnjih analiza, zanimalo nas i je koliko su usporedive biološke replike međusobno. Stoga smo korelirali vrijednosti *LFQ* vrijednosti intenziteta u svim replikama pojedinih uzoraka. Što je vrijednost korelacije bliže vrijednosti 1, korelacija je jača. Kao što je vidljivo sa Slike 66., *LFQ* vrijednosti intenziteta bioloških replika svakog tumorskog uzorka međusobno jako koreliraju, što pokazuje da su uzorci bioloških replika tehnički dobro pripremljeni za analizu na spektrometru masa.

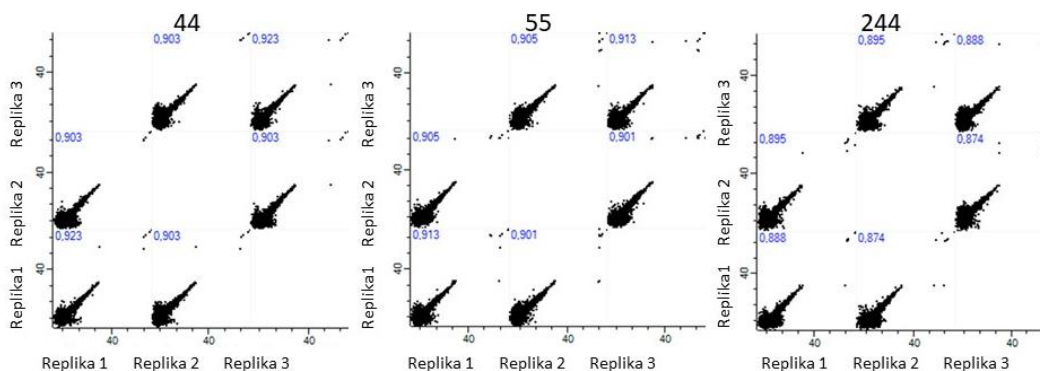


Slika 65. Distribucija vrijednosti *LFQ* intenziteta u uzorcima. Plavim kvadratićem označene su distribucije *LFQ* vrijednosti tumorskih uzoraka, a zelenim kvadratićem *LFQ* distribucije zdravih kontroli



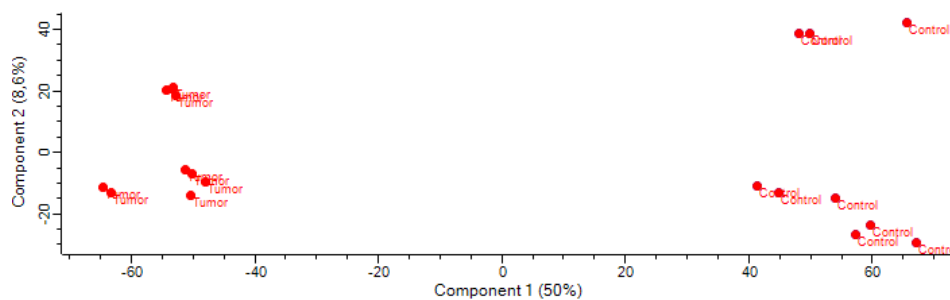
Slika 66. Korelacija *LFQ* vrijednosti intenziteta bioloških replika tumora pluća skvamoznih stanica. Plavom bojom označene su vrijednosti Pearson korelacije *LFQ* vrijednosti intenziteta.

Rezultati korelacije *LFQ* vrijednosti intenziteta bioloških replika zdravih kontroli također pokazuju da su sve biološke replike tehnički podjednako dobro obrađene i pripremljene za analizu na spektrometru masu (Slika 67.)



Slika 67. Korelacija *LFQ* vrijednosti intenziteta bioloških replika zdravih kontrola. Plavom bojom označene su vrijednosti Pearson korelacije *LFQ* vrijednosti intenziteta.

Kako bi provjerili kako se biološke replike tumorskih i zdravih uzoraka u analizi grupiraju na temelju *LFQ* vrijednosti intenziteta, napravljena je PCA analiza. Kao što je vidljivo sa Slike 68. sve biološke replike tumorskih tkiva se grupiraju zajedno na lijevoj strani grafa, kao i sve biološke replike zdravih tkiva na desnoj strani grafa. Komponenta 1 PCA analize objašnjava najveće varijacije u uzorcima (50 %), dok su varijacije u uzorcima manje objašnjene komponentom 2 (8,6 %). Razdvajanje tumorskih uzoraka od zdravih stoga je pretpostavljeno efekt bioloških varijacija između tumorskih uzoraka i zdravih kontrola, a ne efekt tehničke prirode.

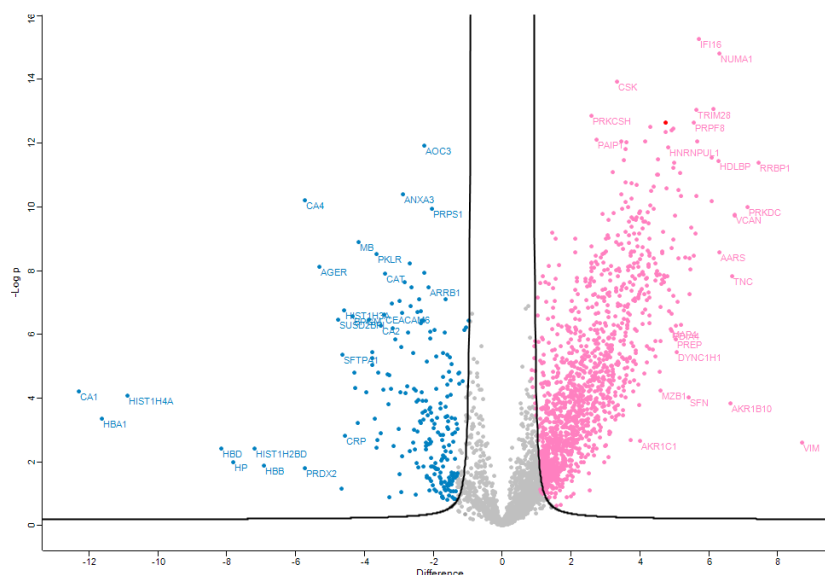


Slika 68. PCA analiza tumorskih i zdravih uzoraka na temelju *LFQ* vrijednosti intenziteta.

4.3.2. Rezultati analize diferencijalne ekspresije proteina

Kako bi ispitali profil proteinske ekspresije koji doprinosi fenotipu tumora pluća skvamoznih stanica, u odnosu na zdrave kontrole, napravljena je analiza diferencijalne ekspresije proteina. Za ovu analizu korištena je matrica kvantificiranih očitavanja na razini proteina, u obliku $\log_2$ LFQ vrijednosti intenziteta i program *Perseus*, koji diferencijalnu ekspresiju proteina identificira pomoću t-testa. Proteini su okarakterizirani kao diferencijalno eksprimirani na temelju postavljene granične vrijednosti $|\log_2 FC| \geq 1$ i prilagođene statističke značajnosti $p < 0,05$, prema metodi Benjamini-Hochberg [4].

Rezultati analize diferencijalne ekspresije proteina pokazali su da 1675 proteina u uzorcima tumora pluća skvamoznih stanica ima promijenjenu ekspresiju u odnosu na zdrave kontrole. Od 1675 proteina, 1404 proteina ima pozitivnu, a 271 protein negativno promijenjenu ekspresiju u tumorskom tkivu, u odnosu na zdrave kontrole. Na Slici 69. prikazani su rezultati diferencijalne ekspresije proteina u obliku *volcano* grafa. Kao što je vidljivo iz slike, statistički najznačajnije promijene detektirane su za proteine IFI16 (engl. *Interferon Gamma Inducible Protein 16*) i NUMA1 (engl. *Nuclear Mitotic Apparatus Protein 1*) sa pozitivnom promijenom, i proteine AOC3 (engl. *Amine Oxidase Copper Containing 3*) i ANXA3 (engl. *Annexin A3*) sa negativnom promjenom razine ekspresije u tumorskom, u odnosu na zdravo tkivo.



Slika 69. Rezultati diferencijalne ekspresije proteina prikazani *Volcano* plotom. Rozom bojom označena su imena proteina sa promijenjenom razine ekspresije, a sivom bojom proteini bez značajne promijene u razini ekspresije. Crna crta označava postavljene granične vrijednosti razlike ekspresije i p vrijednosti.

20 proteina sa najvećom pozitivnom promijenjenom razinom ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole, prikazani su u Tablici 30. Uloga navedenih proteina u stanici je raznolika. Primjerice, intermedijalni filament VIM ima ulogu u održavanju staničnog integriteta te sudjeluje u adheziji i migraciji stanica. PRKDC djeluje kao senzor oštećenja DNA, dok PARP1 ima ulogu u popravku DNA. NUMA1 ima ulogu u formaciji i održavanju polova diobenog vretena te poravnanju kromosoma tijekom mitoze. IFI16 ima ulogu u regulaciji transkripcije i može djelovati kao represor transkripcije.

Tablica 30. Prikaz 20 proteina sa najvećom pozitivnom promjenom u razini ekspresije u tumorskim uzorcima u odnosu na zdrave kontrole

Protein ID	Simbol gena	Opis proteina	log2FC	Prilagođena p vrijednost
P08670	VIM	Vimentin	8,69	<1,0 x10 ⁻¹⁰
Q9P2E9	RRBP1	Ribosome-binding protein 1	7,44	<1,0 x10 ⁻¹⁰
P78527	PRKDC	DNA-dependent protein kinase catalytic subunit	7,11	<1,0 x10 ⁻¹⁰
Q07065	CKAP4	Cytoskeleton-associated protein 4	6,75	<1,0 x10 ⁻¹⁰
P13611	VCAN	Versican core protein	6,75	<1,0 x10 ⁻¹⁰
P24821	TNC	Tenascin	6,67	<1,0 x10 ⁻¹⁰
O60218	AKR1B10	Aldo-ketoreductase family 1 member B10	6,63	<1,0 x10 ⁻¹⁰
P49588	AARS	Alanine-tRNA ligase, cytoplasmic	6,30	<1,0 x10 ⁻¹⁰
Q14980	NUMA1	Nuclear mitotic apparatus protein 1	6,29	<1,0 x10 ⁻¹⁰
Q00341	HDLBP	Vigilin	6,27	<1,0 x10 ⁻¹⁰
P09874	PARP1	Poly[ADP-ribose] polymerase 1	6,13	<1,0 x10 ⁻¹⁰
P53621	COPA	Coatomer subunit alpha	6,08	<1,0 x10 ⁻¹⁰
P49257	LMAN1	Protein ERGIC-53; Lectin, Mannose Binding 1	6,07	<1,0 x10 ⁻¹⁰
Q16666	IFI16	Gamma-interferon-inducible protein 16	5,72	<1,0 x10 ⁻¹⁰
Q92616	GCN1L1	Translational activator GCN1	5,65	<1,0 x10 ⁻¹⁰
Q13263	TRIM28	Transcription intermediary factor 1-beta	5,64	<1,0 x10 ⁻¹⁰
Q86UP2	KTN1	Kinectin	5,62	<1,0 x10 ⁻¹⁰
Q9Y4L1	HYOU1	Hypoxia up-regulated protein 1	5,61	<1,0 x10 ⁻¹⁰
Q6P2Q9	PRPF8	Pre-mRNA-processing-splicing factor 8	5,56	<1,0 x10 ⁻¹⁰
Q04637	EIF4G1	Eukaryotic translation initiation factor 4 gamma 1	5,55	<1,0 x10 ⁻¹⁰

20 proteina sa najvećom negativnom promijenjenom razinom ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole, prikazani su u Tablici 31. Proteini HBB i HBA1 povezuju se sa transportom dušikovog monoksida i kisika, a zajedno sa proteinima PRDX2, HBD i HP sudjeluju

u staničnom odgovoru na toksične supstance i u metaboličkim procesima kisikovih reaktivnih vrsta. Proteini CRP, SFTPA1 i C3 sudjeluju u procesu opsonizacije.

Tablica 31. Prikaz 20 proteina sa najvećom negativnom promjenom u razini ekspresije u tumorskim uzorcima u odnosu na zdrave kontrole

Protein ID	Simbol gena	Opis proteina	log2FC	Prilagođena p vrijednost
P00915	CA1	Carbonic anhydrase 1	-12,28	$<1,0 \times 10^{-10}$
P69905	HBA1	Hemoglobin subunit alpha	-11,62	$<1,0 \times 10^{-10}$
P62805	HIST1H4A	Histone H4	-10,87	$<1,0 \times 10^{-10}$
P02042	HBD	Hemoglobin subunit delta	-8,15	$<1,0 \times 10^{-10}$
P00738	HP	Haptoglobin	-7,81	$<1,0 \times 10^{-10}$
P58876	HIST1H2BD	Histone H2B type 1-D	-7,20	$<1,0 \times 10^{-10}$
P68871	HBB	Hemoglobin subunit beta	-6,93	$8,42 \times 10^{-5}$
P22748	CA4	Carbonic anhydrase 4	-5,73	$<1,0 \times 10^{-10}$
P32119	PRDX2	Peroxiredoxin-2	-5,72	$1,83 \times 10^{-4}$
Q15109	AGER	Advanced glycosylation end product-specific receptor	-5,31	$<1,0 \times 10^{-10}$
Q9UGT4	SUSD2	Sushi domain-containing protein 2	-4,76	$<1,0 \times 10^{-10}$
P01024	C3	Complement C3	-4,67	$1,29 \times 10^{-3}$
Q8IWL2	SFTPA1	Pulmonary surfactant-associated protein A1	-4,63	$<1,0 \times 10^{-10}$
P68431	HIST1H3A	Histone H3.1	-4,58	$<1,0 \times 10^{-10}$
P02741	CRP	C-reactive protein	-4,56	$1,63 \times 10^{-4}$
P07738	BPGM	Bisphosphoglycerate mutase	-4,34	$<1,0 \times 10^{-10}$
P24158	PRTN3	Myeloblastin	-4,28	$6,85 \times 10^{-5}$
P84243	H3F3A	Histone H3.3	-4,26	$8,33 \times 10^{-5}$
P62736	ACTA2	Actin Alpha 2	-4,18	$1,86 \times 10^{-4}$
P02144	MB	Myoglobin	-4,16	$<1,0 \times 10^{-10}$

4.3.3. Rezultati usporedbe ekspresije proteina sa ekspresijom gena i miRNA

Obzirom da je pretpostavka da se promjena u razini ekspresije gena može detektirati i na proteinskoj razini, napravljena je usporedba rezultata dobivenih sekvenciranjem transkriptoma i spektrometrije masa. U ovu analizu uključeni su geni čija je razina ekspresije potvrđena na obje testirane kohorte. Rezultati pokazuju da od ukupno 1404 proteina sa pozitivno promijenjenom razinom ekspresije u tumorskom tkivu, ista promjena na razini mRNA ekspresije zabilježena je za 245 proteina. Prekomjerno eksprimirani proteini imaju ulogu u replikaciji i popravku DNA, biosintezi DNA, kao i u metabolizmu glikogena i glutaciona te staničnoj adheziji. Primjerice, u replikaciji DNA ulogu imaju proteini MCM2-MCM7, PCNA, kao i jedan od visokoumreženih gena FEN1. Drugi visokoumrežen gen, DCPS, koji je također prekomjerno eksprimiran i na mRNA i na proteinskoj razini, ima ulogu u staničnom odgovoru na nutrijente. U diferencijaciji epitela sudjeluju prekomjerno eksprimirani proteini poput KRT19, FASN i HDAC1. 20 proteina sa

najvećom razinom ekspresije na proteinskoj razini, čija je razina povišena i na mRNA, prikazani su u Tablici 32.

Tablica 32. 20 proteina sa najvišom promjenom ekspresije, potvrđenih i na transkripcijskoj razini

Simbol gena	log2FC_protein	padj_protein	log2FC_mRNA	padj_mRNA
PRKDC	7,11	$<1,0 \times 10^{-10}$	1,29	0,0004
CKAP4	6,76	$<1,0 \times 10^{-10}$	1,13	0,0164
TNC	6,68	$<1,0 \times 10^{-10}$	2,73	0,0083
AKR1B10	6,63	$<1,0 \times 10^{-10}$	7,41	0,0010
TRIM28	5,64	$<1,0 \times 10^{-10}$	1,29	0,0012
SFN	5,41	$<1,0 \times 10^{-10}$	3,90	$7,43 \times 10^{-7}$
SERPINB5	4,97	$<1,0 \times 10^{-10}$	7,19	$9,27 \times 10^{-11}$
CAD	4,97	$<1,0 \times 10^{-10}$	2,52	$2,21 \times 10^{-9}$
PDIA4	4,93	$<1,0 \times 10^{-10}$	1,13	0,0153
EFTUD2	4,69	$<1,0 \times 10^{-10}$	1,12	0,0154
FKBP10	4,66	$<1,0 \times 10^{-10}$	1,22	0,0571
MZB1	4,60	$<1,0 \times 10^{-10}$	2,84	0,0061
THBS2	4,48	$<1,0 \times 10^{-10}$	3,24	$2,4 \times 10^{-6}$
ALDH18A1	4,44	$<1,0 \times 10^{-10}$	1,18	$5,66 \times 10^{-8}$
OPA1	4,39	$<1,0 \times 10^{-10}$	1,00	0,0099
VAR5	4,37	$<1,0 \times 10^{-10}$	1,47	$2,07 \times 10^{-7}$
CKMT1A	4,34	$<1,0 \times 10^{-10}$	2,86	0,0568
ERLIN2	4,30	$<1,0 \times 10^{-10}$	1,52	0,1230
MOGS	4,28	$<1,0 \times 10^{-10}$	1,04	0,0390
FBLN2	4,28	$<1,0 \times 10^{-10}$	1,16	0,0872

Od 1085 gena za koje je zabilježena smanjena ekspresija u tumorskom tkivu, na proteinskoj razini zabilježena je smanjena ekspresija za 271 protein. Proteini čija je ekspresija smanjena u tumorskom tkivu na mRNA i proteinskoj razini prikazani su u Tablici 33. Proteini sa smanjenom ekspresijom u tumorskom tkivu imaju ulogu u regulaciji imunološkog odgovora, aktivaciji leukocita, sustava komplemenata, neutrofilnoj degranulaciji, kao i u pozitivnoj regulaciji apoptoze i metabolizmu glutationa. Primjerice, DEFA1 i DEFA3 sudjeluju u odgovoru stanice na stimulans, dok ELANE predstavlja važan faktor neutrofilima posredovane citotoksičnosti.

Tablica 33. 20 proteina sa najviše smanjenom ekspresijom, potvrđenih i na transkripcijskoj razini

Simbol gena	log2FC_protein	padj_protein	log2FC_mRNA	padj_mRNA
MPO	-2,9819	0,0003	-6,1132	0,0002
AZU1	-3,0130	0,0009	-6,4441	0,0357
CLIC5	-3,1763	0,0003	-4,5017	4,75 x10 ⁻⁵
CAV2	-3,2356	0,0003	-1,0988	0,1514
SNCA	-3,2905	0,0003	-1,7406	0,2612
MFAP4	-3,3133	0,0003	-2,7040	0,0011
CAT	-3,4070	0,0001	-2,1344	2,23x10 ⁻¹³
CEACAM6	-3,4232	0,0001	-3,0279	0,0301
ALDOB	-3,4427	0,0003	-5,9859	0,0187
CA2	-3,5292	0,0001	-1,0668	0,2572
ITGA2B	-3,6371	0,0004	-3,9972	0,0077
PPBP	-3,7037	0,0003	-5,6665	0,0017
EPB42	-3,7666	0,0001	-6,8759	0,0515
VWF	-3,7851	0,0001	-1,9185	0,0008
BPI	-3,8644	6,52 x10 ⁻⁵	-7,7199	0,0094
PRTN3	-4,2843	6,86 x10 ⁻⁵	-7,8196	0,0746
SUSD2	-4,7687	<1,0 x10 ⁻¹⁰	-3,1789	0,0002
AGER	-5,3113	<1,0 x10 ⁻¹⁰	-5,8860	7,27 x10 ⁻⁷
CA4	-5,7334	<1,0 x10 ⁻¹⁰	-6,0137	0,0141
CA1	-12,2844	<1,0 x10 ⁻¹⁰	-7,2755	5,57 x10 ⁻⁵

Također je provjerena korelacija razine ekspresije proteina sa genima čija ekspresija utječe na sveukupno preživljenje pacijenata. Od 89 identificiranih prognostičkih gena, na proteinskoj razini potvrđena je prekomjerna ekspresija 9 proteina (GALNT2, PYGB, LAMB3, MCCC1, DCPS, DSG3, PC, FARSA i DTYMK), kao i smanjena ekspresija proteina SLC9A3RN.

Nadalje, zanimalo nas je koji proteini sa sniženom ekspresijom imaju povišenu ekspresiju na transkripcijskoj razini. Pretpostavka je naime da bi razlog smanjene proteinske ekspresije tih gena mogao ležati u povišenoj ekspresiji miRNA, koja bi mogla utišati njihovu translaciju. Za identifikaciju miRNA koje ciljaju proteine sa povišenom ekspresijom korištene su samo validirane interakcije miRNA-mRNA/protein iz baza podataka unutar multiMir paketa u R programu. Naši rezultati pokazuju smanjenu ekspresiju ukupno 7 proteina čija je ekspresija povišena na transkripcijskoj razini: SERPIND1, CMA1, APEH, HMBS, HIST2H3A, RPIA i HIST1H4A. Pretragom miRNA pronađeno je da od navedenih gena, gen HMBS je ciljani gen hsa-miR-181a-5p. Međutim, obzirom da navedena miRNA ima sniženu ekspresiju u tumorskom tkivu, pretpostavka je da snižena ekspresija navedenog proteina nije posljedica djelovanja utišane translacije proteina od strane miRNA. Nadalje, zanimalo nas je koji proteini su prekomjerno ekspimirani i na mRNA razini. Pretpostavka je da bi razlog njihove ekspresije mogao ležati u smanjenoj ekspresiji miRNA.

Od 245 proteina čija je ekspresija povišena i na mRNA razini, pronašli smo da su 57 gena ciljni geni miRNA sa sniženom ekspresijom. Popis miRNA sa sniženom ekspresijom i prekomjernom ekspresijom proteina navedeni su u Tablici 34.

Tablica 34. Prekomjerno eksprimirani proteini i miRNA sa sniženom ekspresijom koje potencijalno reguliraju translaciju proteina

Simbol proteina	Smanjeno eksprimirane miRNA
TNC	hsa-miR-30d-5p, hsa-miR-145-5p, hsa-miR-2110, hsa-miR-676-3p
AKR1B10	hsa-miR-145-5p, hsa-miR-30d-5p, hsa-miR-133a-3p
SFN	hsa-miR-107, hsa-miR-486-3p
SERPINB5	hsa-miR-30a-5p, hsa-miR-107, hsa-miR-130a-3p
CAD	hsa-miR-338-3p, hsa-miR-92a-3p, hsa-miR-107, hsa-miR-133a-3p, hsa-miR-363-3p, hsa-miR-92b-3p, hsa-miR-486-3p
THBS2	hsa-miR-106a-5p, hsa-miR-133a-3p, hsa-miR-126-5p
MXRA5	hsa-miR-34b-5p, hsa-miR-34c-5p
PCNA	hsa-miR-125b-2-3p, hsa-miR-30a-5p, hsa-let-7c-5p, hsa-miR-92a-3p, hsa-miR-107, hsa-miR-30d-5p, hsa-miR-363-3p
UCHL1	hsa-miR-181a-5p, hsa-miR-92a-3p, hsa-miR-34b-5p, hsa-miR-34c-5p, hsa-miR-338-3p, hsa-miR-664a-5p
TTLL12	hsa-miR-30a-5p, hsa-miR-107, hsa-miR-130a-3p, hsa-miR-106a-5p, hsa-miR-30d-5p, hsa-miR-3605-5p
AKR1C2	hsa-miR-107, hsa-miR-130a-3p
IGF2BP3	hsa-miR-320a-3p, hsa-miR-486-3p, hsa-miR-30a-5p, hsa-miR-181a-5p, hsa-miR-664a-5p, hsa-let-7c-5p
ADH7	hsa-miR-34b-5p, hsa-miR-34c-5p
PTGFRN	hsa-miR-30a-5p, hsa-miR-106a-5p, hsa-miR-30d-5p, hsa-miR-92a-3p, hsa-miR-107, hsa-miR-139-5p, hsa-miR-181a-5p, hsa-miR-2110, hsa-miR-3065-3p, hsa-miR-676-3p
TXNDC5	hsa-miR-30a-5p, hsa-miR-30d-5p, hsa-miR-2110
GPC1	hsa-miR-107
PYCR1	hsa-miR-2110
SLC9A3R1	hsa-let-7c-5p, hsa-miR-181a-5p, hsa-miR-34b-5p
PFKP	hsa-miR-338-3p, hsa-miR-3613-3p, hsa-miR-106a-5p
FAP	hsa-miR-30a-5p, hsa-miR-30d-5p
MFAP2	hsa-miR-126-3p
DSP	hsa-miR-30a-5p, hsa-miR-30d-5p, hsa-let-7c-5p, hsa-miR-107, hsa-miR-145-5p, hsa-miR-486-5p
ABCC1	hsa-miR-145-5p, hsa-miR-664a-5p
MMP12	hsa-miR-145-5p
DPP3	hsa-miR-107, hsa-let-7c-5p
FEN1	hsa-miR-107, hsa-miR-126-3p
KRT19	hsa-miR-130a-3p, hsa-miR-133a-3p, hsa-miR-145-5p, hsa-miR-126-3p
MCM2	hsa-miR-107, hsa-miR-130a-3p, hsa-miR-145-5p
MMP1	hsa-miR-145-5p, hsa-miR-30d-5p, hsa-miR-34b-5p, hsa-miR-34c-5p
CYP4F11	hsa-miR-107, hsa-miR-34c-5p
MCM4	hsa-miR-338-3p, hsa-let-7c-5p, hsa-miR-30a-3p, hsa-miR-30a-5p, hsa-miR-107, hsa-miR-1251-5p, hsa-miR-3613-3p
TUBB3	hsa-miR-30a-5p, hsa-miR-92a-3p
ITGB4	hsa-miR-30a-5p, hsa-miR-126-3p
CRABP2	hsa-miR-107, hsa-miR-34b-5p, hsa-miR-34c-5p, hsa-miR-486-3p
HMGB3	hsa-miR-106a-3p, hsa-miR-125a-5p, hsa-miR-181a-5p
CTHRC1	hsa-miR-30a-5p, hsa-miR-30d-5p, hsa-miR-107, hsa-miR-126-3p
SLC7A5	hsa-miR-3065-3p, hsa-miR-30a-5p, hsa-miR-30d-5p, hsa-miR-126-3p, hsa-miR-133a-3p, hsa-miR-126-5p, hsa-miR-338-3p, hsa-miR-652-3p, hsa-miR-139-3p, hsa-miR-125b-2-3p, hsa-miR-1255a
IL4I1	hsa-miR-30a-5p
IPO4	hsa-miR-107, hsa-miR-2110, hsa-miR-92a-3p
COL1A1	hsa-miR-133a-3p, hsa-miR-107, hsa-miR-181a-5p, hsa-miR-130a-3p, hsa-miR-145-5p, hsa-miR-34c-5p, hsa-miR-3605-5p, hsa-let-7c-5p
THY1	hsa-miR-125a-5p, hsa-miR-139-3p, hsa-miR-486-3p

Tablica 34. Nastavak

DTYMK	hsa-miR-107
FERMT1	hsa-miR-181a-5p, hsa-miR-125a-5p
COL5A1	hsa-miR-92a-3p, hsa-miR-92b-3p, hsa-miR-181a-5p, hsa-miR-145-5p, hsa-miR-30a-5p, hsa-miR-107
IGFBP2	hsa-miR-126-3p, hsa-miR-107
PDK1	hsa-miR-30d-5p
KRT6B	hsa-miR-30a-5p
MCM7	hsa-miR-107, hsa-let-7c-5p, hsa-miR-92a-3p, hsa-miR-130a-3p, hsa-miR-125b-2-3p
FSCN1	hsa-miR-133a-3p, hsa-miR-133b, hsa-miR-30a-5p, hsa-miR-107, hsa-miR-326, hsa-miR-145-5p
SULF1	hsa-miR-130a-3p
RAP2B	hsa-miR-30a-5p, hsa-miR-30d-5p, hsa-miR-107, hsa-miR-125a-5p, hsa-miR-126-3p, hsa-miR-490-3p
FAM83H	hsa-miR-139-3p
GPNUMB	hsa-miR-107, hsa-miR-130a-3p, hsa-miR-126-3p, hsa-miR-34b-5p, hsa-miR-34c-5p
MARCKSL1	hsa-miR-2110, hsa-let-7c-5p, hsa-miR-30a-5p, hsa-miR-30d-5p, hsa-miR-34c-5p, hsa-miR-107, hsa-miR-181a-5p, hsa-miR-34b-5p, hsa-miR-125b-2-3p, hsa-miR-3617-5p
BDH1	hsa-miR-30a-5p, hsa-miR-125a-5p
GCLC	hsa-miR-92a-3p, hsa-miR-92b-3p, hsa-miR-30a-5p, hsa-miR-30d-5p
SLC25A10	hsa-miR-107

Također, provjeren je potencijalni utjecaj miRNA na utišavanje translacije prognostičkih gena sa povišenom ekspresijom. Od 56 gena čija se razina ekspresije povezuje sa prognozom bolesti, a čija je ekspresija na mRNA razini povišena, nismo pronašli prognostičke proteine sa negativnom promjenom u razini ekspresije. Također, nismo pronašli prognostičke gene čija je ekspresija smanjena na mRNA razini, a povećana na proteinskoj razini.

Provjerili smo i korelaciju smjera ekspresije za visokoumrežene gene u PPI mreži, kao i visokoumrežene gene iz *sienna3* klastera. Zanimljivo je da smo na proteinskoj razini zabilježili povišenu ekspresiju samo za visokoumrežene gene MCM2 i FEN1, dok je razina ekspresije CDK1 ispod postavljene granične vrijednosti ($\log_2FC=0,63$). Obzirom da je njihova ekspresija povišena na mRNA razini, kao i na proteinskoj, provjerili smo koje miRNA su smanjeno eksprimirane a ciljaju navedene gene/proteine. Za MCM2 pronašli smo hsa-miR-130a-3p, hsa-miR-145-5p i hsa-miR-107. Za FEN1 pronašli smo hsa-miR-126-3p i hsa-miR-107. Za CDK1 pronašli smo hsa-miR-34c-5p, hsa-miR-652-3p, hsa-miR-3617-5p, hsa-miR-107, hsa-miR-92a-3p, hsa-miR-30d-5p, hsa-miR-130a-3p, hsa-miR-126-5p i hsa-miR-126-3p. Također smo provjerili korelaciju smjera ekspresije za gene identificirane kao visokoumreženi unutar *sienna3* klastera. Na proteinskoj razini detekirana su 6 gena, od kojih jedino PRF1 i HLA-DPB1 pokazuju isti smjer ekspresije na obje razine. Geni GIMAP7, GIMAP8 i HLA-E na proteinskoj razini ne pokazuju značajnu promjenu u odnosu na zdravo tkivo, dok je GIMAP4 eksprimiran u suprotnom smjeru od mRNA ekspresije.

5. RASPRAVA

1 tablica

5. RASPRAVA

Tumor pluća skvamoznih stanica je, uz adenokarcinom pluća, najčešće dijagnosticirani histološki tip tumora pluća. Tumor pluća skvamoznih stanica često nastaje u centralnom djelu pluća, obično u području bronha. Trajna ozljeda epitela dišnih putova može dovesti do fenotipa skvamozne metaplazije, a potom i do displazije, a obje su histološke značajke koje se mogu vidjeti u dišnim putovima pušača. Vjeruje se da se tumor pluća skvamoznih stanica razvija kao posljedica niza genetskih i epigenetskih promjena koje mijenjaju epitel od skvamozne metaplazije, zatim do displazije, karcinoma *in situ* i konačno do invazivnog karcinoma [114]. Ranije je pokazano da su temeljni mehanizmi, uključujući razvoj, rast i širenje tumora, prilično različiti između dva najčešća tipa tumora pluća [11]. Primjena visokoprotočnih metoda, poput sekvenciranja slijedeće generacije, doprinjelo je identifikaciji velikog broja pokretačkih mutacija u adenokarcinomu pluća koje su uspješno iskorištene za razvoj lijekova. Ciljane terapije koje inhibiraju prekomjerno aktivirane stanične puteve, kao posljedicu djelovanja pokretačkih mutacija, uspješno su poboljšale stopu preživljenja pacijenata oboljelih od adenokarcinoma pluća. U tumoru pluća skvamoznih stanica do sada je identificiran mali broj pokretačkih mutacija, međutim pokazalo se da one nisu pogodne za razvoj lijekova, što razvoj novih paradigmi liječenja za ovu bolest čini izazovnim. Stoga je potrebna daljnja karakterizacija bioloških osobina ove vrste tumora, kako i na genomskoj, tako i na transkripcijskoj i translacijskoj razini, u svrhu identifikacije novih potencijalnih biljega koji doprinose tumorigenenom fenotipu ove bolesti.

Glavni cilj ove doktorske disertacije je doprinjeti boljem razumijevanju bioloških karakteristika tumora pluća skvamoznih stanica te pronaći potencijalne ekspresijske biomarkere koji bi se mogli koristiti u dijagnostici i prognostici ove bolesti. Fokus disertacije je analiza ekspresijskog profila tumora pluća skvamoznih stanica na transkripcijskoj, post-transkripcijskoj i translacijskoj razini. U sklopu ove disertacije prikupljena su tkiva fiksirana formalinom i uklopljena u parafin (engl. *Formalin-fixed, paraffin embedded*, FFPE) od pacijenata oboljelih od tumora pluća skvamoznih stanica, kao i tkiva pluća zdravih osoba. U istraživanje je uključeno 23 pacijenata kojima prije resekcije pluća nije propisana terapija, kao i 4 zdrave osobe. FFPE tkiva narezana su u 10 µm rezove koji su se koristili za izolaciju ukupne RNA i proteina. Izolirana ukupna RNA korištena je pripremu knjižnica za profiliranje transkriptoma i miRNA sekvenciranjem slijedeće generacije, dok su izolirani tkivni proteini korišteni za analizu na spektrometru masa. Rezultati ekspresije gena potvrđeni su na validacijskoj korti pacijenata

oboljelih od tumora pluća skvamoznih stanica, uključenih u TCGA LUSC projekt, a kao zdrave kontrole uključene su osobe iz GTEx projekta. Obzirom na veći broj korištenih analiza i dobivenih rezultata, rasprava je podijeljena na zasebna poglavlja, kako bi se što preglednije naglasile nove spoznaje.

5.1. Transkriptomski profil tumora pluća skvamoznih stanica

U svrhu ispitivanja profila ekspresije gena na transkriptomskoj razini, analizirani su podaci dobiveni sekvenciranjem slijedeće generacije pomoću nekoliko bioinformatičkih alata, za obje testirane kohorte. Promjena u razini ekspresije gena određena je pomoću DESeq2 alata, a potom su biološke funkcije detektiranih gena sa promijenjenom razinom ekspresije okarakterizirane pomoću analize genske ontologije i analize obogaćivanja setovima gena. Primjena ovih metoda omogućuje pregled dobivenih rezultata u kontekstu bioloških funkcija gena te omogućava identifikaciju staničnih puteva i procesa u kojima svi geni sa promijenjenom razinom ekspresije sudjeluju.

Rezultati ovog istraživanja pokazuju da je u tumorskom tkivu značajno povišena razina ekspresije gena koji sudjeluju u procesima keratinizacije i formacije keratinizirane ovojnice. Najzastupljeniji geni koji doprinose navedenim staničnim procesima su keratini. Na obje testirane kohorte potvrđena je prekomjerna ekspresija ukupno trinaest keratina (KRT5, KRT6A-C, KRT13-KRT17, KRT19, KRT74 i KRT81). Poznato je da keratinski filamenti sudjeluju u formiranju proteinskog strukturnog okvira unutar stanične citoplazme i, kao takvi, imaju važnu ulogu u zaštiti svih tipova epitelnih stanica od (ne)mehaničkih stresora. Nadalje, sudjeluju u diferencijaciji stanice, regulaciji stanične proliferacije, razvoju epiderme i razvoju ektoderma. Ekspresija keratina se obično mijenja u stresnim uvjetima, uz post-translacijske modifikacije (poput fosforilacije) i strukturalnu reorganizaciju. [115]. Primjerice, ekspresija KRT 5 i p63 na matičnim stanicama dišnog sustava povezana je sa regeneracijom plućnog tkiva nakon oštećenja uzrokovanog infekcijom virusa H1N1. p63/KRT5 matične stanice diferenciraju u pneumocite tipa I i II nakon transplantacije u oštećena pluća podupirući regeneraciju tkiva pluća. Regeneracija plućnog tkiva onemogućena je uklanjanjem p63/KRT5 matičnih stanica u *in vivo* modelu što dovodi do pre-fibrotičnih lezija i smanjenja opskrbe kisikom [116]. Navedeno upućuje na potencijal korištenja ovih vrsta stanica u liječenju akutnih i kroničnih oboljenja dišnog sustava. Također je poznato da karcinome skvamoznih stanica, neovisno o mjestu nastanka, karakterizira povišena ekspresija markera epitelne stratifikacije, keratina KRT5, KRT14 i KRT17, i keratina KRT6 i KRT16 koji se

povezuju sa hiperproliferacijom keratinocita. Primjerice, na *in vitro* modelu adenokarcinoma i metastatskog tumora pluća pokazano je da prekomjerna ekspresija KRT6A promovira staničnu proliferaciju povećanjem ekspresije G6PD gena aktivacijom MYC signalnog puta. G6PD gen je enzim koji sudjeluje u metabolizmu glukoze (ciklusu pentoza fosfata) i koji ima ulogu u invaziji tumora [117]. Naši rezultati također ukazuju na povišenu ekspresiju navedenih keratina, kao i prekomjernu aktivaciju Myc signalnog puta. Iako je c-Myc prekomjerno eksprimiran samo u TCGA kohorti, obje kohorte pokazuju prekomjernu ekspresiju MYCL gena, sugerirajući da bi KRT6A/MYCL/G6P os mogla utjecati na proliferaciju tumora pluća skvamoznih stanica. Nadalje, meta analiza prekomjerno eksprimiranih gena u ranim stadijima tumora pluća skvamoznih stanica i adenokarcinoma, pokazala je da su prekomjerno eksprimirani keratini karakteristika tumora pluća skvamoznih stanica, dok je njihova ekspresija smanjena u adenokarcinomu pluća [118].

Geni sa povišenom razinom ekspresije u tumoru pluća skvamoznih stanica, osim u procesu keratinizacije epitela, sudjeluju i u regulaciji staničnog ciklusa. Rezultati ukazuju da su u obje testirane kohorte prekomjerno aktivirani stanični putevi povezani sa diobom stanice, poput FOXM1, p53, p63, PLK1, AURORA i AURORA B staničnih puteva. Pronađena je i prekomjerna aktivacija E2F obitelji transkripcijskih faktora, kao i aktivacija ATR kompleksa u odgovoru na replikacijski stres. Poznato je da je kontrolirana replikacija DNA esencijalna za očuvanje stabilnosti genoma. Narušena kontrola replikacije DNA, uzrokovana endogenim ili egzogenim stresom (uzrokovanim primjerice pušenjem) dovodi do kolapsa replikacijske rašlje i uzrokuje oštećenja DNA, što često dovodi do razvoja tumora. FOXM1 je transkripcijski faktor koji ima važnu ulogu u normalnom perinatalnom razvoju pluća gdje utječe na diferencijaciju plućnih epitelnih stanica, kao i njihovu funkciju. Također, FOXM1 je važan za regulaciju staničnog ciklusa jer kontrolira tranziciju iz G1 u S fazu, G2 u M fazu staničnog ciklusa, kao i progresiju M faze. Nekoliko gena koji su ključni za regulaciju mitoze, poput Skp2, ciklina A i B, Cdc25A, Cdc25C, AURKB, BIRC5 i CENPA, su pod transkripcijskom kontrolom FOXM1 [119]. Svi navedeni geni pokazali su prekomjernu ekspresiju u našim ispitivanim kohortama, a njihova povišena ekspresija pokazana je i u drugim studijama tumora pluća skvamoznih stanica [120], [121]. Jedan od gena koji je također pod transkripcijskom regulacijom FOXM1, je i PLK1 gen. PLK1 također ima nekoliko uloga u regulaciji staničnog ciklusa, poput kontrole prijelaska iz G2 u M fazu, koordinacije centrosoma i regulacije odvajanja kromosoma. U nekoliko vrsta tumora, uključujući tumor pluća skvamoznih stanica, zabilježena je prekomjerna ekspresija ovoga gena [122]. U skladu

sa rezultatima ostalih istraživanja, naši rezultati u obje ispitivane kohorte također pokazuju da je ekspresija ovoga gena znatno povišena u tumorskom tkivu. U nedavno objavljenoj studiji, koja je istraživala razlike u aktiviranim staničnim putevima između adenokarcinoma i tumora pluća skvamoznih stanica, pronađena je prekomjerna aktivacija NOTCH1, Wnt, Hedgehog i ErrB signalnih puteva u tumoru pluća skvamoznih stanica [80]. Iako naši rezultati ne ukazuju na povišenu aktivaciju NOTCH1 i ErrB signalnih puteva u ispitivanim kohortama, pronađena je povišena aktivacija Wnt i Hedgehog signalnih puteva. Geni povezani sa popravkom DNA također imaju važnu ulogu u nadzoru i zaštiti integriteta genoma, a njihova deregulacija jedan je od glavnih obilježja agresivnih tumora. Postoji nekoliko mehanizama za popravak DNA u stanicama, a u našim kohortama pronašli smo prekomjerno aktivirane četiri skupine gena koji održavaju integritet genoma. Primjerice, prekomjerno su eksprimirani geni koji sudjeluju u popravku DNA izrezivanjem baza, neusklađenom popravku (engl. *mismatch repair*), homolognoj rekombinaciji i *Fanconi anaemia* signalnom putu. Još jedna skupina gena uključena u kontrolu staničnog ciklusa, za koje je na ispitivanim kohortama zabilježena povišena ekspresija, su E2F transkripcijski faktori. Oni reguliraju ekspresiju gena odgovornih za tranziciju iz G1 u S fazu staničnog ciklusa. Prijašnje studije također su pronašle povećanu ekspresiju ovih gena u tumoru pluća [80]. E2F transkripcijski faktori su skupina gena koje na temelju funkcije možemo podijeliti u dvije skupine: aktivatori transkripcije (E2F1, E2F2 i E2F3a) i represore transkripcije (E2F3b i E2F4-8). Na temelju rezultata nekoliko studija, E2F1, E2F2, E2F3 i E2F8 mogu se smatrati onkogenima koji doprinose razvoju tumora pluća, uključujući skvamozni podtip [123], [124], [125]. Primjerice, prekomjerna ekspresija E2F1 dovodi do pojačane agresivnosti tumora uslijed visoke stope proliferacije te rezistencije na kemoterapiju [123]. Također, u studiji koja je istraživala razliku ekspresijskog profila gena između adenokarcinoma i tumora pluća skvamoznih stanica, pokazano je da E2F8 prekomjerno eksprimiran u oba tipa tumora pluća, E2F2 u adenokarcinomu, a E2F1 u tumoru pluća skvamoznih stanica [80]. Prekomjerna ekspresija E2F3 obilježje je svih podtipova tumora pluća ne-malih stanica, dok ekspresija E2F4 i E2F6 nije zabilježena na tumoru pluća [126]. Naši rezultati potvrđuju rezultate navedenih studija i pokazuju da su E2F1, E2F7 i E2F8 prekomjerno eksprimirani u tumoru pluća skvamoznih stanica. E2F2 čija se prekomjerna ekspresija ranije povezivala samo sa adenokarcinomom, također je prekomjerno eksprimiran i u našim testiranim kohortama. Nedavna studija na tumoru pluća skvamoznih stanica pokazala je da je E2F signalni put prekomjerno aktiviran nakon terapije pacijenata oboljelih od tumora pluća skvamoznih stanica

zračenjem, te da visoka ekspresija uzrokuje slabiji odgovor na terapiju zračenjem [127]. Nadalje, jedna studija pokazala je na *in vitro* modelu da je FOXM1 ciljni gen E2F1 transkripcijskog faktora, te da je proliferacija tumora pluća posredovana lncRNA-HIT-E2F1-FOXM1 signalnim putem [128]. Iako nismo pronašli povećanu ekspresiju lncRNA i HIT gena, zabilježena je povećana ekspresija E2F1 i FOXM1, sugerirajući da je i u tumoru pluća skvamoznih stanica E2F1-FOXM1 os bitna za proliferaciju tumorskih stanica, ali se točni E2F1 uzvodni mehanizam regulacije u ovoj vrsti tumora tek treba utvrditi budućim studijama.

Tumorski mikrookoliš sve više postaje zanimljiva niša za pronalazak novih prediktivnih i prognostičkih biomarkera, kao i novih mehanizama za razvoj terapija tumora. U tumorskom mikrookolišu, osim tumorskih stanica, djeluju i stanice urođenog i adaptivnog imunološkog odgovora. Stanice urođenog imunološkog odgovora su bazofili, dendritične stanice, eozinofili, monociti, makrofazi, neutrofili i NK stanice. U normalnim fiziološkim uvjetima, stanice urođenog imunološkog sustava stalno su u stanju pripravnosti, za prepoznavanje patogena koji bi potencijalno mogli narušiti homeostazu plućnog tkiva. Epitelne stanice dišnog sustava djeluju kao prva linija obrane protiv antigena koji su izbjegli mukocilijarno čišćenje. Ekspimiraju receptore za prepoznavanje molekularnih uzoraka pridruženih patogenima, poput *Toll-like*, *NOD-like* i *RIG-like* receptora. Najbolje istraženi receptori urođene imunosti su Toll-like receptori, a osim pitelnih stanica, ekspimiraju ih i dendritične stanice, makrofazi i fibroblasti. Aktivacija *Toll-like* receptora pokreće signalnu kaskadu koja dovodi do aktivacije *NF-kB* transkripcijskog faktora i oslobađanja pro-upalnih citokina i kemokina, uključujući interferon, interleukin 6 i interleukin 8 [129]. Rezultati našeg istraživanja pokazuju da je ekspresija *Toll-like* receptora TLR2 i TLR4 smanjena u tumoru pluća, dok je razina ekspresije TLR10 povišena. Nadalje, smanjena je ekspresija NF-kB, a povišena ekspresija njegovog inhibitora NFKBIB, ukazujući na smanjenu aktivaciju NF-kB signalnog puta u tumoru pluća skvamoznih stanica. U tumoru pluća, uloga urođenog imunološkog sustava je kontroverzna zbog dokaza o pro- i protu-tumorskim učincima. Primjerice, TLR4 je primarni receptor koji prepoznaje lipopolisaharid bakterija, ali može imati ulogu u tumorigenezi. Na mišjem modelu adenokarcinoma, pokazano je da su miševi s nedostatkom TLR4 osjetljiviji na upalu pluća i promicanje tumora pluća, u usporedbi s miševima divljeg tipa [129]. Medijatori upale koje luče stanice urođenog imunološkog sustava bitna su komponenta imunološkog odgovora tijekom ozljede ili infekcije. Primjerice, interleukin 8 je kemokin koji je odgovoran za privlačenje neutrofila. Neutrofili su jedan od prvih cirkulirajućih tipova imunoloških stanica koji se regrutiraju

u pluća kao odgovor na upalu. Glavni mehanizmi urođene obrambene funkcije neutrofila su fagocitoza, degranulacija i oslobađanje neutrofilnih ekstracelularnih zamki [130]. Osim IL-8, regrutiranje neutrofila može biti potaknuto lučenjem GM-CSF i upalnog proteina 2 od makrofaga. Njihov izlazak iz koštane srži i zadržavanje u plućima je dinamičan proces koji zahtijeva zajedničku ekspresiju kemokinskih receptora CXCR2 i CXCR4. Neki od najzastupljenijih proteolitičkih enzima sadržanih u granulama neutrofila su neutrofilna elastaza (ELANE) i katepsin G (CTSG). Nakon degranulacije, ovi enzimi stimuliraju proizvodnju sluzi i cijepaju propeptide citokina kako bi pojačali imunološki odgovor [130]. Rezultati naše studije, pokazuju značajno sniženu degranulaciju neutrofila u tumorskom tkivu. Razlog tome vjerojatno leži u sniženoj ekspresiji CXCR2, ELANE i CTSG gena koju smo detektirali u našim kohortama. Sada je poznato da imunološki sustav ima potencijal uništiti stanice tumora i inhibirati rast tumora putem urođenog i adaptivnog imunološkog odgovora. Adaptivni imunološki odgovori su specifični za antigen, razvijaju se sporije, nude imunološku memoriju i obuhvaćaju humoralnu i staničnu imunost posredovanu B i T stanicama. [131].

5.2. Klasteri gena i visokoumreženi geni u tumoru pluća skvamoznih stanica

U sklopu ove disertacije, napravljena je i analiza mreže proteinskih interakcija za gene sa promijenjenom razinom ekspresije. Mrežni model, napravljen u STRING alatu je vizualiziran pomoću Cytoscape programa. Vizualizirana mreža proteinskih interakcija sadržavala je 1298 čvorišta sa 13 575 interakcija između gena. Unutar mreže analizirane su visokoumrežene regije gena, kao i visokoumreženi geni, sa ciljem identifikacije gena koji potencijalno imaju važnu ulogu u tumorigenezi pluća. Naši rezultati pokazuju postojanje 11 visokoumreženih regija gena, odnosno klastera, u tumoru pluća skvamoznih stanica. U najznačajnijem klasteru, Klasteru 1, grupiralo se 106 gena, sa više od 4,5 tisuća međuinterakcija. Rezultati GSEA analize pokazali su da ti geni uglavnom imaju ulogu u regulaciji staničnog ciklusa. Naime, 75 posto obogaćenih staničnih puteva u Klasteru 1 bilo je povezano sa regulacijom različitih faza staničnog ciklusa. Nedavno objavljena studija također je identificirala visokoumrežene regije u tumoru pluća skvamoznih stanica pomoću analize mreže proteinskih interakcija i na neki način je potvrdila naše rezultate za najznačajniji klaster. Naime, geni grupirani u Klaster 1 važni su za regulaciju staničnog ciklusa, kao i kod najznačajnijeg klastera u navedenoj studiji [128]. U drugi najznačajniji klaster, Klaster 2, grupiralo se 49 gena, sa 402 interakcije. Geni u Klasteru 2, većinski imaju ulogu u regulaciji imunološkog

odgovora. Ovi rezultati nadopunjuju ranije diskutirane rezultate genske ontologije i staničnih puteva. Naime, oni također pokazuju da su geni povezani sa regulacijom staničnog ciklusa prekomjerno eksprimirani, prvenstveno u Klasteru 1, dok su geni koji imaju ulogu u imunološkom odgovoru smanjeno eksprimirani, prvenstveno u Klasteru 2. Nadalje, od 11 identificiranih klastera, za 10 klastera pronašli smo izvorišne (engl. *seed*) gene oko kojih su geni u klasteru grupirani. Pronađeni izvorišni geni za klastere od 1-10 su redom TRIP13, S1PR1, GNG11, DQX1, S100A9, COL6A6, POU2AF1, NCAPD2, SLX4 i ADAMTS18. Geni koji imaju ulogu u kontroli staničnog ciklusa često su prekomjerno eksprimirani u tumorima, a njihova prekomjerna ekspresija korelira sa nestabilnošću kromosoma u tumorskim stanicama. Gubitak tumor-supresora, poput Rb ili p53, dovodi do pojačane transkripcije gena koji kontroliraju stanični ciklus putem E2F transkripcijskih faktora. Stanicama sa visokom razinom kromosomske nestabilnosti, koje imaju smanjenu funkciju tumor-supresor gena, potrebna je ekspresija Mad2 gena koji onemogućava napredovanje staničnog ciklusa sve dok se mikrotubuli diobenog vretena ispravno ne poravnaju. Prekomjerna ekspresija Mad2 u stanicama prolongira mitozu. TRIP13, izvorišni gen Klastera 1, bitan je za inaktivaciju kompleksa koji sadrži Mad2 i kritičan je čimbenik za izlazak iz mitoze u Mad2 prekomjerno eksprimirajućim stanicama. Na staničnom modelu pokazano je da prekomjerna ekspresija Mad2, uz smanjenu ekspresiju TRIP13 gena, značajno smanjuje proliferaciju stanica zbog produljene mitoze. Slično je potvrđeno i na modelu ksenografta, gdje je smanjena ekspresija TRIP13 značajno smanjila sposobnost stanica sa prekomjernom ekspresijom Mad2 da formiraju tumor [132]. Naši rezultati pokazuju da su TRIP13 i Mad2 (MAD2L1) prekomjerno eksprimirani u tumoru pluća skvamoznih stanica, a prekomjerna ekspresija Mad2 u tumoru pluća skvamoznih stanica pokazana je i na drugim studijama [133]. Nadalje, na staničnom modelu tumora pluća ne-malih stanica, pokazano je da TRIP13 ima ulogu onkogeni i da potiče proliferaciju i invaziju tumorskih stanica [134,135]. Iako je onkogeni uloga TRIP13 pokazana na tumoru pluća ne-malih stanica, u navedene studije nisu bili uključeni stanični modeli tumora pluća skvamoznih stanica. Koliko nam je poznato, onkogeni mehanizmi TRIP13 nisu do sada ispitivani na ovom tipu tumora, pogotovo u kontekstu prekomjerne ekspresije Mad2 gena, te su potrebna daljnja istraživanja kako bi se utvrdio terapijski potencijal djelovanja na ekspresiju TRIP13 u tumoru pluća skvamoznih stanica sa prekomjernom ekspresijom Mad2.

S1PR1, izvorišni gen Klastera 2, je receptor za sfingozin-fosfat (S1P). Eksprimiran je na vaskularnim endotelnim stanicama i ima važnu ulogu u embrionalnom razvoju vaskularnog

sustava. Međutim, studije ukazuju da S1PR1 i S1P imaju ulogu i u tumorskom mikrookolišu koja se povezuje sa migracijom i polarizacijom makrofaza [136], te proliferacijom i diferencijacijom limfocita [137]. Na tumoru pluća skvamoznih stanica pokazana je smanjena ekspresija S1PR1 koja utječe na lošije preživljenje pacijenata [138]. Nadalje, na mišjem modelu pokazano je da povećana ekspresija S1PR1 ima protu-tumorsko djelovanje te da je odgovor na kemostatik doxorubicin i imunoterapiju efikasniji u miševima sa prekomjerno eksprimiranim S1PR1 genom [139]. Međutim, njegovo djelovanje je i potencijalno pro-tumorigeno. Naime, studija je pokazala da u tumoru pluća ne-malih stanica Apolipoprotein M promovira proliferaciju i invaziju stanica tumora pluća putem povećane ekspresije S1PR1 i aktivacije ERK1/2 i PI3K/AKT signalnih puteva. Obzirom na oprečne rezultate istraživanja, imunomodularnu ulogu S1PR1 u tumoru pluća treba dodatno istražiti u budućim studijama.

Osim visokoumreženih regija, u sklopu ove disertacije, identificirano je i 45 visokoumreženih gena. Većina visokoumreženih gena (36/45) pripada Klasteru 1, što njihovu funkciju povezuje sa regulacijom staničnog ciklusa. Među najznačajnijim visokoumreženim genima koje smo pronašli su CDK1, CDC20, CCNA2, CCNB1, PLK1, KIF11, CDC45, BUB1B i MKI67. Ranije spomenuta studija također je identificirala visokoumrežene gene u tumoru pluća skvamoznih stanica pomoću analize mreže proteinskih interakcija. Kao gene koji mogu imati važnu ulogu u progresiji tumora u navedenoj studiji identificirani su TOP2A, CCNA2, CDC20, AURKA, AURKB i FEN1 geni [140]. Svi navedeni geni, uz iznimku AURKA, identificirani su kao visokoumreženi geni tumora pluća skvamoznih stanica i u ovoj disertaciji, što pokazuje da ovi geni mogu biti ključni u tumorigenezi ove bolesti. Ciklin 2A (CCNA2), osim u regulaciji staničnog ciklusa, također može podupirati epitelno-mezenhimalnu tranziciju i proces metastaziranja [141], a njegovu proliferativnu ulogu može inhibirati hsa- miR-137, kako je pokazano na *in vitro* modelu tumora pluća ne-malih stanica [142]. Geni AURKA i AURKB su članovi obitelji aurora kinaza i imaju važnu ulogu u regulaciji progresije staničnog ciklusa iz G2 u M fazu. AURKB ima ulogu u regulaciji pravilnog poravnavanja i segregacije kromosoma. Na temelju bioinformatičkih analiza, AURKB je identificiran kao važan gen i u tumorigenezi adenokarcinoma pluća [143]. Njegova povećana ekspresija zabilježena je u tumoru pluća ne-malih stanica, uključujući tumor pluća skvamoznih stanica [121] i povezuje se sa lošijim sveukupnim preživljenjem pacijenata [144]. FEN1 pripada obitelji nukleaza specifičnih za strukturu Rad2 i igra ključnu ulogu u održavanju stabilnosti genoma, replikaciji i popravku DNA. FEN1 također podupire progresiju tumora pluća

ne-malih stanica i posreduje rezistenciji na cisplatinu. Supresija FEN1 gena dovodi do smanjene razine DNA replikacije i akumulacije oštećene DNA, što dovodi do induciranja apoptoze [145]. Navedeno upućuje na zaključak da je ekspresija FEN1 neophodna za brzu proliferaciju tumorskih stanica pluća. CCNB2, jedan od visokoumreženih gena u našoj analizi, prekomjerno je eksprimiran i u adenokarcinomu. Na *in vitro* modelu, pokazano je da utišavanje CCNB2 dovodi do inhibicije proliferacije, migracije i invazije staničnog modela adenokarcinoma. Također je pokazano da prekomjerna ekspresija miR-335-5p inhibira proliferaciju i metastaziranje tumorskih stanica te da miR-335-5p može preokrenuti trend promicanja prekomjerne ekspresije CCNB2 na napredak stanica raka [146].

U sklopu ove disertacije zanimalo nas je i dali ekspresija određene skupine gena korelira sa kliničkim parametrima pacijenata, poput primjerice spoli, dobi, stadija tumora, sveukupnog preživljenja i KOPB statusa. Za pronalazak korelacije skupina gena sa kliničkim parametrima korištena je metoda ponderirane korelacijske mreže gena (WGCNA) i matrica sa ekspresijskim vrijednostima gena. Na temelju ekspresijskog profila identificirano je 30 skupina visokokoreliranih gena, odnosno 30 genskih modula te je ispitana njihova povezanost sa kliničkim parametrima. Naši rezultati ukazuju na negativnu korelaciju modula *Darkgreen* i *Darkturquoise* sa statusom pacijenata, negativnu korelaciju *Plum* modula sa spolom, pozitivnu korelaciju *Skyblue* modula sa dobi pacijenata, kao i pozitivnu korelaciju *sienna3* modula sa prisutnošću metastaza. Na temelju dobivenih rezultata, odlučili smo detaljnije istražiti ulogu 115 gena iz modula *sienna3* jer je taj modul pokazao pozitivnu korelaciju sa procesom metastaziranja. Analiza genske ontologije pokazala je da ti geni sudjeluju uglavnom u regulaciji imunološkog odgovora. Usporedbom rezultata ekspresije gena, pokazalo se da većina gena objedinjenih u modulu *sienna3* ima smanjenu ekspresiju u tumorskom tkivu, u odnosu na zdrave kontrole. Obzirom da izbjegavanje imunološkog odgovora, ili stvaranje imunosupresivnog mikrookruženja tumora, pogoduje njegovom daljnjem rastu, nije iznenađujuće da su se u ovom modulu objedinili geni sa smanjenom ekspresijom koji imaju ulogu u regulaciji imunološkog odgovora. Naime, analiza genske ontologije pokazala je da se neodgovarajući imunološki odgovor posredovan limfocitima, NK stanicama, INF γ signalnim putem, kao i procesuiranje i prezentacija endogenih i egzogenih antigena pomoću MHC kompleksa klase I u tumorskom tkivu, povezuju sa procesom metastaziranja. Zanimalo nas je i koji geni su visokoumreženi geni unutar *sienna3* modula, jer oni potencijalno imaju krucijalnu ulogu u procesu metastaziranja tumora pluća skvamoznih stanica. Rezultati su pokazali da su visokoumreženi geni

ovoga modula FGL2, FYN, GIMAP4, GIMAP5, GIMAP7, GIMAP8, HLA-DPB1, HLA-E, KLRF1, PRF1, STAT5B i UNC13D. FGL2 (engl. *Fibrinogen-like protein 2*) ima važnu ulogu u urođenom i adaptivnom imunološkom odgovoru. Djeluje kao efektorna molekula Treg stanica, potiskuje proliferaciju T limfocita, inducira apoptozu u B stanicama, promiče aktivaciju makrofaga i inhibira sazrijevanje dendritičnih stanica [147,148]. Na mišjem modelu glioma, pokazano je da povećana ekspresija FGL2 potiče imunosupresivni mikrookoliš povećanjem ekspresije PD-1, Treg stanica i M2 makrofaga [149]. Nadalje, na mišjem modelu tumora pluća pokazano je da FGL2 iz tumorske strome može inducirati pro-tumorigeni fenotip tumoru-pridruženih fibroblasta [150]. Naši rezultati pokazuju da je ekspresija FGL2 smanjena u eksperimentalnoj kohorti, dok je u TCGA LUSC kohorti ekspresija FGL2 tek blago smanjena ($\log_2FC=-0,3$) u odnosu na zdrave kontrole. Jedna studija pokazala je da je ekspresija FGL2 smanjena u adenokarcinomu pluća i da smanjena ekspresija korelira sa lošijim preživljenjem pacijenata i smanjenom aktivacijom T limfocita, posebice CD8+ T limfocita [151]. Analiza sveukupnog preživljenja u našem istraživanju također je pokazala da je povećana ekspresija FGL2 u tumoru pluća skvamoznih stanica povezana sa boljim preživljenjem pacijenata ($p=0,03$), što podupire rezultate ranije navedene studije na adenokarcinomu. Uloga FGL2 u tumoru i tumorskom mikrookolišu pluća nije dovoljno istražena, posebice u tumoru pluća skvamoznih stanica. Pregledom literature nisu pronađeni raniji radovi koji bi FGL2 povezali sa procesom metastaziranja u tumoru pluća skvamoznih stanica, sugerirajući da je FGL2 potencijalno novi marker progresije ove vrste tumora.

5.3. Utjecaj razine ekspresije gena na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica

U sklopu ove disertacije napravljena je i analiza utjecaja razine ekspresije gena na sveukupno preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica. Utjecaj ekspresije gena na sveukupno preživljenje provjereno je za ukupno 2466 gena čija je razina ekspresije promijenjena na obje testirane kohorte, kao i za 45 identificiranih visokoumreženih gena, 10 *seed* gena unutar PPI mreže te 12 visokoumreženih gena unutar *sienna3* klastera. Obzirom da je eksperimentalna kohorta relativno mala, utjecaj ekspresije gena na preživljenje pacijenata ispitan je na validacijskoj kohorti od 223 pacijenta oboljelih od tumora pluća skvamoznih stanica. Od svih testiranih gena, pronađena je povezanost razine ekspresije ukupno 89 gena sa sveukupnim

preživljenjem pacijenata. Geni čija promjena razine ekspresije utječe na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica prikazani su u Tablici 35.

Tablica 35. Prikaz gena čija se razina ekspresije povezuje sa sveukupnim preživljenjem pacijenata oboljelih od tumora pluća skvamoznih stanica. Geni su prikazani prema p vrijednostima *log-rank* testa, sa najznačajnijim genima istaknutim pri vrhu tablice

Gen	<i>log-rank p</i> vrijednost	Povećana ekspresija,	Gen	<i>log-rank p</i> vrijednost	Povećana ekspresija,
HOXC4	0,001	lošije preživljenje	VNN2	0,004	bolje preživljenje
GALNT2	0,003	lošije preživljenje	TM4SF19	0,008	bolje preživljenje
FAM65A	0,004	lošije preživljenje	GSTZ1	0,009	bolje preživljenje
PYGB	0,004	lošije preživljenje	LILRA1	0,009	bolje preživljenje
PALM3	0,005	lošije preživljenje	MTCH2	0,010	bolje preživljenje
SMOX	0,006	lošije preživljenje	MAD2L1	0,012	bolje preživljenje
RNFT2	0,006	lošije preživljenje	C1orf106	0,013	bolje preživljenje
OLFM2	0,007	lošije preživljenje	SNX31	0,013	bolje preživljenje
LOXL2	0,010	lošije preživljenje	C1orf168	0,013	bolje preživljenje
FSD1	0,010	lošije preživljenje	C14orf169	0,014	bolje preživljenje
L1CAM	0,010	lošije preživljenje	CLCA2	0,014	bolje preživljenje
IL4R	0,010	lošije preživljenje	TSFM	0,015	bolje preživljenje
RGS11	0,011	lošije preživljenje	WDHD1	0,016	bolje preživljenje
NKD2	0,012	lošije preživljenje	ENTPD3	0,017	bolje preživljenje
SLC4A3	0,013	lošije preživljenje	GRHL3	0,017	bolje preživljenje
ARHGEF10	0,013	lošije preživljenje	LILRA5	0,018	bolje preživljenje
TBXA2R	0,013	lošije preživljenje	ROPN1L	0,018	bolje preživljenje
WSF1	0,013	lošije preživljenje	CCDC78	0,020	bolje preživljenje
LAMB3	0,020	lošije preživljenje	THOC3	0,021	bolje preživljenje
COL22A1	0,024	lošije preživljenje	CLEC7A	0,021	bolje preživljenje
EFNB3	0,024	lošije preživljenje	WDR38	0,021	bolje preživljenje
NOC2L	0,024	lošije preživljenje	MCCC1	0,023	bolje preživljenje
SLC35F2	0,025	lošije preživljenje	C3orf58	0,024	bolje preživljenje
DPP6	0,026	lošije preživljenje	CLIC2	0,024	bolje preživljenje
RALGAPA2	0,026	lošije preživljenje	DCPS	0,026	bolje preživljenje
DNAH17	0,027	lošije preživljenje	KEAP1	0,027	bolje preživljenje
SEZ6L2	0,027	lošije preživljenje	DSG3	0,029	bolje preživljenje
CHPF	0,030	lošije preživljenje	EPT1	0,031	bolje preživljenje
FLNC	0,030	lošije preživljenje	CCDC34	0,032	bolje preživljenje
CEP72	0,031	lošije preživljenje	KCNH1	0,033	bolje preživljenje
LAMC3	0,031	lošije preživljenje	KIR3DL1	0,033	bolje preživljenje

Tablica 35. Nastavak

PC	0,033	lošije preživljenje	HELLS	0,034	bolje preživljenje
BCL9	0,035	lošije preživljenje	TMPRSS13	0,034	bolje preživljenje
FAT1	0,036	lošije preživljenje	SLC25A15	0,043	bolje preživljenje
ARMCX1	0,037	lošije preživljenje	FARSA	0,044	bolje preživljenje
SLC9A3R2	0,037	lošije preživljenje	DTYMK	0,046	bolje preživljenje
APLP1	0,038	lošije preživljenje	FGF2	0,046	bolje preživljenje
CASKIN2	0,039	lošije preživljenje	CCNA2	0,048	bolje preživljenje
AGRN	0,040	lošije preživljenje	ACSL1	0,048	bolje preživljenje
MACF1	0,040	lošije preživljenje	STGGAL2	0,049	bolje preživljenje
BLCAP	0,041	lošije preživljenje			
RNF146	0,041	lošije preživljenje			
PNPLA2	0,044	lošije preživljenje			
HOXC6	0,046	lošije preživljenje			
DUSP8	0,046	lošije preživljenje			
DDX11	0,047	lošije preživljenje			
TCF3	0,047	lošije preživljenje			
MPV17L	0,048	lošije preživljenje			

Naše rezultate povezanosti ekspresije gena sa preživljenjem pacijenata oboljelih od tumora pluća skvamoznih stanica, potvrdile su i druge studije. Primjerice, pokazano je da RNFT2, gen koji sudjeluje u regulaciji urođenog imunološkog odgovora negativnom regulacijom IL-3R α [152], kao dio seta gena sa prognostičkim potencijalom, može utjecati na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica [153]. Utjecaj povišene ekspresije L1CAM [154,155] i LOXL2 gena [156] na lošije preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica također je potvrđeno drugim studijama. Nadalje, rezultati naše analize pokazuju povezanost povišene ekspresije MAD2L1 gena sa boljim preživljenjem pacijenata. Prognostički potencijal ovoga gena potvrdila je studija na tumoru pluća ne-malih stanica [157]. Rezultati navedene studije pokazuju da je povišena ekspresija ovoga gena povezana sa lošijim preživljenjem pacijenata, što je rezultat suprotan od pronađenog u našoj analizi. Razlog nesuglasja rezultata može biti posljedica toga što su autori ispitivali utjecaj razine MAD2L1 gena na uzorcima tumora pluća ne-malih stanica zajedno, a ne zasebno za svaki podtip tumora pluća. Nadalje, studije na adenokarcinomu, kao i ostalim podtipovima tumora pluća ne-malih stanica, pronašli su povezanost povišene razine ekspresije GALNT2 [158], PYGB [159] i SMOX [160] gena sa lošijim preživljenjem pacijenata, što je u skladu sa našim rezultatima. Koliko nam je poznato, razina njihove ekspresije do sada se

nije povezivala sa preživljenjem pacijenata oboljelih od tumora pluća skvamoznih stanica. Nadalje, pregled literature ukazuje da su geni HOXC4, C1orf106, SNX31 i, ranije diskutirani FGL2, potencijalni novi prognostički biomarkeri tumora pluća skvamoznih stanica. HOX geni pripadaju obitelji *homeobox* transkripcijskih faktora. Iako se njihova primarna funkcija povezuje sa regulacijom embrionalnog razvoja, pokazano je da je njihova ekspresija u odrasloj dobi tkivno specifična i da deregulacija njihove ekspresije može podupirati tumorigenezu [161]. Povećana ekspresija HOX gena, uključujući HOXC4, pokazana je i u tumoru pluća ne-malih stanica. Nadalje, pokazano je da blokiranjem vezanja HOX gena i njegovog ko-faktora PBX dolazi do smanjene aktivnosti HOX proteina i apoptoze na staničnom modelu adenokarcinoma, te smanjenog rasta adenokarcinoma na mišjem modelu [162]. Rezultati navedene studije ukazuju da interakcija HOX i PBX proteina predstavlja potencijal za razvoj novih terapija. Također, nekoliko studija pokazalo je da ekspresija HOX gena može biti regulirana od strane miRNA [163,164], te da je razina ekspresije HOX gena povezana sa preživljenjem pacijenata oboljelih od glioblastoma [165] i tumora želuca [166]. Obzirom na mali broj studija koje su proučavale funkciju HOX gena, uloga HOXC4 gena u tumoru pluća skvamoznih stanica nije poznata i razjašnjena. Stoga su potrebne daljnje funkcionalne studije koje će detaljnije istražiti mehanizme djelovanja HOXC4 gena u ovoj bolesti. Ekspresija C1orf106 gena, poznatog i kao aktivator urođene imunosti (engl. innate immunity activator, INAVA) povezuje se sa rizikom od nastajanja Crohn-ove bolesti. Iako uloga ovoga gena u tumorigenezi pluća nije poznata, pronađena je njegova povećana ekspresija u adenokarcinomu pluća [167], indicirajući da bi mogao imati ulogu u tumorigenezi pluća. Povećana ekspresija C1orf106 gena do sada je povezana sa tumorigenezom i metastaziranjem papilarnog tumora štitnjače. Predloženi mehanizam na koji C1orf106 podupire agresivnost papilarnog tumora štitnjače podrazumijeva prekomjernu ekspresiju ovoga gena koja utječe na povišenje ekspresije faktora rasta fibroblasta (FGF1) što posljedično uzrokuje povišenu razinu metaloproteinaze 9 (MMP9) [168]. Rezultati ove disertacije također pokazuju povišenu ekspresiju navedenih gena u tumoru pluća skvamoznih stanica, indicirajući da bi ovaj mehanizam mogao imati ulogu i u tumorigenezi pluća skvamoznih stanica.

5.4. Profil tumorskog mikrokoliša i utjecaj na preživljenje pacijenata

Anatomske i stanične karakteristike normalnih pluća služe kao obrambena barijera protiv stranih patogena i čestica. Međutim, u upalnim stanjima kao što je kronična opstruktivna plućna

bolest (KOPB), mikrookoliš pluća pokazuje značajke koje mogu podržati karcinogenezu. Stoga je u posljednjem desetljeću prepoznata središnja uloga tumorskog mikrookoliša (TME) u inicijaciji i progresiji tumora pluća [55]. TME predstavlja zanimljivu nišu za istraživanje novih biomarkera i potencijalnih meta novih lijekova. Lijekovi koji djeluju na različite komponente TME-a, uključujući vaskularni endotelni faktor rasta (VEGF), aromatazu i imunološke kontrolne točke, već su odobreni za kliničku upotrebu. U tumoru pluća ne-malih stanica, infiltracija tumora imunološkim stanicama ovisi o stadiju, što sugerira da TME može imati prognostički značaj [57]. Imunohistokemija, kao jedna od najčešće korištenih metoda za procjenu stupnja infiltriranosti tumora imunološkim stanicama, je jedina metoda u širokoj upotrebi koja daje kvantitativne informacije zajedno sa informacijom o prostornoj lokalizaciji infiltrata u tumorsko tkivo. U posljednjih nekoliko godina sve više studija umjesto imunohistokemije koristi visokoprotodne metode, poput sekvenciranja transkriptoma, za kvantifikaciju stupnja infiltracije imunoloških stanica u tumorsko tkivo, iako ove metode ne pružaju informaciju o prostornoj lokalizaciji infiltrata. Zahvaljujući takvim studijama, prikupljeno je puno informacija o profilu stanica koje infiltriraju različite vrste tumora, kao i informacije o utjecaju imunoloških infiltrata na preživljenje pacijenata. Primjerice, jedna studija istraživala je profil TME-a u 33 vrste tumora bioinformatičkim metodama, koristeći rezultate sekvenciranja slijedeće generacije, u svrhu identifikacije imunoloških podtipova i utjecaja stupnja infiltracije na preživljenje. Rezultati navedene studije pokazali su grupiranje tumorskih uzoraka u 6 skupina, nazvanih Imunološki podtipovi C1-C6. Karakterizacija identificiranih podtipova pokazala je da uzorci grupirani u podtip C1 imaju povišenu ekspresiju gena uključenih u proces angiogeneze, visoki stupanj proliferacije stanica, kao i pristranost Th2 stanica prema infiltraciji stanicama stečene imunosti. C2 podtip karakterizira polarizacija M1/M2 makrofaga, jaka infiltracija CD8⁺ T limfocitima, raznolikost receptora T limfocita, kao i povećan stupanj proliferacije. C3 podtip pokazuje povišenu ekspresiju Th17 i Th1 gena, niski/srednji stupanj proliferacije i manji broj somatskih promjena u odnosu na druge podtipove. C4 podtip okarakteriziran je supresijom Th1 stanica i M2 fenotipom makrofaga. C5 podtip pokazuje najmanji stupanj infiltracije sa Th1 stanicama i najveći stupanj infiltracije M2 makrofazima. C6 podtip pokazao je visoki stupanj infiltracije T limfocitima tipa 1 i 2 [169]. Uzorci tumora pluća skvamoznih stanica u njihovoj studiji uglavnom su se grupirali u C1 podtip, zajedno sa tumorom dojki i skvamoznim tumorom glave i vrata. Nadalje, pokazali su da se grupiranje u identificirane imunosne podtipove povezuje sa preživljenjem pacijenata te da C3 podtip ima

očekivano najbolju prognozu, a C4 i C6 najlošiju očekivanu prognozu. Druga studija također je identificirala 6 imunskih podtipova u tumorima skvamoznih stanica. *Immune-cold* podtip imao je najmanju razinu infiltracije tumora limfocitima i visoki stupanj aneuploidije. Pacijenti grupirani u ovaj podtip imali su najlošiju očekivanu prognozu. *Immune-hot* podtip, s druge je strane pokazao najveći stupanj infiltracije tumora CD8⁺ limfocitima, aktiviranim NK stanicama i pojačani odgovor na IFN- γ . Pacijenti grupirani u ovaj podtip imali su najbolju očekivanu prognozu [170]. Kao što je vidljivo iz navedenih studija, stupanj infiltracije tumora imunološkim stanicama može imati utjecaj na preživljenje pacijenata. U sklopu ove disertacije, stoga nas je zanimala razina infiltracije imunoloških stanica u tumorsko tkivo za sve pacijente iz eksperimentalne kohorte. Inicijalna procjena razine infiltriranosti tumorskog tkiva sa stromalnim i imunološkim stanicama napravljena je pomoću *ESTIMATE* paketa u R programu. Obzirom da su rezultati ove analize ukazivali da analizirani tumori imaju različiti stupanj infiltracije navedenim stanicama, napravili smo detaljniju procjenu tumorskog mikrookoliša pomoću *GSVA* metode i 16 skupina gena specifičnih za određene stanice imunološkog sustava. Obzirom da se smatra da aktivirani T limfociti, gama delta T stanice i NK stanice imaju najveću ulogu u suzbijanju tumorskih stanica, geni specifični za ove stanice objedinjeni su u zaseban set gena nazvan Citotoksične stanice. Na temelju izračunatih *GSVA* vrijednosti obogaćivanja i grupiranja gena klasteriranjem, pronađene su 4 podskupine pacijenata, nazvanih Podtip1-4, koji se međusobno razlikuju u profilu i stupnju infiltracije imunološkim stanicama. Naši rezultati pokazali su da pacijenti grupirani u Podtip 1 i Podtip 2 imaju manju razinu infiltriranosti tumora sa NK *bright*, T folikularnim pomoćničkim stanicama i eozinofilima u odnosu na Podtip 4. Obzirom na ulogu navedenih stanica, izgledno je da Podtipovi 1 i 2 manju protu-tumorsku aktivnost, u odnosu na Podtip 3 i 4, koji imaju više infiltriranih efektornih i centralnih memorijskih T stanica, CD8⁺ T limfocita i eozinofila. Veći procijenjeni stupanj infiltracije tumora navedenim stanicama u Podtipu 3 i 4 ukazuje na potencijalnu veću protu-tumorsku aktivnost. Nadalje, zanimljivo je primjetiti da su svi podtipovi podjednako (ne)infiltrirani dendritičnim stanicama. Razlog tome možda leži u činjenici da regulacija diferencijacije dendritičnih stanica ovisi o aktivaciji PI3K/PKB, JAK/STAT, MAPK, i NF- κ B staničnim putevima, koji su generalno smanjeno aktivirani u našoj eksperimentalnoj kohorti. Također, rezultate naše analize teško je direktno uspoređivati sa objavljenim studijama, iz razloga što je naša kohorta relativno mala te su korišteni drugačiji setovi gena za identifikaciju podtipova. Primjerice, navedene studije koristile su setove gena specifične za M1 i M2 fenotip makrofaga,

dok u našu analizu nismo uključili gene koji bi diferencirali ove dvije skupine makrofaga, kao niti gene koji bi razdvojili Th1 od Th17 stanica. S druge strane, u našu analizu uključeni su geni koji razdvajaju NK stanice na dva fenotipa, *NKbright* i *NKdim*, koje navedene studije nisu uključile u svoja istraživanja.

Obzirom da se u literaturi veća razina infiltracije tumorskog tkiva sa imunološkim stanicama često može povezati sa preživljenjem pacijenata, zanimalo nas je dali pripadnost uzoraka eksperimentalne kohorte određenom imunosnom podtipu ima utjecaj na sveukupno preživljenje pacijenata. Napravljena je analiza sveukupnog preživljenja Kaplan-Meierovom metodom u GraphPad programu, a razlika u preživljenju između podtipova ispitana je *log-rank* testom. Rezultati analize preživljenja pokazali su da pripadnost određenom podtipu nema značajan utjecaj na preživljenje pacijenata ($p=0,09$). Iako nema statistički značajne razlike u sveukupnom preživljenju pacijenata, zanimljivo je da su svi pacijenti grupirani u Podtip3 bili živi tijekom zadnjeg praćenja. U Podtipu3 je pronađena najviša razina infiltriranosti tumora citotoksičnim stanicama, a najmanja razina infiltriranosti makrofazima, što sugerira najbolji protu-tumorski odgovor u odnosu na ostale podtipove. Međutim, važno je naglasiti da je eksperimentalna kohorta dosta mala, stoga je potrebno u budućim studijama provjeriti reproducibilnost pronađenih podtipova na većem broju pacijenata, kako bi se mogao izvući konkretniji zaključak.

5.5. Profil ekspresije miRNA u tumoru pluća skvamoznih stanica

Za proces tumorigeneze odgovorne su promjene u staničnim procesima uzrokovane, ne samo mutacijama ili promjenama u razini ekspresije gena, već i promjenama u epigenetskoj regulaciji staničnih procesa. Jedan od mehanizama epigenetske regulacije staničnih procesa je regulacija putem ekspresije miRNA, koje mogu inhibirati učinak genske ekspresije vezanjem i razgradnjom miRNA na mRNA transkript, ili inhibirati translaciju mRNA transkripta u protein. Većina miRNA može regulirati ekspresiju više gena, kao što i više miRNA može regulirati ekspresiju istog gena. Studije su pokazale da miRNA imaju bitnu ulogu u različitim biološkim procesima, uključujući diferencijaciju i proliferaciju stanica te apoptozu. Nadalje, pokazano je da miRNA imaju dvojaku ulogu u tumorigenezi pluća promovirajući ekspresiju onkogeni ili inhibirajući tumor-supresorske gene [171–175]. Kako bi bolje razumijeli promjene u staničnim procesima tumora pluća skvamoznih stanica na epigenetskoj razini, istražili smo profil promijenjene ekspresije miRNA u tumorskom tkivu, u odnosu na zdravo tkivo pluća. Pomoću

analize diferencijalne ekspresije, zabilježili smo promjenu u razini ekspresije ukupno 122 miRNA, od kojih 75 miRNA ima povišenu ekspresiju, a 47 miRNA ima smanjenu ekspresiju u tumorskom tkivu. Među pronađenim miRNA sa promijenjenom ekspresijom u tumorskom tkivu, mnoge su već poznate i okarakterizirane u drugim studijama [176]. Primjerice, pokazano je da smanjena ekspresija hsa-miR-126 inhibira proliferaciju tumora pluća ciljajući EGFL7 i SLC7A5 gene [177,178]. U našoj studiji hsa-miR-126 također pokazuje smanjenu ekspresiju u tumorskom tkivu koja korelira sa povećanom ekspresijom SLC7A5 gena, indicirajući da je hsa-miR-126 potencijalni tumor-supresor u tumoru pluća skvamoznih stanica. Nadalje, hsa-miR-21-3p i hsa-miR-21-5p su također prekomjerno eksprimirane u tumorskom tkivu naše kohorte, a prijašnje studije okarakterizirale su ih kao onkomire u tumoru pluća [179,180]. Najznačajnije promjene u ekspresiji miRNA u našoj analizi pokazali su hsa-miR-708-5p i hsa-miR-708-3p. Nekoliko drugih studija također je potvrdilo naše rezultate povišene ekspresije navedenih miRNA na tumoru pluća skvamoznih stanica, kao i na adenokarcinomu pluća [181–183]. Studija koja je pratila promjenu u profilu ekspresije miRNA od zdravog tkiva do tkiva tumora pluća skvamoznih stanica, u svojim istraživanjima pokazala je prekomjernu ekspresiju hsa-miR-944 i hsa-miR-31-5p, te smanjenu ekspresiju hsa-miR-34c-3p, hsa-miR-145-5p, hsa-miR-486-5p i hsa-miR-133a-3p [184]. Rezultati našeg istraživanja potvrđuju promjene u ekspresiji navedenih miRNA u tumoru pluća skvamoznih stanica. Navedena studija također sugerira da je prekomjerna ekspresija hsa-miR-183 potencijalni onkogeni pokretač tumora pluća skvamoznih stanica jer podupire formaciju, rast i progresiju tumora. Naši rezultati potvrđuju prekomjernu ekspresiju ove miRNA u tumoru pluća skvamoznih stanica. Hsa-miR-577 također ima opisanu ulogu u tumorigenezi pluća. Primjerice, pokazano je na staničnom modelu adenokarcinoma pluća da je u tumoru pluća ne-malih stanica smanjena ekspresija hsa-miR-577 koja korelira sa povećanom ekspresijom WNT2B gena i aktivacijom Wnt/ β -kateninskog signalnog puta [185], te pojačanom proliferacijom i invazijom stanica povezanom sa pojačanom ekspresijom HOXA1 gena [186]. Naši rezultati pokazuju povećanu ekspresiju hsa-miR-577 u tumoru pluća skvamoznih stanica, te nepromijenjenu ekspresiju WNT2B i povećanu ekspresiju HOXA1 gena. Navedeno sugerira da aktivacija Wnt signalnog puta, kao i ekspresija HOXA1 gena nije pod regulacijom hsa-miR-577 u tumoru pluća skvamoznih stanica. Druga studija pokazala je povećanu ekspresiju hsa-miR-577 na tumoru pluća ne-malih stanica i predložila GNG11, EFEMP1, SCEL, GKN2 i ATG16L2 kao ciljene gene navedene miRNA [187]. Rezultati naše studije pokazuju da je ekspresija navedenih gena negativno promijenjena u tumoru pluća

skvamoznih stanica, sugerirajući da bi hsa-miR-577 mogla biti regulator ekspresije navedenih gena u ovom tipu tumora. Rezultati naše studije također pokazuju da miRNA sa sniženom ekspresijom u tumoru pluća skvamoznih stanica utječu na gene koji imaju funkciju u DNA replikaciji i popravku DNA, organizaciji vanstaničnog matriksa i staničnom ciklusu. Hsa-miR-107 pokazuje sniženu ekspresiju u našoj studiji. Zanimljivo je primjetiti da hsa-miR-107 cilja 40 od 45 gena identificiranih kao visokumreženi geni unutar PPI mreže, indicirajući da hsa-miR-107 može imati krucijalnu ulogu u regulaciji staničnog ciklusa u tumoru pluća skvamoznih stanica. Druge studije potvrdile su tumor-supresorsku ulogu miRNA-107 u drugim tumorima, uključujući tumor pluća ne-malih stanica [188–192]. Pregledom literature, utvrdili smo da potencijalna tumor-supresorska uloga do sada nije pronađena u tumor pluća skvamoznih stanica. Također, signalni put E2F1-FOXM1 je pod potencijalnom regulacijom miRNA. Naime, pronašli smo negativnu korelaciju ekspresije FOXM1 gena i hsa-miR-107, hsa-miR-130a-3p i hsa-miR-125a-5p. Osim regulacije staničnog ciklusa, naši rezultati pokazuju i da miRNA sa povećanom ekspresijom utječu na smanjenu ekspresiju gena koji su povezani sa imunološkim odgovorom, posebice imunološkim odgovorom posredovanim neutrofilima.

5.6. Profil ekspresije proteina u tumoru pluća skvamoznih stanica

Za bolje razumijevanje procesa tumorigeneze pluća, izuzetno je važna i karakterizacija profila ekspresije proteina, obzirom da su to ključne izvršne molekule za funkciju stanice. Korištenje visokoprotočnih metoda za profiliranje proteina, poput spektrometrije masa, omogućuje identifikaciju većeg broja proteina u uzorku u odnosu na druge manje osjetljive metode. U sklopu ove disertacije, profiliranje proteinske ekspresije napravljeno je u suradnji sa Laboratorijem za sistemsku biologiju i prijenos signala iz njemačkog Centra za istraživanje raka. Obzirom da je profiliranje proteinske ekspresije iz tkiva uklopljenih u parafin još u povojima, ovaj dio istraživanja napravljen je na razini pilot studije u koju smo uključili 3 tkiva tumora pluća skvamoznih stanica i 3 tkiva zdravih pluća. U svrhu ispitivanja profila proteinske ekspresije koji doprinosi fenotipu tumora pluća skvamoznih stanica, napravljena je analiza diferencijalne ekspresije proteina pomoću *Perseus* programa. Rezultati su pokazali povećanu razinu ekspresije 1404 proteina, dok je smanjena razina ekspresije pokazana za 271 protein. Najveća razlika u ekspresiji između testiranih skupina detektirana je za proteine VIM ($\log_2FC=8,69$) i CA1 ($\log_2FC=-12,28$). Karbonatna anhidraza (CA1) pripada obitelji enzima koji imaju važnu ulogu u transportu iona, regulaciji

stanične pH, retenciji Na⁺, kalcifikaciji, prijenosu signala, glukoneogenezi, ureagenezi i lipogenezi. Povišene razine CA1 pronađene su u tumorima debelog crijeva, bubrega i melanoma [193]. Također, pokazano je da povećana ekspresija CA1 doprinosi tumorigenezi i migraciji stanica tumora dojke [194]. Nadalje, povišena ekspresija CA1, pronađena je u serumu pacijenata oboljelih od tumora pluća ne-malih stanica, i dovodi se u korelaciju sa hipoksičnim mikrookolišem tumora [195]. Naši rezultati ne potvrđuju rezultate drugih studija, obzirom da je ekspresija CA1 značajno smanjena u tumoru pluća skvamoznih stanica. Naime, mehanizam djelovanja i posljedice promjenjene ekspresije proteina CA1 u tumoru pluća skvamoznih stanica još nije opisan, stoga su potrebne daljnje studije koje bi bolje okarakterizirale ulogu CA1 u ovom tipu tumora. Vimentin (VIM) kao glavna komponenta intermedijarnih filamenata održava integritet stanica [196]. Nalazi se u fokalnim adhezijama koje povezuju stanični citoskelet sa vanstaničnim matriksom, a studije su pokazale da ima ulogu u migraciji i invaziji stanica tijekom epitelno-mezenhimalne tranzicije [197,198]. Do sada je pokazana povezanost povišene ekspresije vimentina i lošije prognoze pacijenata oboljelih od različitih vrsta tumora, uključujući tumor dojke [199], želuca [200] i pluća [201]. Rezultati naše analize također su pokazali povišenu ekspresiju vimentina na proteinskoj razini. Nedavno objavljena studija na adenokarcinomu [202] sugerira da aktivnost vimentina može biti post-translacijski regulirana fosforilacijom potaknutom od PLK1 proteina koji se u našim analizama pokazao kao visokoumrežen gen unutar PPI mreže. Međutim, točni regulatorni mehanizmi utjecaja povišene ekspresije vimentina na epitelno-mezenhimalnu tranziciju, kao i uloga u procesu metastaziranja tumora pluća skvamoznih stanica tek trebaju biti okarakterizirani. Naši rezultati također pokazuju smanjenu ekspresiju proteina peroksiredoksina 2 (PRDX2), međutim druga istraživanja pronašla su povišene razine ovog proteina u tumorima. Primjerice, pokazano je da tumori pluća pokazuju povišenu razinu PRDX2 koja korelira sa lošijim preživljenjem pacijenata, prvenstveno oboljelih od adenokarcinoma [203]. Također, na staničnom modelu pokazano je da prekomjerna ekspresija PRDX2 potiče proliferaciju, migraciju i invaziju tumorskih staničnih linija, sugerirajući da PRDX2 ima protumorigenu ulogu u tumoru pluća [203,204]. Mehanizmi kojima promjenjena ekspresija PRDX2 utječe na proliferaciju i invazivnost tumorskih stanica, koliko je nama poznato, do sada nije istraživana u tumoru pluća skvamoznih stanica. Versican (VCAN) je važna komponenta vanstaničnog matriksa, a do sada je pokazano da ima ulogu u tumorigenezi potičući rast, migraciju i invaziju tumorskih stanica, kao i angiogenezu i metastaziranje [205]. Također, pokazano je da VCAN ima potencijal prognostičkog biomarkera.

Primjerice, prekomjerna ekspresija VCAN gena povezana je sa lošijim preživljenjem pacijenata oboljelih od tumora želuca [206] i mjehura [207]. Studija koja je profilirala tumore pluća skvamoznog karcinoma i adenokarcinoma [208] pronašla je prekomjernu ekspresiju VCAN proteina u tumoru pluća skvamoznih stanica. Naši rezultati pokazuju isti trend ekspresije u ovom tipu tumora. Međutim, i navedena studija, kao i naša, uključivala je tri testirana uzoraka tumora pluća skvamoznih stanica, te je potrebno dodatno validirati ekspresiju ovog proteina na većoj kohorti pacijenata.

5.7. Interakcija mRNA, proteina i miRNA u tumoru pluća skvamoznih stanica

Rezultati ove doktorske disertacije pokazuju da su keratini prekomjerno eksprimirani u tumoru pluća skvamoznih stanica na mRNA razini. Međutim, većinu keratina detektiranih na transkripcijskoj razini, nismo detektirali na proteinskoj razini. Prekomjernu ekspresiju na obje razine potvrdili smo za KRT6B, KRT18 i KRT19. Nadalje, MYCL, član predloženog mehanizma koji potencijalno doprinosi proliferaciji tumorskih stanica preko KRT6A-MYCL-G6P osi, nije detektiran na proteinskoj razini, dok je za G6P uočeno blago povišenje ekspresije u tumoru ($\log_2FC=0,58$). Naši rezultati sugeriraju da bi hsa-let-7c-5p mogla inhibirati translaciju gena MYCL, što posljedično utječe na razinu ekspresije gena G6P. Rezultati profiliranja transkriptoma također ukazuju da su geni koji sudjeluju u regulaciji staničnog ciklusa prekomjerno eksprimirani na mRNA razini, međutim nismo uspjeli pokazati prekomjernu ekspresiju na proteinskoj razini za sve gene. Konzistentnu ekspresiju na obje razine pokazali smo za 245 gena/proteina sa prekomjernom ekspresijom i 144 sa smanjenom ekspresijom. Primjerice, članovi predložene osi koja također doprinosi proliferaciji tumorskih stanica, E2F1-FOXM1-ciljni geni FOXM1 transkripcijskog faktora, nisu detektirani na proteinskoj razini. Geni Skp2, CCNA2, CCNB1, Cdc25A, AURKB, BIRC5 i CENPA prekomjerno su eksprimirani na transkripcijskoj razini, dok njihova ekspresija na proteinskoj razini nije detektirana. I drugi geni uključeni u regulaciju staničnog ciklusa nisu identificirani na proteinskoj razini, poput gena PLK1, MAD2L1, i CCNB2, a prekomjerno su eksprimirani na mRNA razini. S druge strane potvrđena je prekomjerna ekspresija gena FEN1 i na proteinskoj razini. Rezultati transkriptomskih analiza pokazuju da su prekomjerno eksprimirani geni koji sudjeluju u Wnt, Myc i Hedgehog staničnim putevima, poznatim putevima koji podupiru tumorigeni fenotip [209–211]. Primjerice, na mRNA razini povišena je ekspresija članova Wnt signalnog puta, poput gena WNT3, WNT5A, WNT7B,

WNT10A, FZD1, FZD6-9. Niti jedan od navedenih gena nije pokazao ekspresiju na proteinskoj razini. S druge strane, na proteinskoj razini potvrdila se prekomjerna ekspresija proteina koji sudjeluju u Myc signalnom putu, poput primjerice MCM2, MCM4, MCM7, TRIM28, CTSP1 i EIF3B. Razlog velikog broja gena detektiranih na mRNA razini koji nisu detektirani na proteinskoj razini, vjerojatno leži u činjenici da je spektrometrijom masa identificiran relativno maleni broj proteina u odnosu na broj detektiranih gena. Primjerice, u sva tri uzorka tumorskog tkiva detektirano je 3670 proteina, dok je broj detektiranih proteina u sva tri uzorka zdravog tkiva još manji, 2061. Te brojke su deset puta manje od broja detektiranih gena na RNA-Seq-u. Također, u našem istraživanju sve analize rađene su na *FFPE* tkivima, koja su dosta zahtjevnja za pripremu za visokoprotlačne analize. Međutim, kompleti za pripremu knjižnica za sekvenciranje na neki način kompenziraju nedostatke *FFPE* tkiva, i olakšavaju daljnje analize. S druge strane, korištenje *FFPE* uzoraka za analizu na spektrometru masa je još u povojima. Nezna se dovoljno o utjecaju fiksacije formalinom i parafinom na analizu proteoma, što je posebice problem kada se istražuju proteinske modifikacije. Potrebna su daljnja proteomska istraživanja na većoj kohorti koja će detaljnije istražiti proteom ove vrste tumora.

Rezultati ove disertacije također pokazuju da 89 gena sa promijenjenom razinom ekspresije ima utjecaj na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica. Od 89 identificiranih gena, rezultati ove disertacije ukazuju da su geni GALNT2, PYGB, SMOX, HOXC4, C1orf106, SNX31 i FGL2 novi potencijali prognostički biomarkeri tumora pluća skvamoznih stanica, a FGL2 i potencijalni marker progresije bolesti. Također, identificirali smo hsa-miR-107, hsa-miR-126 i hsa-let-7c-5p kao miRNA koje djeluju kao potencijalni tumor-supresori u tumoru pluća skvamoznih stanica. Rezultati našeg istraživanja pokazali su da navedene miRNA imaju sniženu ekspresiju, a njihovi ciljni geni su prekomjerno eksprimirani i sudjeluju u regulaciji staničnog ciklusa. Obzirom da njihova uloga u ovom tipu tumora još nije okarakterizirana, potrebne su daljnje studije koje će potvrditi naše rezultate.

6. ZAKLJUČCI

6. ZAKLJUČCI

1. Tumori pluća skvamoznih stanica prekomjerno aktiviraju stanične signalne puteve NP63, FOXM1, p53, PLK1 i AURORA/B, kao i Wnt, Myc i Hedgehog signalni putevi
2. U tumorima pluća skvamoznih stanica smanjeno su aktivirani stanični putevi koji reguliraju imunološki odgovor
3. Prekomjerna aktivacija KRT6A-MYCL-G6P, kao i E2F1-FOXM1-AURKB ili PLK1 osi potencijalno doprinosi tumorigenezi pluća
4. Na temelju stupnja povezanosti gena sa promijenjenom razinom ekspresije u tumoru pluća skvamoznih stanica, identificirano je 11 klastera gena unutar PPI mreže.
5. Unutar PPI mreže identificirana su 45 visokoumrežena gena (engl. *hub genes*), poput CDK1, CDC20, CCNA2, CCNB1 KIF11. Od 45 *hub* gena, 35 ih se grupiralo u Klaster 1.
6. Smanjena ekspresija gena GIMAP7, GIMAP4, FGL2, UNC13D, GIMAP8, GIMAP5, KLRF1, PRF1, HLA-DPB1, STAT5B, HLA-E i FYN doprinosi metastatskom fenotipu tumora pluća skvamoznih stanica
7. U tumoru pluća skvamoznih stanica identificirana su 4 imunosna podtipa, a pripadnost određenom imunosnom podtipu ne utječe na sveukupno preživljenje pacijenata.
8. Ekspresija 89 gena povezuje sa prognozom pacijenata oboljelih od tumora pluća skvamoznih stanica.
9. Najznačajnije pozitivno promijenjene miRNA u tumoru su hsa-miR-577 i hsa-miR-615-5p, a najznačajnije negativno promijenjene miRNA su hsa-miR-490-3p i hsa-miR-1251-5p.
10. Ciljni geni prekomjerno eksprimiranih miRNA imaju ulogu u aktivaciji i degradaciji neutrofila te regulaciji apoptoze u endotelnim stanicama.
11. Ciljni geni miRNA sa smanjenom ekspresijom imaju ulogu u DNA metaboličkim procesima, DNA replikaciji i popravku DNA.
12. Na proteinskoj razini, najveća prekomjerna ekspresija u tumorskom tkivu zabilježena je za VIM, RRBP1, PRKDC, CKAP4 i VCAN proteine.

- 13.** Najmanja ekspresija u tumorskom tkivu zabilježena je za proteine CA1, HBA1, HBD, HIST1H4A i HP.
- 14.** Na proteinskoj razini potvrđena je prekomjerna ekspresija 245 gena, i smanjena ekspresija za 144 gena.
- 15.** Povećana ekspresija 57 proteina potencijalno je rezultat smanjene ekspresije određenih miRNA.
- 16.** Od 89 gena sa prognostičkim potencijalom, podudaranje sa povećanom razinom mRNA ekspresije pronađeno je za 9 proteina.
- 17.** GALNT2, PYGB, SMOX, HOXC4, C1orf106, SNX31 i FGL2 su novi potencijalni prognostički markeri tumora pluća skvamoznih stanica
- 18.** hsa-miR-107, hsa-miR-126 i hsa-let-7c-5p su nove potencijalne tumor-supresorske miRNA u tumoru pluća skvamoznih stanica

7. LITERATURA

7. LITERATURA

1. I., A. *Primary malignant growth of the lung and bronchi.*; New York: Longman, Green, Company, 1912;
2. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **2018**, 68, 394–424, doi:10.3322/caac.21492.
3. Howlader N, N.A.K.M. et al (eds) *SEER Cancer Statistics Review, 1975-2016*; 2019;
4. Ahmadzada, T.; Kao, S.; Reid, G.; Boyer, M.; Mahar, A.; Cooper, W. An Update on Predictive Biomarkers for Treatment Selection in Non-Small Cell Lung Cancer. *J. Clin. Med.* **2018**, 7, 153, doi:10.3390/jcm7060153.
5. Organization, W.H. *Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019.*; 2020;
6. Za Rak, R.; Hrvatske, R. *HRVATSKI ZAVOD ZA JAVNO ZDRAVSTVO INCIDENCIJA RAKA U HRVATSKOJ CANCER INCIDENCE IN CROATIA*; 2015;
7. WHO World Health Organization (WHO). *Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019.* **2020**.
8. Allemani, C.; Matsuda, T.; Di Carlo, V.; Harewood, R.; Matz, M.; Nikšić, M.; Bonaventure, A.; Valkov, M.; Johnson, C.J.; Estève, J.; et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* **2018**, 391, 1023–1075, doi:10.1016/S0140-6736(17)33326-3.
9. Planchard, D.; Popat, S.; Kerr, K.; Novello, S.; Smit, E.F.; Faivre-Finn, C.; Mok, T.S.; Reck, M.; Van Schil, P.E.; Hellmann, M.D. *by the ESMO Guidelines Committee Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up* † 29 (suppl 4): iv192-iv237; 2018;
10. Cooper, W.A.; O'toole, S.; Boyer, M.; Horvath, L.; Mahar, A. What's new in non-small cell lung cancer for pathologists: The importance of accurate subtyping, EGFR mutations

- and ALK rearrangements. *Pathology* **2011**, *43*, 103–115, doi:10.1097/PAT.0b013e328342629d.
11. Hammerman, P.S.; Voet, D.; Lawrence, M.S.; Voet, D.; Jing, R.; Cibulskis, K.; Sivachenko, A.; Stojanov, P.; McKenna, A.; Lander, E.S.; et al. Comprehensive genomic characterization of squamous cell lung cancers. *Nat. 2012 4897417* **2012**, *489*, 519–525, doi:10.1038/nature11404.
 12. Collisson, E.A.; Campbell, J.D.; Brooks, A.N.; Berger, A.H.; Lee, W.; Chmielecki, J.; Beer, D.G.; Cope, L.; Creighton, C.J.; Danilova, L.; et al. Comprehensive molecular profiling of lung adenocarcinoma. *Nat. 2014 5117511* **2014**, *511*, 543–550, doi:10.1038/nature13385.
 13. Bernhardt, E.B.; Jalal, S.I. Small Cell Lung Cancer. *Cancer Treat. Res.* **2016**, *170*, 301–322, doi:10.1007/978-3-319-40389-2_14.
 14. Osmani, L.; Askin, F.; Gabrielson, E.; Li, Q.K. Current WHO Guidelines and the Critical Role of Immunohistochemical Markers in the Subclassification of Non-Small Cell Lung Carcinoma (NSCLC). Moving from Targeted Therapy to Immunotherapy., doi:10.1016/j.semcancer.2017.11.019.
 15. Nicholson, A.G.; Tsao, M.S.; Beasley, M.B.; Borczuk, A.C.; Brambilla, E.; Cooper, W.A.; Dacic, S.; Jain, D.; Kerr, K.M.; Lantuejoul, S.; et al. The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015. *J. Thorac. Oncol.* **2022**, *17*, 362–387, doi:10.1016/J.JTHO.2021.11.003/ATTACHMENT/4766B675-4B9A-480B-8751-F5E206D6E384/MMC1.DOCX.
 16. Šutić, M.; Vukić, A.; Baranašić, J.; Försti, A.; Džubur, F.; Samaržija, M.; Jakopović, M.; Brčić, L.; Knežević, J. Diagnostic, Predictive, and Prognostic Biomarkers in Non-Small Cell Lung Cancer (NSCLC) Management. *J. Pers. Med.* **2021**, *11*, 1102, doi:10.3390/JPM11111102.
 17. Zheng, M. Classification and Pathology of Lung Cancer. **2016**, doi:10.1016/j.soc.2016.02.003.
 18. Yanagawa, N.; Shiono, S.; Abiko, M.; Ogata, S.Y.; Sato, T.; Tamura, G. New

IASLC/ATS/ERS Classification and Invasive Tumor Size are Predictive of Disease Recurrence in Stage I Lung Adenocarcinoma. *J. Thorac. Oncol.* **2013**, *8*, 612–618, doi:10.1097/JTO.0B013E318287C3EB.

19. Park, H.J.; Cha, Y.J.; Kim, S.H.; Kim, A.; Kim, E.Y.; Chang, Y.S. Keratinization of Lung Squamous Cell Carcinoma Is Associated with Poor Clinical Outcome. *Tuberc. Respir. Dis. (Seoul)*. **2017**, *80*, 179, doi:10.4046/TRD.2017.80.2.179.
20. Detterbeck, F.C.; Boffa, D.J.; Kim, A.W.; Tanoue, L.T. The Eighth Edition Lung Cancer Stage Classification. *Chest* 2017, *151*, 193–203.
21. Kadara, H.; Scheet, P.; Wistuba, I.I.; Spira, A.E. Early Events in the Molecular Pathogenesis of Lung Cancer. *Cancer Prev. Res.* **2016**, *9*, 518–527, doi:10.1158/1940-6207.CAPR-15-0400.
22. Chalela, R.; Curull, V.; Enríquez, C.; Pijuan, L.; Bellosillo, B.; Gea, J. Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy. *J. Thorac. Dis.* **2017**, *9*, 2142–2158, doi:10.21037/JTD.2017.06.20.
23. Newcomb, P.A.; Carbone, P.P. The health consequences of smoking: Cancer. *Med. Clin. North Am.* **1992**, *76*, 305–331, doi:10.1016/S0025-7125(16)30355-8.
24. Akopyan G, B.B. Understanding tobacco smoke carcinogen NNk and lung tumorigenesis. *Int J Oncol.* **2006**, *29*, 745–752.
25. Bade B. C, D.C.C.. Lung Cancer 2020. *Clin. Chest Med.* **2020**, *41*, 1–24, doi:doi:10.1016/j.ccm.2019.10.001.
26. Dela Cruz, C.S.; Tanoue, L.T.; Matthay, R.A. Lung cancer: epidemiology, etiology, and prevention. *Clin. Chest Med.* **2011**, *32*, 605–644, doi:10.1016/J.CCM.2011.09.001.
27. Wang, Y.; McKay, J.D.; Rafnar, T.; Wang, Z.; Timofeeva, M.N.; Broderick, P.; Zong, X.; Laplana, M.; Wei, Y.; Han, Y.; et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat. Genet.* **2014**, *46*, 736, doi:10.1038/NG.3002.
28. McKay, J.D.; Hung, R.J.; Han, Y.; Zong, X.; Carreras-Torres, R.; Christiani, D.C.; Caporaso, N.E.; Johansson, M.; Xiao, X.; Li, Y.; et al. Large-scale association analysis

- identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat. Genet.* 2017 497 **2017**, 49, 1126–1132, doi:10.1038/ng.3892.
29. Velmurugan, K.R.; Varghese, R.T.; Fonville, N.C.; Garner, H.R. High-depth, high-accuracy microsatellite genotyping enables precision lung cancer risk classification. *Oncogene* 2017 3646 **2017**, 36, 6383–6390, doi:10.1038/onc.2017.256.
 30. Brenner, D.R.; McLaughlin, J.R.; Hung, R.J. Previous Lung Diseases and Lung Cancer Risk: A Systematic Review and Meta-Analysis. *PLoS One* **2011**, 6, e17479, doi:10.1371/JOURNAL.PONE.0017479.
 31. Sekine, Y.; Katsura, H.; Koh, E.; Hiroshima, K.; Fujisawa, T. Early detection of COPD is important for lung cancer surveillance. *Eur. Respir. J.* **2012**, 39, 1230–1240, doi:10.1183/09031936.00126011.
 32. Yoo, H.; Jeong, B.H.; Chung, M.J.; Lee, K.S.; Kwon, O.J.; Chung, M.P. Risk factors and clinical characteristics of lung cancer in idiopathic pulmonary fibrosis: A retrospective cohort study. *BMC Pulm. Med.* **2019**, 19, 1–8, doi:10.1186/S12890-019-0905-8/TABLES/4.
 33. Callister, M.E.J.; Baldwin, D.R.; Akram, A.R.; Barnard, S.; Cane, P.; Draffan, J.; Franks, K.; Gleeson, F.; Graham, R.; Malhotra, P.; et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* **2015**, 70 Suppl 2, ii1–ii54, doi:10.1136/THORAXJNL-2015-207168.
 34. Bernardi, F.D.C.; Bernardi, M.D.C.; Takagaki, T.; Siqueira, S.A.C.; Dolhnikoff, M. Lung cancer biopsy: Can diagnosis be changed after immunohistochemistry when the H&E-Based morphology corresponds to a specific tumor subtype? *Clinics* **2018**, 73.
 35. Gregg, J.P.; Li, T.; Yoneda, K.Y. Molecular testing strategies in non-small cell lung cancer: optimizing the diagnostic journey. *Transl. Lung Cancer Res.* **2019**, 8, 286, doi:10.21037/TLCR.2019.04.14.
 36. Duma, N.; Santana-Davila, R.; Molina, J.R. Non–Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. *Mayo Clin. Proc.* **2019**, 94, 1623–1640,

doi:10.1016/j.mayocp.2019.01.013.

37. Spiro, S.C.; Silvestri, G.A. One Hundred Years of Lung Cancer. <https://doi.org/10.1164/rccm.200504-531OE> **2012**, *172*, 523–529, doi:10.1164/RCCM.200504-531OE.
38. Rockswold GL, Ramsey HE, B.G. The results of treatment of lung cancer by surgery, radiation and chemotherapy at a USPHS hospital. *Mil Med.* **1970**, *135*, 362–368.
39. Group, N.C.L.C.C. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *BMJ* **1995**, *311*, 899–909.
40. D’Addario, G.; Pintilie, M.; Leighl, N.B.; Feld, R.; Cerny, T.; Shepherd, F.A. Platinum-based versus non-platinum-based chemotherapy in advanced non-small-cell lung cancer: a meta-analysis of the published literature. *J. Clin. Oncol.* **2005**, *23*, 2926–2936, doi:10.1200/JCO.2005.03.045.
41. Johnson, R.M.G.; Dong, H. Functional expression of programmed death-ligand 1 (B7-H1) by immune cells and tumor cells. *Front. Immunol.* **2017**, *8*, 1–9, doi:10.3389/fimmu.2017.00961.
42. Dong, H.; Strome, S.E.; Salomao, D.R.; Tamura, H.; Hirano, F.; Flies, D.B.; Roche, P.C.; Lu, J.; Zhu, G.; Tamada, K.; et al. Tumor-associated B7-H1 promotes T-cell apoptosis: A potential mechanism of immune evasion. *Nat. Med.* **2002**, *8*, 793–800, doi:10.1038/nm730.
43. Reck, M.; Rodriguez-Abreu, D.; Robinson, A.G.; Hui, R.; Csöszi, T.; Fülöp, A.; Gottfried, M.; Peled, N.; Tafreshi, A.; Cuffe, S.; et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* **2016**, *375*, 1823–1833, doi:10.1056/NEJMoa1606774.
44. Garrido, P.; Conde, E.; de Castro, J.; Gómez-Román, J.J.; Felip, E.; Pijuan, L.; Isla, D.; Sanz, J.; Paz-Ares, L.; López-Ríos, F. Updated guidelines for predictive biomarker testing in advanced non-small-cell lung cancer: a National Consensus of the Spanish Society of Pathology and the Spanish Society of Medical Oncology. *Clin. Transl. Oncol.* **2019**,

doi:10.1007/s12094-019-02218-4.

45. Wu, Y.L.; Planchard, D.; Lu, S.; Sun, H.; Yamamoto, N.; Kim, D.W.; Tan, D.S.W.; Yang, J.C.H.; Azrif, M.; Mitsudomi, T.; et al. Pan-asian adapted clinical practice guidelines for the management of patients with metastatic non-small-cell lung cancer: A csc-esmo initiative endorsed by jsmo, ksmo, mos, sso and tos. *Ann. Oncol.* 2019, 30, 171–210.
46. Lindeman, N.I.; Cagle, P.T.; Aisner, D.L.; Arcila, M.E.; Mary, ; Beasley, B.; Bernicker, E.H.; Colasacco, C.; Dacic, S.; Hirsch, F.R.; et al. Updated Molecular Testing Guideline for the Selection of Lung Cancer Patients for Treatment With Targeted Tyrosine Kinase Inhibitors Guideline From the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *J. Thorac. Oncol. J. Mol. Diagnostics. Arch Pathol Lab Med* **2018**, 142, 321–346, doi:10.5858/arpa.2017-0388-CP.
47. Kohno, T.; Nakaoku, T.; Tsuta, K.; Tsuchihara, K.; Matsumoto, S.; Yoh, K.; Goto, K. Beyond ALK-RET, ROS1 and other oncogene fusions in lung cancer. *Transl. Lung Cancer Res.* **2015**, 4, 156–164, doi:10.3978/j.issn.2218-6751.2014.11.11.
48. Schoenhals, J.E.; Seyedin, S.N.; Anderson, C.; Brooks, E.D.; Li, Y.R.; Younes, A.I.; Niknam, S.; Li, A.; Barsoumian, H.B.; Cortez, M.A.; et al. Uncovering the immune tumor microenvironment in non-small cell lung cancer to understand response rates to checkpoint blockade and radiation. *Transl. Lung Cancer Res.* **2017**, 6, 148–158, doi:10.21037/TLCR.2017.03.06.
49. Diamond, G.; Legarda, D.; Ryan, L.K. The innate immune response of the respiratory epithelium. *Immunol. Rev.* **2000**, 173, 27–38, doi:10.1034/J.1600-065X.2000.917304.X.
50. Martin, T.R.; Frevert, C.W. Innate immunity in the lungs. *Proc. Am. Thorac. Soc.* **2005**, 2, 403–411, doi:10.1513/PATS.200508-090JS.
51. Wong, M.T.; Ong, D.E.H.; Lim, F.S.H.; Teng, K.W.W.; McGovern, N.; Narayanan, S.; Ho, W.Q.; Cerny, D.; Tan, H.K.K.; Anicete, R.; et al. A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures. *Immunity* **2016**, 45, 442–456, doi:10.1016/J.IMMUNI.2016.07.007.

52. Houghton, A.M.G. Mechanistic links between COPD and lung cancer. *Nat. Rev. Cancer* **2013**, *13*, 233–245, doi:10.1038/NRC3477.
53. Banat, G.A.; Tretyn, A.; Pullamsetti, S.S.; Wilhelm, J.; Weigert, A.; Olesch, C.; Ebel, K.; Stiewe, T.; Grimminger, F.; Seeger, W.; et al. Immune and Inflammatory Cell Composition of Human Lung Cancer Stroma. *PLoS One* **2015**, *10*, doi:10.1371/JOURNAL.PONE.0139073.
54. Kargl, J.; Busch, S.E.; Yang, G.H.Y.; Kim, K.H.; Hanke, M.L.; Metz, H.E.; Hubbard, J.J.; Lee, S.M.; Madtes, D.K.; McIntosh, M.W.; et al. Neutrophils dominate the immune cell composition in non-small cell lung cancer. *Nat. Commun.* **2017**, *8*, 1–11, doi:10.1038/ncomms14381.
55. Hanahan, D.; Coussens, L.M. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **2012**, *21*, 309–322, doi:10.1016/J.CCR.2012.02.022.
56. Quail, D.F.; Joyce, J.A. Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* **2013**, *19*, 1423–1437, doi:10.1038/NM.3394.
57. Altorki, N.K.; Markowitz, G.J.; Gao, D.; Port, J.L.; Saxena, A.; Stiles, B.; McGraw, T.; Mittal, V. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nat. Rev. Cancer* **2018**, *19*, 9–31, doi:10.1038/s41568-018-0081-9.
58. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **2011**, *144*, 646–674, doi:10.1016/J.CELL.2011.02.013.
59. Weichand, B.; Popp, R.; Dziuambla, S.; Mora, J.; Strack, E.; Elwakeel, E.; Frank, A.C.; Scholich, K.; Pierre, S.; Syed, S.N.; et al. S1PR1 on tumor-associated macrophages promotes lymphangiogenesis and metastasis via NLRP3/IL-1 β . *J. Exp. Med.* **2017**, *214*, 2695–2713, doi:10.1084/JEM.20160392.
60. Caetano, M.S.; Zhang, H.; Cumpian, A.M.; Gong, L.; Unver, N.; Ostrin, E.J.; Daliri, S.; Chang, S.H.; Ochoa, C.E.; Hanash, S.; et al. IL6 Blockade Reprograms the Lung Tumor Microenvironment to Limit the Development and Progression of K-ras-Mutant Lung Cancer. *Cancer Res.* **2016**, *76*, 3189–3199, doi:10.1158/0008-5472.CAN-15-2840.

61. Lei, X.; Lei, Y.; Li, J.K.; Du, W.X.; Li, R.G.; Yang, J.; Li, J.; Li, F.; Tan, H.B. Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett.* **2020**, *470*, 126–133, doi:10.1016/J.CANLET.2019.11.009.
62. Kastenmüller, W.; Kastenmüller, K.; Kurts, C.; Seder, R.A. Dendritic cell-targeted vaccines--hope or hype? *Nat. Rev. Immunol.* **2014**, *14*, 705–711, doi:10.1038/NRI3727.
63. Schneider, T.; Hoffmann, H.; Dienemann, H.; Schnabel, P.A.; Enk, A.H.; Ring, S.; Mahnke, K. Non-small cell lung cancer induces an immunosuppressive phenotype of dendritic cells in tumor microenvironment by upregulating B7-H3. *J. Thorac. Oncol.* **2011**, *6*, 1162–1168, doi:10.1097/JTO.0B013E31821C421D.
64. Dumitriu, I.E.; Dunbar, D.R.; Howie, S.E.; Sethi, T.; Gregory, C.D. Human dendritic cells produce TGF-beta 1 under the influence of lung carcinoma cells and prime the differentiation of CD4+CD25+Foxp3+ regulatory T cells. *J. Immunol.* **2009**, *182*, 2795–2807, doi:10.4049/JIMMUNOL.0712671.
65. Lee, J.M.; Lee, M.H.; Garon, E.; Goldman, J.W.; Salehi-Rad, R.; Baratelli, F.E.; Schae, D.; Wang, G.; Rosen, F.; Yanagawa, J.; et al. Phase I trial of intratumoral injection of CCL21 gene–modified dendritic cells in lung cancer elicits tumor-specific immune responses and CD8+ T-cell infiltration. *Clin. Cancer Res.* **2017**, *23*, 4556–4568, doi:10.1158/1078-0432.CCR-16-2821.
66. Almatroodi, S.A.; McDonald, C.F.; Darby, I.A.; Pouniotis, D.S. Characterization of M1/M2 Tumour-Associated Macrophages (TAMs) and Th1/Th2 Cytokine Profiles in Patients with NSCLC. *Cancer Microenviron.* **2016**, *9*, 1–11, doi:10.1007/S12307-015-0174-X.
67. Ohri, C.M.; Shikotra, A.; Green, R.H.; Waller, D.A.; Bradding, P. Macrophages within NSCLC tumour islets are predominantly of a cytotoxic M1 phenotype associated with extended survival. *Eur. Respir. J.* **2009**, *33*, 118–126, doi:10.1183/09031936.00065708.
68. Pei, B.X.; Sun, B.S.; Zhang, Z.F.; Wang, A.L.; Ren, P. Interstitial tumor-associated macrophages combined with tumor-derived colony-stimulating factor-1 and interleukin-6, a novel prognostic biomarker in non-small cell lung cancer. *J. Thorac. Cardiovasc. Surg.* **2014**, *148*, 1208-1216.e2, doi:10.1016/J.JTCVS.2014.05.003.

69. Ge, Z.; Ding, S. The Crosstalk Between Tumor-Associated Macrophages (TAMs) and Tumor Cells and the Corresponding Targeted Therapy. *Front. Oncol.* **2020**, *10*, 2404, doi:10.3389/FONC.2020.590941/BIBTEX.
70. Aloe, C.; Wang, H.; Vlahos, R.; Irving, L.; Steinfort, D.; Bozinovski, S. Emerging and multifaceted role of neutrophils in lung cancer. *Transl. Lung Cancer Res.* **2021**, *10*, 2806–2818, doi:10.21037/TLCR-20-760.
71. Spits, H.; Cupedo, T. Innate Lymphoid Cells: Emerging Insights in Development, Lineage Relationships, and Function. <https://doi.org/10.1146/annurev-immunol-020711-075053> **2012**, *30*, 647–675, doi:10.1146/ANNUREV-IMMUNOL-020711-075053.
72. Carrega, P.; Ferlazzo, G. Natural Killers Are Made Not Born: How to Exploit NK Cells in Lung Malignancies. *Front. Immunol.* **2017**, *8*, doi:10.3389/FIMMU.2017.00277.
73. Dunn, G.P.; Old, L.J.; Schreiber, R.D. The three Es of cancer immunoediting. *Annu. Rev. Immunol.* **2004**, *22*, 329–360, doi:10.1146/ANNUREV-IMMUNOL.22.012703.104803.
74. Ganesan, A.-P.; Johansson, M.; Ruffell, B.; Beltran, A.; Lau, J.; Jablons, D.M.; Coussens, L.M. Tumor-infiltrating regulatory T cells inhibit endogenous cytotoxic T cell responses to lung adenocarcinoma. *J. Immunol.* **2013**, *191*, 2009–2017, doi:10.4049/JIMMUNOL.1301317.
75. Lafont, V.; Sanchez, F.; Laprevotte, E.; Michaud, H.A.; Gros, L.; Eliaou, J.F.; Bonnefoy, N. Plasticity of $\gamma\delta$ T Cells: Impact on the Anti-Tumor Response. *Front. Immunol.* **2014**, *5*, doi:10.3389/FIMMU.2014.00622.
76. Wakita, D.; Sumida, K.; Iwakura, Y.; Nishikawa, H.; Ohkuri, T.; Chamoto, K.; Kitamura, H.; Nishimura, T. Tumor-infiltrating IL-17-producing gammadelta T cells support the progression of tumor by promoting angiogenesis. *Eur. J. Immunol.* **2010**, *40*, 1927–1937, doi:10.1002/EJI.200940157.
77. Sutton, C.E.; Mielke, L.A.; Mills, K.H.G. IL-17-producing $\gamma\delta$ T cells and innate lymphoid cells. *Eur. J. Immunol.* **2012**, *42*, 2221–2231, doi:10.1002/EJI.201242569.
78. Ma, Y.; Aymeric, L.; Locher, C.; Mattarollo, S.R.; Delahaye, N.F.; Pereira, P.; Boucontet, L.; Apetoh, L.; Ghiringhelli, F.; Casares, N.; et al. Contribution of IL-17-producing

- gamma delta T cells to the efficacy of anticancer chemotherapy. *J. Exp. Med.* **2011**, *208*, 491–503, doi:10.1084/JEM.20100269.
79. Koper, A.; Zeef, L.A.H.; Joseph, L.; Kerr, K.; Gosney, J.; Lindsay, M.A.; Booton, R. Whole Transcriptome Analysis of Pre-invasive and Invasive Early Squamous Lung Carcinoma in Archival Laser Microdissected Samples. *Respir. Res.* **2017**, *18*, 1–10, doi:10.1186/S12931-016-0496-3/FIGURES/4.
 80. Anusewicz, D.; Orzechowska, M.; Bednarek, A.K. Lung squamous cell carcinoma and lung adenocarcinoma differential gene expression regulation through pathways of Notch, Hedgehog, Wnt, and ErbB signalling. **2020**, *10*, 1–15.
 81. Chen, J.W.; Dhahbi, J. Lung adenocarcinoma and lung squamous cell carcinoma cancer classification, biomarker identification, and gene expression analysis using overlapping feature selection methods. *Sci. Reports 2021 111* **2021**, *11*, 1–15, doi:10.1038/s41598-021-92725-8.
 82. Li, Y.; Gu, J.; Xu, F.; Zhu, Q.; Ge, D.; Lu, C. Transcriptomic and functional network features of lung squamous cell carcinoma through integrative analysis of GEO and TCGA data. *Sci. Rep.* **2018**, *8*, doi:10.1038/S41598-018-34160-W.
 83. Wilkerson, M.D.; Yin, X.; Hoadley, K.A.; Liu, Y.; Hayward, M.C.; Cabanski, C.R.; Muldrew, K.; Miller, R.; Randell, S.H.; Socinski, M.A.; et al. Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important and correspond to different normal cell types. *Clin Cancer Res* **2010**, *16*, 4864–4875, doi:10.1158/1078-0432.CCR-10-0199.
 84. Wu, D.; Pang, Y.; Wilkerson, M.D.; Wang, D.; Hammerman, P.S.; Liu, J.S. Gene-expression data integration to squamous cell lung cancer subtypes reveals drug sensitivity. *Br. J. Cancer* **2013**, *109*, 1599–1608, doi:10.1038/bjc.2013.452.
 85. Wu, K.L.; Tsai, Y.M.; Lien, C.T.; Kuo, P.L.; Hung, J.Y. The roles of microRNA in lung cancer. *Int. J. Mol. Sci.* **2019**, *20*.
 86. Yu, H.; Pang, Z.; Li, G.; Gu, T. Bioinformatics analysis of differentially expressed miRNAs in non-small cell lung cancer. *J. Clin. Lab. Anal.* **2021**, *35*, e23588,

doi:10.1002/JCLA.23588.

87. Gallach, S.; Jantus-Lewintre, E.; Calabuig-Fariñas, S.; Montaner, D.; Alonso, S.; Sirera, R.; Blasco, A.; Usó, M.; Guijarro, R.; Martorell, M.; et al. MicroRNA profiling associated with non-small cell lung cancer: Next generation sequencing detection, experimental validation, and prognostic value. *Oncotarget* **2017**, *8*, 56143–56157, doi:10.18632/oncotarget.18603.
88. Ma, J.; Mannoer, K.; Gao, L.; Tan, A.; Guarnera, M.A.; Zhan, M.; Shetty, A.; Stass, S.A.; Xing, L.; Jiang, F. Characterization of microRNA transcriptome in lung cancer by next-generation deep sequencing. *Mol. Oncol.* **2014**, *8*, 1208–1219, doi:10.1016/j.molonc.2014.03.019.
89. Gilad, S.; Lithwick-Yanai, G.; Barshack, I.; Benjamin, S.; Krivitsky, I.; Edmonston, T.B.; Bibbo, M.; Thurm, C.; Horowitz, L.; Huang, Y.; et al. Classification of the four main types of lung cancer using a microRNA-based diagnostic assay. *J. Mol. Diagnostics* **2012**, *14*, 510–517, doi:10.1016/j.jmoldx.2012.03.004.
90. Gao, X.; Wang, Y.; Zhao, H.; Wei, F.; Zhang, X.; Su, Y.; Wang, C.; Li, H.; Ren, X. *Plasma miR-324-3p and miR-1285 as diagnostic and prognostic biomarkers for early stage lung squamous cell carcinoma*; Vol. 7;.
91. Stewart, P.A.; Welsh, E.A.; Slebos, R.J.C.; Fang, B.; Izumi, V.; Chambers, M.; Zhang, G.; Cen, L.; Pettersson, F.; Zhang, Y.; et al. Proteogenomic landscape of squamous cell lung cancer. *Nat. Commun.* **2019**, *10*, 1–17, doi:10.1038/s41467-019-11452-x.
92. Yang, S.; Chen, L.; Chan, D.W.; Li, Q.K.; Zhang, H. Protein signatures of molecular pathways in non-small cell lung carcinoma (NSCLC): Comparison of glycoproteomics and global proteomics. *Clin. Proteomics* **2017**, *14*, 1–15, doi:10.1186/S12014-017-9166-9/FIGURES/7.
93. Shankha Satpathy, Karsten Krug, Pierre M. Jean Beltran, Sara R. Savage, Francesca Petralia, Chandan Kumar-Sinha, Yongchao Dou, Boris Reva, M. Harry Kane, Shayan C. Avanesian, Suhas V. Vasaikar, Azra Krek, Jonathan T. Lei, Eric J. Jaehnig, Tatiana Omelche, Z.H. A proteogenomic portrait of lung squamous cell carcinoma - ScienceDirect. *Cell* **2021**, *184*, 4348–4371.

94. Weeden, C.; Solomon, B.; Asselin-Labat, M.L. FGFR1 inhibition in lung squamous cell carcinoma: questions and controversies. *Cell Death Discov.* 2015 11 **2015**, 1, 1–5, doi:10.1038/cddiscovery.2015.49.
95. Yuan, G.; Flores, N.M.; Hausmann, S.; Lofgren, S.M.; Kharchenko, V.; Angulo-Ibanez, M.; Sengupta, D.; Lu, X.; Czaban, I.; Azhibek, D.; et al. Elevated NSD3 histone methylation activity drives squamous cell lung cancer. *Nat.* 2021 5907846 **2021**, 590, 504–508, doi:10.1038/s41586-020-03170-y.
96. Vivian, J.; Rao, A.A.; Nothhaft, F.A.; Ketchum, C.; Armstrong, J.; Novak, A.; Pfeil, J.; Narkizian, J.; Deran, A.D.; Musselman-Brown, A.; et al. Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* 2017 354 **2017**, 35, 314–316, doi:10.1038/nbt.3772.
97. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, 29, 15, doi:10.1093/BIOINFORMATICS/BTS635.
98. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 2017 144 **2017**, 14, 417–419, doi:10.1038/nmeth.4197.
99. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, 15, doi:10.1186/S13059-014-0550-8.
100. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, 57, 289–300, doi:10.1111/J.2517-6161.1995.TB02031.X.
101. Carbon, S.; Douglass, E.; Dunn, N.; Good, B.; Harris, N.L.; Lewis, S.E.; Mungall, C.J.; Basu, S.; Chisholm, R.L.; Dodson, R.J.; et al. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **2019**, 47, D330, doi:10.1093/NAR/GKY1055.
102. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.

Proc. Natl. Acad. Sci. **2005**, *102*, 15545–15550, doi:10.1073/PNAS.0506580102.

103. Safari-Alighiarloo, N.; Taghizadeh, M.; Rezaei-Tavirani, M.; Goliaei, B.; Peyvandi, A.A. Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol. Hepatol. From Bed to Bench* **2014**, *7*, 17, doi:10.22037/ghfbb.v7i1.511.
104. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **2008**, *9*, 1–13, doi:10.1186/1471-2105-9-559/FIGURES/4.
105. Fridman, W.H.; Pagès, F.; Sautès-Fridman, C.; Galon, J. The immune contexture in human tumours: Impact on clinical outcome. *Nat. Rev. Cancer* **2012**, *12*, 298–306, doi:10.1038/nrc3245.
106. Yoshihara, K.; Shahmoradgoli, M.; Martínez, E.; Vegesna, R.; Kim, H.; Torres-Garcia, W.; Treviño, V.; Shen, H.; Laird, P.W.; Levine, D.A.; et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **2013**, *4*, 1–11, doi:10.1038/ncomms3612.
107. Bindea, G.; Mlecnik, B.; Tosolini, M.; Kirilovsky, A.; Waldner, M.; Obenauf, A.C.; Angell, H.; Fredriksen, T.; Lafontaine, L.; Berger, A.; et al. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* **2013**, *39*, 782–795, doi:10.1016/J.IMMUNI.2013.10.003.
108. Charoentong, P.; Finotello, F.; Angelova, M.; Mayer, C.; Efremova, M.; Rieder, D.; Hackl, H.; Trajanoski, Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* **2017**, *18*, 248–262, doi:10.1016/J.CELREP.2016.12.019.
109. Juurikka, K.; Butler, G.S.; Salo, T.; Nyberg, P.; Åström, P. Molecular Sciences The Role of MMP8 in Cancer: A Systematic Review. **2019**, doi:10.3390/ijms20184506.
110. Meijer, B.; Gearry, R.B.; Day, A.S. The Role of S100A12 as a Systemic Marker of Inflammation. *Int. J. Inflam.* **2012**, *2012*, doi:10.1155/2012/907078.
111. Xu, D.; Lu, W. Defensins: A Double-Edged Sword in Host Immunity. *Front. Immunol.* **2020**, *11*, 764, doi:10.3389/FIMMU.2020.00764/BIBTEX.

112. Zhang Q, Dong Y, Hao S, Tong Y, Luo Q, A.P. The oncogenic role of TRIP13 in regulating proliferation, invasion, and cell cycle checkpoint in NSCLC cells - PubMed. *Int J Clin Exp Pathol.* **2019**, 12, 3357–3366.
113. Wang, P.; Yuan, Y.; Lin, W.; Zhong, H.; Xu, K.; Qi, X. Roles of sphingosine-1-phosphate signaling in cancer. *Cancer Cell Int.* 2019 191 **2019**, 19, 1–12, doi:10.1186/S12935-019-1014-8.
114. Ooi, A.T.; Gower, A.C.; Zhang, K.X.; Vick, J.L.; Hong, L.; Nagao, B.; Wallace, W.D.; Elashoff, D.A.; Walser, T.C.; Dubinett, S.M.; et al. Molecular profiling of premalignant lesions in lung squamous cell carcinomas identifies mechanisms involved in stepwise carcinogenesis. *Cancer Prev. Res.* **2014**, 7, 487–495, doi:10.1158/1940-6207.CAPR-13-0372/36314/AM/MOLECULAR-PROFILING-OF-PREALIGNANT-LESIONS-IN.
115. Karantza, V. Keratins in health and cancer: more than mere epithelial cell markers. *Oncogene* 2011 302 **2010**, 30, 127–138, doi:10.1038/onc.2010.456.
116. Zuo, W.; Zhang, T.; Wu, D.Z.A.; Guan, S.P.; Liew, A.A.; Yamamoto, Y.; Wang, X.; Lim, S.J.; Vincent, M.; Lessard, M.; et al. p63+Krt5+ distal airway stem cells are essential for lung regeneration. *Nat.* 2014 5177536 **2014**, 517, 616–620, doi:10.1038/nature13903.
117. Che, D.; Wang, M.; Sun, J.; Li, B.; Xu, T.; Lu, Y.; Pan, H.; Lu, Z.; Gu, X. KRT6A Promotes Lung Cancer Cell Growth and Invasion Through MYC-Regulated Pentose Phosphate Pathway. *Front. Cell Dev. Biol.* **2021**, 9, 1331, doi:10.3389/FCELL.2021.694071/BIBTEX.
118. Wang, T.; Zhang, L.; Tian, P.U.; Tian, S. Identification of differentially-expressed genes between early-stage adenocarcinoma and squamous cell carcinoma lung cancer using meta-analysis methods. *Oncol. Lett.* **2017**, 13, 3314, doi:10.3892/OL.2017.5838.
119. Li, Y.; Wu, F.; Tan, Q.; Guo, M.; Ma, P.; Wang, X.; Zhang, S.; Xu, J.; Luo, P.; Jin, Y. The multifaceted roles of FOXM1 in pulmonary disease. *Cell Commun. Signal.* 2019 171 **2019**, 17, 1–16, doi:10.1186/S12964-019-0347-1.
120. Zhong, K.; Yang, F.; Han, Q.; Chen, J.; Wang, J. Skp2 expression has different clinicopathological and prognostic implications in lung adenocarcinoma and squamous

- cell carcinoma. *Oncol. Lett.* **2018**, *16*, 2873, doi:10.3892/OL.2018.9000.
121. Vischioni, B.; Oudejans, J.J.; Vos, W.; Rodriguez, J.A.; Giaccone, G. Frequent overexpression of aurora B kinase, a novel drug target, in non-small cell lung carcinoma patients. *Mol. Cancer Ther.* **2006**, *5*, 2905–2913, doi:10.1158/1535-7163.MCT-06-0301.
 122. Liu, Z.; Sun, Q.; Wang, X. PLK1, A Potential Target for Cancer Therapy. *Transl. Oncol.* **2017**, *10*, 22–32, doi:10.1016/J.TRANON.2016.10.003.
 123. Huang, C.L.; Liu, D.; Nakano, J.; Yokomise, H.; Ueno, M.; Kadota, K.; Wada, H. E2F1 Overexpression Correlates with Thymidylate Synthase and Survivin Gene Expressions and Tumor Proliferation in Non–Small-Cell Lung Cancer. *Clin. Cancer Res.* **2007**, *13*, 6938–6946, doi:10.1158/1078-0432.CCR-07-1539.
 124. Chen, L.; Yu, J.H.; Lu, Z.H.; Zhang, W. E2F2 induction in related to cell proliferation and poor prognosis in non-small cell lung carcinoma. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 10545.
 125. Jin, D.H.; Kim, Y.; Lee, B. Bin; Han, J.; Kim, H.K.; Shim, Y.M.; Kim, D.H. Metformin induces cell cycle arrest at the G1 phase through E2F8 suppression in lung cancer cells. *Oncotarget* **2017**, *8*, 101509–101519, doi:10.18632/ONCOTARGET.21552.
 126. Sun, C.C.; Zhou, Q.; Hu, W.; Li, S.J.; Zhang, F.; Chen, Z.L.; Li, G.; Bi, Z.Y.; Bi, Y.Y.; Gong, F.Y.; et al. Transcriptional E2F1/2/5/8 as potential targets and transcriptional E2F3/6/7 as new biomarkers for the prognosis of human lung carcinoma. *Aging (Albany NY)* **2018**, *10*, 973, doi:10.18632/AGING.101441.
 127. Wang, C.; Gu, X.; Zhang, X.; Zhou, M.; Chen, Y. Development and Validation of an E2F-Related Gene Signature to Predict Prognosis of Patients With Lung Squamous Cell Carcinoma. *Front. Oncol.* **2021**, *11*, 4331, doi:10.3389/FONC.2021.756096/BIBTEX.
 128. Yu, L.; Fang, F.; Lu, S.; Li, X.; Yang, Y.; Wang, Z. lncRNA-HIT promotes cell proliferation of non-small cell lung cancer by association with E2F1. *Cancer Gene Ther.* **2017**, *24*, 221–226, doi:10.1038/cgt.2017.10.
 129. Gu, J.; Liu, Y.; Xie, B.; Ye, P.; Huang, J.; Lu, Z. Roles of toll-like receptors: From inflammation to lung cancer progression. *Biomed. Reports* **2018**, *8*, 126, doi:10.3892/BR.2017.1034.

130. Milette, S.; Fiset, P.O.; Walsh, L.A.; Spicer, J.D.; Quail, D.F. The innate immune architecture of lung tumors and its implication in disease progression. *J. Pathol.* **2019**, *247*, 589–605, doi:10.1002/PATH.5241.
131. Carbone, D.P.; Gandara, D.R.; Antonia, S.J.; Zielinski, C.; Paz-Ares, L. Non–Small Cell Lung Cancer: Role of the Immune System and Potential for Immunotherapy. *J. Thorac. Oncol.* **2015**, *10*, 974, doi:10.1097/JTO.0000000000000551.
132. Marks, D.H.; Thomas, R.; Chin, Y.; Shah, R.; Khoo, C.; Benezra, R. Mad2 Overexpression Uncovers a Critical Role for TRIP13 in Mitotic Exit. *Cell Rep.* **2017**, *19*, doi:10.1016/j.celrep.2017.05.021.
133. Kato, T.; Daigo, Y.; Aragaki, M.; Ishikawa, K.; Sato, M.; Kondo, S.; Kaji, M. Overexpression of MAD2 predicts clinical outcome in primary lung cancer patients. *Lung Cancer* **2011**, *74*, 124–131, doi:10.1016/J.LUNGCAN.2011.01.025.
134. LU, R.; ZHOU, Q.; JU, L.; CHEN, L.; WANG, F.; SHAO, J. Upregulation of trip13 promotes the malignant progression of lung cancer via the emt pathway. *Oncol. Rep.* **2021**, *46*, 1–10, doi:10.3892/OR.2021.8123/HTML.
135. Zhang, Q.; Dong, Y.; Hao, S.; Tong, Y.; Luo, Q.; Aerxiding, P. The oncogenic role of TRIP13 in regulating proliferation, invasion, and cell cycle checkpoint in NSCLC cells. *Int. J. Clin. Exp. Pathol.* **2019**, *12*, 3357.
136. Rodriguez, Y.I.; Campos, L.E.; Castro, M.G.; Aladhami, A.; Oskeritzian, C.A.; Alvarez, S.E. Sphingosine-1 Phosphate: A New Modulator of Immune Plasticity in the Tumor Microenvironment. *Front. Oncol.* **2016**, *6*, doi:10.3389/FONC.2016.00218.
137. Aoki, M.; Aoki, H.; Ramanathan, R.; Hait, N.C.; Takabe, K. Sphingosine-1-Phosphate Signaling in Immune Cells and Inflammation: Roles and Therapeutic Potential. *Mediators Inflamm.* **2016**, *2016*, doi:10.1155/2016/8606878.
138. Zhong, L.; Xie, L.; Yang, Z.; Li, L.; Song, S.; Cao, D.; Liu, Y. Prognostic value of S1PR1 and its correlation with immune infiltrates in breast and lung cancers. *BMC Cancer* **2020**, *20*, 1–17, doi:10.1186/S12885-020-07278-2/FIGURES/7.
139. Cartier, A.; Leigh, T.; Liu, C.H.; Hla, T. Endothelial sphingosine 1-phosphate receptors

- promote vascular normalization and antitumor therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 3157–3166, doi:10.1073/PNAS.1906246117/-/DCSUPPLEMENTAL.
140. Gao, M.; Kong, W.; Huang, Z.; Xie, Z. Identification of Key Genes Related to Lung Squamous Cell Carcinoma Using Bioinformatics Analysis. *Int. J. Mol. Sci.* **2020**, *Vol. 21*, Page 2994 **2020**, *21*, 2994, doi:10.3390/IJMS21082994.
 141. Cheung, C.T.; Bendris, N.; Paul, C.; Hamieh, A.; Anouar, Y.; Hahne, M.; Blanchard, J.M.; Lemmers, B. Cyclin A2 modulates EMT via β -catenin and phospholipase C pathways. *Carcinogenesis* **2015**, *36*, 914–924, doi:10.1093/CARCIN/BGV069.
 142. Chen, R.; Zhang, Y.; Zhang, C.; Wu, H.; Yang, S. miR-137 inhibits the proliferation of human non-small cell lung cancer cells by targeting SRC3. *Oncol. Lett.* **2017**, *13*, 3905–3911, doi:10.3892/OL.2017.5904.
 143. Jin, X.; Liu, X.; Li, X.; Guan, Y. Integrated Analysis of DNA Methylation and mRNA Expression Profiles Data to Identify Key Genes in Lung Adenocarcinoma. *Biomed Res. Int.* **2016**, *2016*, doi:10.1155/2016/4369431.
 144. Smith, S.L.; Bowers, N.L.; Betticher, D.C.; Gautschi, O.; Ratschiller, D.; Hoban, P.R.; Booton, R.; Santibáñez-Koref, M.F.; Heighway, J. Overexpression of aurora B kinase (AURKB) in primary non-small cell lung carcinoma is frequent, generally driven from one allele, and correlates with the level of genetic instability. *Br. J. Cancer* **2005**, *93*, 719, doi:10.1038/SJ.BJC.6602779.
 145. Kathera, C.; Zhang, J.; Janardhan, A.; Sun, H.; Ali, W.; Zhou, X.; He, L.; Guo, Z. Interacting partners of FEN1 and its role in the development of anticancer therapeutics. *Oncotarget* **2017**, *8*, 27593–27602, doi:10.18632/ONCOTARGET.15176.
 146. Wang, X.; Xiao, H.; Wu, D.; Zhang, D.; Zhang, Z.

miR-335-5p Regulates Cell Cycle and Metastasis in Lung Adenocarcinoma by Targeting CCNB2

. *Onco. Targets. Ther.* **2020**, *13*, 6255–6263, doi:10.2147/OTT.S245136.
 147. Joller, N.; Lozano, E.; Burkett, P.R.; Patel, B.; Xiao, S.; Zhu, C.; Xia, J.; Tan, T.G.; Sefik, E.; Yajnik, V.; et al. Treg cells expressing the coinhibitory molecule TIGIT selectively inhibit proinflammatory Th1 and Th17 cell responses. *Immunity* **2014**, *40*, 569–581,

doi:10.1016/J.IMMUNI.2014.02.012.

148. Chan, C.W.Y.; Kay, L.S.; Khadaroo, R.G.; Chan, M.W.C.; Lakatoo, S.; Young, K.J.; Zhang, L.; Gorczynski, R.M.; Catral, M.; Rotstein, O.; et al. Soluble fibrinogen-like protein 2/fibroleukin exhibits immunosuppressive properties: suppressing T cell proliferation and inhibiting maturation of bone marrow-derived dendritic cells. *J. Immunol.* **2003**, *170*, 4036–4044, doi:10.4049/JIMMUNOL.170.8.4036.
149. Yan, J.; Kong, L.Y.; Hu, J.; Gabrusiewicz, K.; Dibra, D.; Xia, X.; Heimberger, A.B.; Li, S. FGL2 as a Multimodality Regulator of Tumor-Mediated Immune Suppression and Therapeutic Target in Gliomas. *JNCI J. Natl. Cancer Inst.* **2015**, *107*, 137, doi:10.1093/JNCI/DJV137.
150. Zhu, Y.; Zhang, L.; Zha, H.; Yang, F.; Hu, C.; Chen, L.; Guo, B.; Zhu, B. Stroma-derived Fibrinogen-like Protein 2 Activates Cancer-associated Fibroblasts to Promote Tumor Growth in Lung Cancer. *Int. J. Biol. Sci.* **2017**, *13*, 804, doi:10.7150/IJBS.19398.
151. Yuan, K.; Feng, Y.; Wang, H.; Zhao, L.; Wang, W.; Wang, T.; Feng, Y.; Huang, G.; Xu, A. FGL2 is positively correlated with enhanced antitumor responses mediated by T cells in lung adenocarcinoma. *PeerJ* **2020**, 2020, doi:10.7717/PEERJ.8654/SUPP-3.
152. Tong, Y.; Lear, T.B.; Evankovich, J.; Chen, Y.; Londino, J.D.; Myerburg, M.M.; Zhang, Y.; Popescu, I.D.; McDyer, J.F.; McVerry, B.J.; et al. The RNFT2/IL-3R α axis regulates IL-3 signaling and innate immunity. *JCI Insight* **2020**, *5*, doi:10.1172/JCI.INSIGHT.133652.
153. Zhu, C.Q.; Strumpf, D.; Li, C.Y.; Li, Q.; Liu, N.; Der, S.; Shepherd, F.A.; Tsao, M.S.; Jurisica, I. Prognostic gene expression signature for squamous cell carcinoma of lung. *Clin. Cancer Res.* **2010**, *16*, 5038–5047, doi:10.1158/1078-0432.CCR-10-0612/83439/AM/PROGNOSTIC-GENE-EXPRESSION-SIGNATURE-FOR-SQUAMOUS.
154. L1CAM-positive expression is associated with poorer survival outcomes in resected non-small cell lung cancer patients - PubMed Available online: <https://pubmed.ncbi.nlm.nih.gov/31934096/> (accessed on Mar 8, 2022).

155. Tischler, V.; Pfeifer, M.; Hausladen, S.; Schirmer, U.; Bonde, A.K.; Kristiansen, G.; Sos, M.L.; Weder, W.; Moch, H.; Altevogt, P.; et al. L1CAM protein expression is associated with poor prognosis in non-small cell lung cancer. *Mol. Cancer* **2011**, *10*, 127, doi:10.1186/1476-4598-10-127.
156. Zhan, P.; Lv, X.J.; Ji, Y.N.; Xie, H.; Yu, L.K. Increased lysyl oxidase-like 2 associates with a poor prognosis in non-small cell lung cancer. *Clin. Respir. J.* **2018**, *12*, 712–720, doi:10.1111/CRJ.12584.
157. Wei, R.; Wang, Z.; Zhang, Y.; Wang, B.; Shen, N.; Elii1, L.; Li, X.; Shang, L.; Shang, Y.; Yan, W.; et al. Bioinformatic analysis revealing mitotic spindle assembly regulated NDC80 and MAD2L1 as prognostic biomarkers in non-small cell lung cancer development. *BMC Med. Genomics* **2020**, *13*, 1–14, doi:10.1186/S12920-020-00762-5/TABLES/7.
158. Yu, Y.; Wang, Z.; Zheng, Q.; Li, J. GALNT2/14 overexpression correlate with prognosis and methylation: potential therapeutic targets for lung adenocarcinoma. *Gene* **2021**, *790*, doi:10.1016/J.GENE.2021.145689.
159. Xiao, L.; Wang, W.; Huangfu, Q.; Tao, H.; Zhang, J. PYGB facilitates cell proliferation and invasiveness in non-small cell lung cancer by activating the Wnt- β -catenin signaling pathway. *Biochem. Cell Biol.* **2020**, *98*, 565–574, doi:10.1139/BCB-2019-0445.
160. Huang, Z.; Wang, S.; Zhang, H.-J.; Zhou, Y.L.; Shi, J.-H. SMOX expression predicts the prognosis of non-small cell lung cancer. *Ann. Transl. Med.* **2021**, *9*, 1048–1048, doi:10.21037/ATM-21-998.
161. Takahashi, Y.; Hamada, J.I.; Murakawa, K.; Takada, M.; Tada, M.; Nogami, I.; Hayashi, N.; Nakamori, S.; Monden, M.; Miyamoto, M.; et al. Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. *Exp. Cell Res.* **2004**, *293*, 144–153, doi:10.1016/J.YEXCR.2003.09.024.
162. Plowright, L.; Harrington, K.J.; Pandha, H.S.; Morgan, R. HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). *Br. J. Cancer* **2009**, *100*, 470–475, doi:10.1038/sj.bjc.6604857.

163. Wu, S.; Chen, H.; Zuo, L.; Jiang, H.; Yan, H. Suppression of long noncoding RNA MALAT1 inhibits the development of uveal melanoma via microRNA-608-mediated inhibition of HOXC4. *Am. J. Physiol. - Cell Physiol.* **2020**, *318*, C903–C912, doi:10.1152/AJPCELL.00262.2019/ASSET/IMAGES/LARGE/ZH00032086370006.JPEG .
164. Xiao, F.; Bai, Y.; Chen, Z.; Li, Y.; Luo, L.; Huang, J.; Yang, J.; Liao, H.; Guo, L. Downregulation of HOXA1 gene affects small cell lung cancer cell survival and chemoresistance under the regulation of miR-100. *Eur. J. Cancer* **2014**, *50*, 1541–1554, doi:10.1016/J.EJCA.2014.01.024.
165. Yu, M.; Yu, S.; Zhou, W.; Yi, B.; Liu, Y. HOXC6/8/10/13 predict poor prognosis and associate with immune infiltrations in glioblastoma. *Int. Immunopharmacol.* **2021**, *101*, 108293, doi:10.1016/J.INTIMP.2021.108293.
166. Wang, M.Q.; Yin, Q.Y.; Chen, Y.R.; Zhu, S.L. Diagnostic and prognostic value of HOXC family members in gastric cancer. *Future Oncol.* **2021**, *17*, 4907–4923, doi:10.2217/FON-2021-0291.
167. Zhang, Y.; Wang, H.; Wang, J.; Bao, L.; Wang, L.; Huo, J.; Wang, X. Global analysis of chromosome 1 genes among patients with lung adenocarcinoma, squamous carcinoma, large-cell carcinoma, small-cell carcinoma, or non-cancer. *Cancer Metastasis Rev.* **2015**, *34*, 249–264, doi:10.1007/S10555-015-9558-0.
168. Guan, H.; Guo, Y.; Liu, L.; Ye, R.; Liang, W.; Li, H.; Xiao, H.; Li, Y. INAVA promotes aggressiveness of papillary thyroid cancer by upregulating MMP9 expression. *Cell Biosci.* **2018**, *8*, 26, doi:10.1186/S13578-018-0224-4.
169. Thorsson, V.; Gibbs, D.L.; Brown, S.D.; Wolf, D.; Bortone, D.S.; Ou Yang, T.H.; Porta-Pardo, E.; Gao, G.F.; Plaisier, C.L.; Eddy, J.A.; et al. The Immune Landscape of Cancer. *Immunity* **2018**, *48*, 812, doi:10.1016/J.IMMUNI.2018.03.023.
170. Li, B.; Cui, Y.; Nambiar, D.K.; Sunwoo, J.B.; Li, R. The Immune Subtypes and Landscape of Squamous Cell Carcinoma. *Clin. Cancer Res.* **2019**, *25*, 3528–3537, doi:10.1158/1078-0432.CCR-18-4085.

171. Fabbri, M.; Garzon, R.; Cimmino, A.; Liu, Z.; Zanesi, N.; Callegari, E.; Liu, S.; Alder, H.; Costinean, S.; Fernandez-Cymering, C.; et al. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15805–15810, doi:10.1073/PNAS.0707628104/ASSET/21C70ECD-D19B-4415-988D-A3E8A9A9B6BE/ASSETS/GRAPHIC/ZPQ0390777250006.JPEG.
172. Osada, H.; Takahashi, T. let-7 and miR-17-92: Small-sized major players in lung cancer development. *Cancer Sci.* **2011**, *102*, 9–17, doi:10.1111/J.1349-7006.2010.01707.X.
173. Kumar, M.S.; Erkeland, S.J.; Pester, R.E.; Chen, C.Y.; Ebert, M.S.; Sharp, P.A.; Jacks, T. Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 3903–3908, doi:10.1073/PNAS.0712321105/SUPPL_FILE/12321FIG12A.PDF.
174. Nana-Sinkam, S.P.; Croce, C.M. Non-coding RNAs in cancer initiation and progression and as novel biomarkers. *Mol. Oncol.* **2011**, *5*, 483–491, doi:10.1016/J.MOLONC.2011.10.003.
175. Calin, G.A.; Croce, C.M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* **2006**, *6*, 857–866, doi:10.1038/nrc1997.
176. Takamizawa, J.; Konishi, H.; Yanagisawa, K.; Tomida, S.; Osada, H.; Endoh, H.; Harano, T.; Yatabe, Y.; Nagino, M.; Nimura, Y.; et al. Reduced Expression of the let-7 MicroRNAs in Human Lung Cancers in Association with Shortened Postoperative Survival. *Cancer Res.* **2004**, *64*, 3753–3756, doi:10.1158/0008-5472.CAN-04-0637.
177. Miko, E.; Margitai, Z.; Czimmerer, Z.; Várkonyi, I.; Dezs, B.; Lányi, Á.; Bacsó, Z.; Scholtz, B. miR-126 inhibits proliferation of small cell lung cancer cells by targeting SLC7A5. *FEBS Lett.* **2011**, *585*, 1191–1196, doi:10.1016/J.FEBSLET.2011.03.039.
178. Sun, Y.; Bai, Y.; Zhang, F.; Wang, Y.; Guo, Y.; Guo, L. miR-126 inhibits non-small cell lung cancer cells proliferation by targeting EGFL7. *Biochem. Biophys. Res. Commun.* **2010**, *391*, 1483–1489, doi:10.1016/J.BBRC.2009.12.098.
179. Markou, A.; Tsaroucha, E.G.; Kaklamanis, L.; Fotinou, M.; Georgoulas, V.; Lianidou,

- E.S. Prognostic Value of Mature MicroRNA-21 and MicroRNA-205 Overexpression in Non-Small Cell Lung Cancer by Quantitative Real-Time RT-PCR. *Clin. Chem.* **2008**, *54*, 1696–1704, doi:10.1373/CLINCHEM.2007.101741.
180. Seike, M.; Goto, A.; Okano, T.; Bowman, E.D.; Schetter, A.J.; Horikawa, I.; Mathe, E.A.; Jen, J.; Yang, P.; Sugimura, H.; et al. MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 12085–12090, doi:10.1073/PNAS.0905234106/SUPPL_FILE/0905234106SI.PDF.
 181. Jang, J.S.; Jeon, H.S.; Sun, Z.; Aubry, M.C.; Tang, H.; Park, C.H.; Rakhshan, F.; Schultz, D.A.; Kolbert, C.P.; Lupu, R.; et al. Increased miR-708 expression in NSCLC and its association with poor survival in lung adenocarcinoma from never smokers. *Clin. Cancer Res.* **2012**, *18*, 3658–3667, doi:10.1158/1078-0432.CCR-11-2857.
 182. Molina-Pinelo, S.; Gutiérrez, G.; Pastor, M.D.; Hergueta, M.; Moreno-Bueno, G.; García-Carbonero, R.; Nogal, A.; Suárez, R.; Salinas, A.; Pozo-Rodríguez, F.; et al. MicroRNA-dependent regulation of transcription in non-small cell lung cancer. *PLoS One* **2014**, *9*, doi:10.1371/JOURNAL.PONE.0090524.
 183. Xing, L.; Todd, N.W.; Yu, L.; Fang, H.; Jiang, F. Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers. *Mod. Pathol.* **2010**, *23*, 1157–1164, doi:10.1038/MODPATHOL.2010.111.
 184. Wang, J.; Li, Z.; Ge, Q.; Wu, W.; Zhu, Q.; Luo, J.; Chen, L. Characterization of microRNA transcriptome in tumor, adjacent, and normal tissues of lung squamous cell carcinoma. *J. Thorac. Cardiovasc. Surg.* **2015**, *149*, 1404-1414.e4, doi:10.1016/J.JTCVS.2015.02.012.
 185. Wang, B.; Sun, L.; Li, J.; Jiang, R. miR-577 suppresses cell proliferation and epithelial-mesenchymal transition by regulating the WNT2B mediated Wnt/ β -catenin pathway in non-small cell lung cancer. *Mol. Med. Rep.* **2018**, *18*, 2753–2761, doi:10.3892/MMR.2018.9279.
 186. Men, L.; Nie, D.; Nie, H. microRNA-577 inhibits cell proliferation and invasion in non-small cell lung cancer by directly targeting homeobox A1. *Mol. Med. Rep.* **2019**, *19*, 1875–1882, doi:10.3892/MMR.2019.9804.

187. Lee, H.Y.; Han, S.S.; Rhee, H.; Park, J.H.; Lee, J.S.; Oh, Y.M.; Choi, S.S.; Shin, S.H.; Kim, W.J. Differential expression of microRNAs and their target genes in non-small-cell lung cancer. *Mol. Med. Rep.* **2015**, *11*, 2034–2040, doi:10.3892/MMR.2014.2890/HTML.
188. Luo, Z.; Zheng, Y.; Zhang, W. Pleiotropic functions of miR107 in cancer networks. *Onco. Targets. Ther.* **2018**, *11*, 4113–4124, doi:10.2147/OTT.S151236.
189. Wang, S.; Ma, G.; Zhu, H.; Lv, C.; Chu, H.; Tong, N.; Wu, D.; Qiang, F.; Gong, W.; Zhao, Q.; et al. miR-107 regulates tumor progression by targeting NF1 in gastric cancer. *Sci. Reports* **2016**, *6*, 1–8, doi:10.1038/srep36531.
190. Sharma, P.; Saini, N.; Sharma, R. MIR-107 functions as a tumor suppressor in human esophageal squamous cell carcinoma and targets Cdc42. *Oncol. Rep.* **2017**, *37*, 3116–3127, doi:10.3892/OR.2017.5546/HTML.
191. Cai, P.; Li, J.; Chen, G.; Peng, B.; Yu, L.; Zhao, B.; Yu, Y. MicroRNA-107 may regulate lung cancer cell proliferation and apoptosis by targeting TP53 regulated inhibitor of apoptosis 1. *Oncol. Lett.* **2020**, *19*, 1958, doi:10.3892/OL.2020.11248.
192. Xia, H.; Li, Y.; Lv, X. MicroRNA-107 inhibits tumor growth and metastasis by targeting the BDNF-mediated PI3K/AKT pathway in human non-small lung cancer. *Int. J. Oncol.* **2016**, *49*, 1325–1333, doi:10.3892/IJO.2016.3628/HTML.
193. Mboge, M.Y.; Mahon, B.P.; McKenna, R.; Frost, S.C. Carbonic Anhydrases: Role in pH Control and Cancer. *Metabolites* **2018**, *8*, doi:10.3390/METABO8010019.
194. Zheng, Y.; Xu, B.; Zhao, Y.; Gu, H.; Li, C.; Wang, Y.; Chang, X. CA1 contributes to microcalcification and tumourigenesis in breast cancer. *BMC Cancer* **2015**, *15*, doi:10.1186/S12885-015-1707-X.
195. Wang, D. bin; Lu, X. ke; Zhang, X.; Li, Z. gang; Li, C. xia Carbonic anhydrase 1 is a promising biomarker for early detection of non-small cell lung cancer. *Tumour Biol.* **2016**, *37*, 553–559, doi:10.1007/S13277-015-3834-Z.
196. Goldman, R.D.; Khuon, S.; Chou, Y.H.; Opal, P.; Steinert, P.M. The function of intermediate filaments in cell shape and cytoskeletal integrity. *J. Cell Biol.* **1996**, *134*, 971–983, doi:10.1083/JCB.134.4.971.

197. Mendez, M.G.; Kojima, S.; Goldman, R.D. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. *FASEB J.* **2010**, *24*, 1838–1851, doi:10.1096/FJ.09-151639.
198. Liu, C.Y.; Lin, H.H.; Tang, M.J.; Wang, Y.K. Vimentin contributes to epithelial-mesenchymal transition cancer cell mechanics by mediating cytoskeletal organization and focal adhesion maturation. *Oncotarget* **2015**, *6*, 15966–15983, doi:10.18632/ONCOTARGET.3862.
199. Yamashita, N.; Tokunaga, E.; Kitao, H.; Hisamatsu, Y.; Taketani, K.; Akiyoshi, S.; Okada, S.; Aishima, S.; Morita, M.; Maehara, Y. Vimentin as a poor prognostic factor for triple-negative breast cancer. *J. Cancer Res. Clin. Oncol.* **2013**, *139*, 739–746, doi:10.1007/S00432-013-1376-6.
200. Yin, S.; Chen, F. fang; Yang, G. fang Vimentin immunohistochemical expression as a prognostic factor in gastric cancer: A meta-analysis. *Pathol. Res. Pract.* **2018**, *214*, 1376–1380, doi:10.1016/J.PRP.2018.07.014.
201. Ye, Z.; Zhang, X.; Luo, Y.; Li, S.; Huang, L.; Li, Z.; Li, P.; Chen, G. Prognostic Values of Vimentin Expression and Its Clinicopathological Significance in Non-Small Cell Lung Cancer: A Meta-Analysis of Observational Studies with 4118 Cases. *PLoS One* **2016**, *11*, doi:10.1371/JOURNAL.PONE.0163162.
202. Jang, H.R.; Shin, S.B.; Kim, C.H.; Won, J.Y.; Xu, R.; Kim, D.E.; Yim, H. PLK1/vimentin signaling facilitates immune escape by recruiting Smad2/3 to PD-L1 promoter in metastatic lung adenocarcinoma. *Cell Death Differ.* **2021**, *28*, 2745–2764, doi:10.1038/s41418-021-00781-4.
203. Chen, Y.; Yang, S.; Zhou, H.; Su, D. PRDX2 Promotes the Proliferation and Metastasis of Non-Small Cell Lung Cancer in Vitro and in Vivo. *Biomed Res. Int.* **2020**, *2020*, doi:10.1155/2020/8359860.
204. Jing, X.; Du, L.; Niu, A.; Wang, Y.; Wang, Y.; Wang, C. Silencing of PRDX2 Inhibits the Proliferation and Invasion of Non-Small Cell Lung Cancer Cells. *Biomed Res. Int.* **2020**, *2020*, doi:10.1155/2020/1276328.

205. Du, W.W.; Yang, W.; Yee, A.J. Roles of versican in cancer biology--tumorigenesis, progression and metastasis. *Histol. Histopathol.* **2013**, *28*, 701–713, doi:10.14670/HH-28.701.
206. Li, W.; Han, F.; Fu, M.; Wang, Z. High expression of VCAN is an independent predictor of poor prognosis in gastric cancer. *J. Int. Med. Res.* **2020**, *48*, 1–11, doi:10.1177/0300060519891271.
207. Zhang, Q.; Wu, J.; Chen, X.; Zhao, M.; Zhang, D.; Gao, F. Upregulation of Versican Associated with Tumor Progression, Metastasis, and Poor Prognosis in Bladder Carcinoma. *Biomed Res. Int.* **2021**, *2021*, doi:10.1155/2021/6949864.
208. Yang, S.; Chen, L.; Chan, D.W.; Li, Q.K.; Zhang, H. Protein signatures of molecular pathways in non-small cell lung carcinoma (NSCLC): comparison of glycoproteomics and global proteomics. *Clin. Proteomics* **2017**, *14*, 31, doi:10.1186/S12014-017-9166-9.
209. Skoda, A.M.; Simovic, D.; Karin, V.; Kardum, V.; Vranic, S.; Serman, L. The role of the Hedgehog signaling pathway in cancer: A comprehensive review. *Bosn. J. basic Med. Sci.* **2018**, *18*, 8–20, doi:10.17305/BJBMS.2018.2756.
210. Gabay, M.; Li, Y.; Felsher, D.W. MYC Activation Is a Hallmark of Cancer Initiation and Maintenance. *Cold Spring Harb. Perspect. Med.* **2014**, *4*, doi:10.1101/CSHPERSPECT.A014241.
211. Pai, S.G.; Carneiro, B.A.; Mota, J.M.; Costa, R.; Leite, C.A.; Barroso-Sousa, R.; Kaplan, J.B.; Chae, Y.K.; Giles, F.J. Wnt/beta-catenin pathway: Modulating anticancer immune response. *J. Hematol. Oncol.* **2017**, *10*, 1–12, doi:10.1186/S13045-017-0471-6/FIGURES/3.

8. POPIS KRATICA

8. POPIS KRATICA

µg- mikro gram

µm- mikro metara

A- adenzin

A260/230- omjer apsorbancije 260 i 230 nanometara

A260/280- omjer apsorbancije 260 i 280 nanometara

ADK- adenokarcinom

ANOVA-analiza varijanci

BCA-bicinkonična kiselina

BR-engl. *broad range*

BSA-albumin iz goveđeg seruma

C-citozin

cDNA- komplementarna deoksiribonukleinska kiselina (engl. *complementary DNA*)

CTLA-4-engl. *Cytotoxic T-Lymphocyte-Associated protein 4*

DEG-diferencijalno eksprimirani geni (engl. *differentially expressed genes*)

DESeq2-engl. *Differential analysis of count data*

DKFZ-Njemački centar za istraživanje raka

DNA- deoksiribonukleinska kiselina

dTTP-deoksitimidin-trifosfat

ES- engl. *Enrichment Score*

ESTIMATE-engl. *Estimation of Stromal and Immune cells in Malignant Tumors using Expression data*

FDR-stopa lažnog otkrivanja (engl. *False Discovery Rate*)

FFPE-engl. *Formalin-fixed, paraffin-embedded*

FISH-engl. *fluorescence in situ hybridization*

G-gvanin

GLM-generalizirani linearni model

GO-genska ontologija

G-referentna sekvenca genoma

GS- značajnost gena (engl. *Gene significance*)

GSEA- engl. *Gene Set Enrichment Analysis*

GSVA- engl. *Gene Set Variation Analysis*

GTEX-engl. *Genotype–Tissue Expression*

GWAS- engl. *Genome-wide association study*

IFN γ -interferon gamma

IPF-intersticijska plućna fibroza

IS-engl. *Immune Score*

KBC- Klinički-bolnički centar

K-matrica

KOPB- kronična opstruktivna bolest pluća

L-lista gena

LDCT-niskodozna računalna tomografija

LFQ -engl. *label-free quantification*

log2FC-logaritamska promjena ekspresije gena

LUSC-tumor pluća skvamoznih stanica (engl. *Lung squamous cell carcinoma*)

MCODE- engl. *Molecular Complex Detection*

MHC-engl. *Major histocompatibility complex*

miRNA-mikro ribonukleinska kiselina

MM- pripadnost modulu (engl. *Module membership*)

MML- maksimalna duljina sekvence koja se može poravnati na referentni genom

MMP-engl. *Maximal Mappable Prefix*

mqH20-ultračista voda, tip 1 (MERCK MILLIPORE, Synergy®)

mRNA- engl. *messenger RNA*

MSigDB-engl. *Molecular Signature Database*

mtRNA-mitohondrijska RNA

NaCl- natrijev klorid

NaN3- natrijev azid

ncRNA-nekodirajuće RNA

NES-engl. *Normalized Enrichment Score*

NET- neuroendokrini tumor pluća

ng-nanogram

NGS- engl. *next generation sequencing*

NK-prirodnoubilačke stanice

NSCLC- tumor pluća ne-malih stanica (engl. *non-small cell lung cancer*)

OS- sveukupno preživljenje pacijenata

PCA- engl. *Principal Component Analysis*

PCR- lančana reakcija polimerazom (engl. *Polymerase chain reaction*)

PD-1-engl. *Programmed cell Death protein*

PD-L1-engl. *Programmed Death-Ligand 1*

PET-CT-pozitronska emisijska tomografija

PET-računalna tomografija

PFS-preživljenje pacijenata bez progresije bolesti

pg/ μ L- poikogram/mikrolitar

PPI- engl. *Protein-protein interaction network*

ppm- engl. *parts per million*

Q10-30- Phred vrijednost kvalitete sekvenciranja 10-30

RIN-engl. *RNA Integrity Number*

rlog-engl. *Regularized logarithm transformation*

RNA- ribonukleinska kiselina

RNA-Seq- sekvenciranje RNA

R-očitana sekvenca

rpm-engl. *revolutions per minute*

rRNA- ribosomalna RNA

RT-PCR- engl. *reverse transcription-PCR*

SAD- Sjedinjene Američke Države

SCLC- tumor pluća malih stanica

SS-engl. *Stromal score*

S-setovi gena

ssGSEA-analiza obogaćivanja setovima gena na razini uzorka (engl. *single-sample Gene Set Enrichment Anylsis*)

STAR-engl. *Spliced Transcripts Alignment to a Reference*

STRING-engl. *Search Tool for the Retrieval of Interaction Genes*

SZO- Svjetska zdravstvena organizacija

TAM- tumoru pridruženi makrofazi

TARGET-engl. *Therapeutically Applicable Research to Generate Effective Treatments*

TCGA-engl. *The Cancer Genome Atlas*

TGF- β -transformirajući faktor rasta beta

TME- tumorski mikrookoliš

TNM-engl. *tumor, nodes and metastases*

TOM-matricu topološke mjere preklapanja (engl. *Topological Overlap Matrix*)

TP- tumor pluća

TPMS-tumor pluća malih stanica

TPNS- tumor pluća ne-malih stanica

TPSS-tumor pluća skvamoznih stanica

Treg- T regulatorne stanice

T-timidin

UV-ultraljubičasto zračenje

WGCNA- engl. *Weighted gene co-expression network analysis*

$\gamma\delta$ T stanice- gamma delta T stanice

9. POPIS ILUSTRACIJA

9. Popis ilustracija

9.1. Slike

Slika 1. 10 najčešćih uzročnika smrti na svjetskoj razini u razdoblju 2000.-2019. godine

Slika 2. Globalna stopa incidencije 10 najčešćih vrsta tumora u 2020.

Slika 3. Stopa smrtnosti od 10 najčešćih vrsta tumora u 2020.

Slika 4. Podjela tumora pluća prema klasifikaciji Svjetske zdravstvene organizacije iz 2021.

Slika 5. Histološke karakteristike tumora pluća skvamoznih stanica.

Slika 6. Shema plana istraživanja

Slika 7. Shema protokola koji je korišten u sklopu UCSC Toil RNA-seq recompute compendium projekta.

Slika 8. Shema protokola za pripremu transkriptomске knjižnice.

Slika 9. Shema protokola za pripremu miRNA knjižnica

Slika 10. Primjer rezultata modula Sažetak kvalitete sekvence po bazama

Slika 11. Primjer rezultata u modulu Kvaliteta očitanih sekvenc po bazama

Slika 12. Primjer rezultata za modul Sažetak vrijednosti kvalitete po očitanoj sekvenci.

Slika 13. Primjer rezultata za modul Provjera očitanih sekvenci po bazama

Slika 14. Primjer rezultata za modul Udio GC baza u očitanim sekvencama

Slika 15. Primjer rezultata za modul Udio N očitavanja u sekvenci.

Slika 16. Primjer rezultata za modul Duplicirane sekvence

Slika 17. Toplinska mapa ekspresije gena u uzorcima.

Slika 18. Rezultati diferencijalne ekspresije gena prikazani Volcano plotom

Slika 19. Analiza genske ontologije za gene sa pozitivno promijenjenom ekspresijom u kategoriji Biološki procesi.

Slika 20. Analiza genske ontologije za gene sa pozitivno promijenjenom ekspresijom u kategoriji Molekularne funkcije

Slika 21. Analiza genske ontologije za gene sa pozitivno promijenjenom ekspresijom u kategoriji Stanične komponente

Slika 22. Rezultati analize genske ekspresije za gene sa negativnom promijenom razine ekspresije u tumoru u odnosu na zdrave kontrole za kategoriju Biološki procesi.

Slika 23. Rezultati analize genske ekspresije za gene sa negativnom promijenom razine ekspresije u tumoru u odnosu na zdrave kontrole za kategoriju Molekularne funkcije.

Slika 24. Rezultati analize genske ekspresije za gene sa negativnom promijenom razine ekspresije u tumoru u odnosu na zdrave kontrole za kategoriju Stanične komponente.

Slika 25. Prikaz 20 najznačajnijih staničnih puteva s povišenom i 20 najznačajnijih puteva s sniženom regulacijom u tumorskim uzorcima.

Slika 26. Mreža predviđene interakcije proteina na temelju razine genske ekspresije

Slika 27. Prikaz gena iz Klastera 1 identificiranih pomoću MCODE aplikacije u Cytoscape programu

Slika 28. Prikaz gena iz Klastera 2 identificiranih pomoću MCODE aplikacije u Cytoscape programu.

Slika 29. Preklop top 50 visoko umreženih gena prema izračunatim vrijednostima 4 algoritma i CytoHubba aplikacije.

Slika 30. Identificirani visokoumreženi geni unutar PPI mreže.

Slika 31. Detekcija outlier uzoraka u WGCNA analizi klasteriranjem.

Slika 32. Odabir granične vrijednosti (β) za konstrukciju mreže bez skale

Slika 33. Identifikacija klastera klasteriranjem gena

Slika 34. Identificirani klasteri gena WGCNA analizom

Slika 35. Korelacija identificiranih klastera i kliničkih parametara pacijenata oboljelih od tumora pluća skvamoznih stanica.

Slika 36. Rezultati analize genske ontologije u kategoriji Biološki procesi za sienna3 klaster.

Slika 37. Korelacija vrijednosti pripadnosti sienna3 klasteru (MM) i vrijednosti značajnosti gena (GS)

Slika 38. Vizualizacija interakcije gena u sienna3 klasteru.

Slika 39. Procjena razine infiltriranosti tumora imunosnim i stromalnim stanicama

Slika 40. Procjena razine infiltriranosti tumorskog tkiva sa specifičnim imunosnim stanicama pomoću GSVA metode.

Slika 41. Procjena razine infiltriranosti tumorskog tkiva sa specifičnim imunosnim stanicama prema pripadnosti imunosnim podtipovima.

Slika 42. Procjena razine infiltriranosti tumorskog tkiva sa citotoksičnim imunološkim stanicama (T limfociti, T $\gamma\delta$ i NK stanicama) prema pripadnosti imunosnim podtipovima

Slika 43. Utjecaj razine infiltracije tumora imunosnim stanicama na sveukupno preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica

Slika 44. Vennov dijagram za gene sa pozitivnom (A) i negativnom (B) promjenom u razini ekspresije, u obje ispitivane kohorte pacijenata oboljelih od tumora pluća skvamoznih stanica

Slika 45. Rezultati GO analize za gene sa pozitivnom promjenom u razini ekspresije u obje testirane kohorte

Slika 46. Rezultati GO analize za gene sa negativnom promjenom u razini ekspresije u obje testirane kohorte

Slika 47. Utjecaj ekspresije gena sa povišenom razinom ekspresije u ispitivanim kohortama na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica.

Slika 48. Utjecaj ekspresije gena sa smanjenom razinom ekspresije u ispitivanim kohortama na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica.

Slika 49. Utjecaj ekspresije visokoumreženih gena u eksperimentalnoj kohorti na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica

Slika 50. Utjecaj ekspresije visokoumreženog FGL2 gena u sienna3 klasteru na preživljenje pacijenata oboljelih od tumora pluća skvamoznih stanica

Slika 51. Primjer rezultata u modulu Kvaliteta očitanih sekvenca po bazama

Slika 52. Primjer rezultata za modul Sažetak vrijednosti kvalitete po očitanoj sekvenci.

Slika 53. Primjer rezultata za modul Provjera očitanih sekvenci po bazama

Slika 54. Primjer rezultata distribucija baza kojima su uklonjeni adapteri

Slika 55. Primjer rezultata distribucije duljine očitanih sekvenci prije i nakon uklonjenih sekvenci adaptera

Slika 56. Distribucija malih molekula RNA na primjeru poravnanja sekvenci iz jednog uzorka.

Slika 57. Rezultati diferencijalne ekspresije miRNA, prikazani u obliku volcano grafa.

Slika 58. Prikaz interaktoma miRNA sa pozitivnom i mRNA sa negativnom promijenom ekspresije u uzorcima skvamoznog tumora pluća

Slika 59. Prikaz interaktoma miRNA sa negativnom i mRNA sa pozitivnom promijenom ekspresije u uzorcima skvamoznog tumora pluća

Slika 60. Analiza genske ontologije za ciljne gene miRNA sa pozitivnom promjenom u razini ekspresije koji su promijenjeno eksprimirani u suprotnom smjeru.

Slika 61. Analiza genske ontologije za ciljne gene miRNA sa negativnom promjenom u razini ekspresije koji su promijenjeno eksprimirani u suprotnom smjeru.

Slika 62. Primjer spektara mase i naboja dobivenih analizom na spektrometru masa.

Slika 63. Prikaz ukupnog broja identificiranih proteina (A) i broja identificiranih proteina u uzorcima (B) spektrometrijom masa u tumorskom tkivu.

Slika 64. Prikaz ukupnog broja identificiranih proteina (A) i broja identificiranih proteina u uzorcima (B) spektrometrijom masa u zdravom tkivu.

Slika 65. Distribucija vrijednosti LFQ intenziteta u uzorcima. Plavim kvadratićem označene su distribucije LFQ vrijednosti tumorskih uzoraka, a zelenim kvadratićem LFQ distribucije zdravih kontroli

Slika 66. Korelacija LFQ vrijednosti intenziteta bioloških replika tumora pluća skvamoznih stanica.

Slika 67. Korelacija LFQ vrijednosti intenziteta bioloških replika zdravih kontroli

Slika 68. PCA analiza tumorskih i zdravih uzoraka na temelju LFQ vrijednosti intenziteta

Slika 69. Rezultati diferencijalne ekspresije proteina prikazani Volcano plotom

Slika 70. miRNA koje potencijalno utišavaju translaciju identificiranih prognostičkih gena

9.2. Tablice

Tablica 1. Opis TNM parametara na temelju osmog izdanja TNM klasifikacije iz 2017. godine.

Tablica 2. Određivanje stadija tumora pluća prema TNM parametrima.

Tablica 3. Biološke karakteristike tumora pluća

Tablica 4. Najčešće mutacije tumora pluća za koje je dostupna ciljana terapija.

Tablica 5. Prikaz količine ukupne RNA korištene za pripremu knjižnice TruSeq RNA Exome kitom

Tablica 6. Specifikacija TruSeq RNA Exome esaja korištenog za pripremu mRNA knjižnica

Tablica 7. Definicija osnovnih pojmova korištenih u WGCNA analizi

Tablica 8. Setovi gena korišteni u GSVA analizi

Tablica 9. Sažetak rezultata kvaliteta očitanih sekvenci po bazama za sve uzorke u analizi

Tablica 10. Sažetak rezultata za modul Udio GC baza u očitanim sekvencama za sve uzorke u analizi

Tablica 11. Sažetak rezultata poravnavanja očitanih sekvenci na referentni genom

Tablica 12. Sažetak rezultata poravnavanja očitanih sekvenci na određene regije genoma.

Tablica 13. Detektirane fuzije gena u uzorcima

Tablica 14. Prikaz 20 gena sa najvećom pozitivnom i negativnom promjenom u razini ekspresije u tumorskim uzorcima u odnosu na zdrave kontrole

Tablica 15. Geni sa pozitivno promijenjenom ekspresijom koji se pojavljuju u svim pet najznačajnijim pojmovima genske ontologije u kategoriji Biološki procesi

Tablica 16. Geni sa negativno promijenjenom ekspresijom koji se pojavljuju u svih pet najznačajnijih pojmova genske ontologije u kategoriji Biološki procesi

Tablica 17. Prikaz obogaćenih staničnih puteva za Klaster 1.

Tablica 18. Prikaz obogaćenih staničnih puteva za Klaster 2.

Tablica 19. Prikaz obogaćenih staničnih puteva za Klaster 3-11

Tablica 20. Identificirani visokoumreženi geni unutar PPI mreže

Tablica 21. Geni grupirani u klasteru sienna3

Tablica 22. Identificirani visokoumreženi geni u sienna3 klasteru

Tablica 23. Popis gena čija je promjena u razini ekspresije potvrđena u obje ispitivane kohorte.

Tablica 24. Validacija ekspresije visokoumreženih gena identificiranih u eksperimentalnoj kohorti.

Tablica 25. Validacija ekspresije *seed* gena u obje testirane kohorte

Tablica 26. Prikaz broja očitanih sekvenci i baza prije i nakon uklanjanja sekvenci adaptera

Tablica 27. Sažetak rezultata poravnavanja očitanih sekvenci na referentni genom

Tablica 28. Prikaz 10 najviše promijenjenih miRNA sa pozitivnom (A) i negativnom (B) promijenom u razini ekspresije u tumorskim uzorcima, u odnosu na zdrave kontrole

Tablica 29. Pronađeni *target* geni za has-miR-577 i has-miR-139-3p

Tablica 30. Prikaz 20 proteina sa najvećom pozitivnom promjenom u razini ekspresije u tumorskim uzorcima u odnosu na zdrave kontrole

Tablica 32. Prikaz 20 proteina sa najvećom negativnom promjenom u razini ekspresije u tumorskim uzorcima u odnosu na zdrave kontrole

Tablica 33. 20 proteina sa najvišom promjenom ekspresije, potvrđenih i na transkripcijskoj razini

Tablica 34. 20 proteina sa najviše smanjenom ekspresijom, potvrđenih i na transkripcijskoj razini

Tablica 35. Prekomjerno eksprimirani proteini i miRNA sa sniženom ekspresijom koje potencijalno reguliraju translaciju proteina

10. Životopis

OSOBNİ PODACI

Ime i prezime: Maja Šutić

Adresa: Ulica Francesca Tenchinija 3, 10000 Zagreb (Hrvatska)

Mob: (+385)922231547

E-mail: suticmaja@yahoo.com

Datum rođenja 08/06/1987

Državljanstvo hrvatsko

RADNO ISKUSTVO

-2017.- danas: Suradnik na HRZZ projektu „Genetički i epigenetički biomarkeri urođene imunosti u tumoru pluća i KOPB-u“ pod voditeljstvom doc. dr. sc. Jelene Knežević

-21.12.2016. - danas: Asistent u Laboratoriju za naprednu genomiku i - travanj 2015.–prosinac 2016. volonter u Laboratoriju za naprednu genomiku, Institut Ruđer Bošković

- travanj 2015. -21.12.2016.- volonter u Laboratoriju za naprednu genomiku, Institut Ruđer Bošković

OBRAZOVANJE

- 2016.–2022. Poslijediplomski studij "Medicinska kemija", Sveučilište u Rijeci, Odjel za biotehnologiju, Rijeka

- 2010.–2012. Diplomski studij, Eksperimentalna biologija, modul Fiziologija i imunobiologija, Sveučilište u Zagrebu, PMF, Biološki odsjek, Zagreb

- 2007.–2010. Prediplomski studij, Biologija, Sveučilište u Zagrebu, PMF, Biološki odsjek, Zagreb

OSPOSOBLJAVANJE

13.-16. 09. 2021. Tečaj za osposobljavanje osoba koje rade s pokusnim životinjama i životinjama za proizvodnju bioloških pripravaka (60 sati)

28.11.2019. Radionica "Applications of Next-Generation Sequencing in Oncology and Human Genetics", Ljubljana, Slovenija

22.10.2019. Agilent Genomics Solutions: getting the optimal results from your samples, seminar, Zagreb, Hrvatska

21-28.09.2019. 20th FEBS International Summer School on Immunology- Immune System: Genes, Receptors and Regulations, Hvar, Hrvatska

03.06-28.06.2019. Njemački institut za istraživanje raka, DKFZ, Heidelberg, Njemačka. Priprema uzoraka za proteomske analize. Financirano: Hrvatska zaklada za znanost

16.01.2019. Napredni pristupi molekularne biologije u svrhu stvaranja boljih staničnih modela za budućnost; seminar, Institut Ruđer Bošković, Zagreb, Hrvatska

05.11.-04.12.2018. Njemački institut za istraživanje raka, DKFZ, Heidelberg, Njemačka. Priprema uzoraka za proteomske analize. Financirano: stipendija Instituta Ruđer Bošković za kratkotrajne istraživačke boravke u inozemstvu

15-18.10.2017. EFIS-EJI Ruggero Ceppellini Advance School of Immunology; Tumor immunology: from tissue microenvironment to immunotherapy. Napulj, Italija. Financirano: stipendija EFIS-EJI

10.03.2017. KEKS:Pospremi reference!-Mendeley, radionica, Institut Ruđer Bošković, Zagreb, Hrvatska

KONGRESI I KONFERENCIJE

3.-8. 07. 2021. 41. FEBS kongres: *Molecules of Life: Towards new Horizons*, autor postera *Transcriptomic profiling of squamous cell lung cancer*, virtualni skup

16-19.09.2020. 10. Kongres Hrvatskog torakalnog društva-Toraks 2020., autor postera *Transcriptomic profiling of squamous cell lung cancer*, Zagreb, Hrvatska

10-13.04.2019. 9. Kongres Hrvatskog torakalnog društva- Toraks 2019, pasivno sudjelovanje, Zagreb, Hrvatska

27.04.2018. *International Symposium on Epigenetics in Head and Neck Cancer*, pasivno sudjelovanje, Zagreb, Hrvatska

18-21.04.2018. 8. Kongres Hrvatskog torakalnog društva-Toraks 2018., pasivno sudjelovanje, Zagreb, Hrvatska

08.03.2018. Od istraživanja do komercijalizacije, stručni skup, pasivno sudjelovanje, Zagreb, Hrvatska

20-21-10 2017. *Annual Meeting of the Croatian Immunological Society with EFIS on Tour Symposium*, Zagreb, Hrvatska, autor postera *Hypomethylation of ASC promoter region in tumor tissue of patients with NSCLC*

30.03.2017. Simpozij *New advances in radio and immunotherapy of cancer*, pasivno sudjelovanje, Zagreb, Hrvatska

3-4.11.2016. HDIR-4: *From Bench to Clinic*, Zagreb, Hrvatska, autor postera *Reduced promoter methylation of MYD88 and ASC/TMS1 genes in tumor tissue of patients with lung carcinoma*

STIPENDIJE I NAGRADE

2018. Stipendija Instituta Ruđer Bošković za kratkotrajno istraživanje u inozemstvu

2017. Stipendija EFIS-EJI za sudjelovanje na ljetnoj školi imunologije

2010.-2011. Stipendija Zaklade 'Hrvatska za djecu'

POPIS PUBLIKACIJA

Šutić M, Vukić A, Baranašić J, Försti A, Džubur F, Samaržija M, Jakopović M, Brčić L, Knežević J. Diagnostic, Predictive, and Prognostic Biomarkers in Non-Small Cell Lung Cancer (NSCLC) Management. *J Pers Med*. 2021 Oct 27;11(11):1102. doi: 10.3390/jpm11111102. PMID: 34834454; PMCID: PMC8624402.

Drpa G, Sutić M, Baranasic J, Jakopovic M, Samarzija M, Kukulj S, Knezevic J. Neutrophil-to-lymphocyte ratio can predict outcome in extensive-stage small cell lung cancer. *Radiol Oncol*. 2020 Sep 22;54(4):437-446. doi: 10.2478/raon-2020-0054. PMID: 32960780; PMCID: PMC7585340.

Šutić M, Baranašić J, Bilić LK, Bilić M, Jakovčević A, Brčić L, Seiwerth S, Jakopović M, Samaržija M, Zechner U, Knežević J. CpG islands in MyD88 and ASC/PYCARD/TMS1 promoter regions are differentially methylated in head and neck squamous cell carcinoma and primary lung

squamous cell carcinoma. *Diagn Pathol*. 2021 Feb 26;16(1):17. doi: 10.1186/s13000-021-01078-3. PMID: 33637109; PMCID: PMC7913417.

Ozretić P, da Silva Filho MI, Catalano C, Sokolović I, Vukić-Dugac A, **Šutić M**, Kurtović M, Bubanović G, Popović-Grle S, Skrinjarić-Cincar S, Vugrek O, Jukić I, Rumora L, Bosnar M, Samaržija M, Bals R, Jakopović M, Försti A, Knežević J. Association of NLRP1 Coding Polymorphism with Lung Function and Serum IL-1 β Concentration in Patients Diagnosed with Chronic Obstructive Pulmonary Disease (COPD). *Genes (Basel)*. 2019 Oct 9;10(10):783. doi: 10.3390/genes10100783. PMID: 31601004; PMCID: PMC6826440.

Šutić M, Motzek A, Bubanović G, Linke M, Sabol I, Vugrek O, Ozretić P, Brčić L, Seiwert S, Debeljak Ž, Jakovčević A, Janevski Z, Stančić-Rokotov D, Vukić-Dugac A, Jakopović M, Samaržija M, Zechner U, Knežević J. Promoter methylation status of ASC/TMS1/PYCARD is associated with decreased overall survival and TNM status in patients with early stage non-small cell lung cancer (NSCLC). *Transl Lung Cancer Res*. 2019 Dec;8(6):1000-1015. doi: 10.21037/tlcr.2019.12.08. PMID: 32010578; PMCID: PMC6976376.