

Metabolička analiza utjecaja [2, 3-b] piridinskog spoja na MDA-MB-231 i MCF-7 stanične linije raka dojke

Meić, Luka

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Split, Faculty of Chemistry and Technology / Sveučilište u Splitu, Kemijsko-tehnološki fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:167:102828>

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

Download date / Datum preuzimanja: **2025-04-02**

Repository / Repozitorij:

[Repository of the Faculty of chemistry and technology - University of Split](#)



SVEUČILIŠTE U SPLITU
KEMIJSKO-TEHNOLOŠKI FAKULTET

METABOLOMIČKA ANALIZA UTJECAJA
TIENO[2,3-b]PIRIDINSKOG SPOJA NA MDA-MB-231 I MCF-7
STANIČNE LINIJE RAKA DOJKE

DIPLOMSKI RAD

LUKA MEIĆ

Matični broj: 157

Split, listopad 2022.

SVEUČILIŠTE U SPLITU
KEMIJSKO-TEHNOLOŠKI FAKULTET
DIPLOMSKI STUDIJ KEMIJE
ORGANSKA KEMIJA I BIOKEMIJA

METABOLOMIČKA ANALIZA UTJECAJA
TIENO[2,3-b]PIRIDINSKOG SPOJA NA MDA-MB-231 I MCF-7
STANIČNE LINIJE RAKA DOJKE

DIPLOMSKI RAD

LUKA MEIĆ

Matični broj: 157

Split, listopad 2022.

UNIVERSITY OF SPLIT
FACULTY OF CHEMISTRY AND TECHNOLOGY
GRADUATE STUDY IN CHEMISTRY
ORGANIC CHEMISTRY AND BIOCHEMISTRY

**METABOLOMIC ANALYSIS OF IMPACT OF
THIENO[2,3-b]PYRIDINE COMPOUND ON MDA-MB-231 AND
MCF-7 BREAST CANCER CELLS**

DIPLOMA THESIS

Luka Meić

Parent number: 157

Split, October 2022

TEMELJNA DOKUMENTACIJSKA KARTICA

DIPLOMSKI RAD

Sveučilište u Splitu
Kemijско-tehnološki fakultet u Splitu
Diplomski studij kemije

Znanstveno područje: prirodne znanosti

Znanstveno polje: kemija

Tema rada je prihvaćena na XV. sjednici Fakultetskog vijeća Kemijско-tehnološkog fakulteta u Splitu

Mentor: izv. prof. dr. sc. Mila Radan

METABOLOMIČKA ANALIZA UTJECAJA TIENO[2,3-b]PIRIDINSKOG SPOJA NA MDA-MB-231 I MCF-7 STANIČNE LINIJE RAKA DOJKE

Luka Meić, 157

Sažetak:

Pojam metabolomika se odnosi na sustavnu identifikaciju i kvantifikaciju svih metabolita u određenom organizmu ili biološkom uzorku.

Trenutno se koriste dva komplementarna pristupa za metabolička istraživanja: metaboličko profiliranje i metabolički otisak prsta (eng. *fingerprint*).

Za metabolomičke analize koriste se tehnike za separaciju i tehnike visoke rezolucije za identifikaciju metabolita. Nakon što su analitičkim tehnikama dobiveni sirovi spektri, prije same statističke analize podataka, potrebna je poboljšana obrada neobrađenih podataka. Statistička analiza može biti univarijatna ili multivarijatna.

Cilj ovog rada je istražiti utjecaj spoja (*E*)-3-amino-5-(3-(3-bromfenil)akrilolil)-*N*-(3-kloro-2-metilfenil)-6-metiltieno[2,3-b]piridin-2-karboksamida (u nastavku teksta tieno[2,3-b]piridin) na metabolizam dvije vrste stanica raka dojke MCF-7 i MDA-MB-231.

Za definiranje pikova, poravnanje RT vremena te označavanje i integraciju pikova korišten je softver Agilent MassHunter Qualitative Analysis.

Platforma za analizu metabolomičkih podataka MetaboAnalyst je korištena za analizu različito eksprimiranih metabolita kroz dvije stanične linije.

Rezultati analize pokazuju da tretman stanica tieno[2,3-b]piridinom ima veliki učinak na metabolizam glukoze odnosno energetske cikluse u stanici, a posebno na glikolizu, glukoneogenezu, metabolizam piruvata, pojavu Warburgova efekta i metabolizam inozitola. Tieno[2,3-b]piridin ima veću citotoksičnost za MDA-MB-231 u odnosu na MCF-7 staničnu liniju.

Ključne riječi: metabolomika, GC-MS analiza, MetaboAnalyst, tieno[2,3-b]piridin

Rad sadrži: 38 stranica, 19 slika, 3 tablice, 41 literaturnu referencu

Jezik izvornika: hrvatski

Sastav Povjerenstva za obranu:

1. Izv. prof. dr. sc. Ante Prkić - predsjednik
2. Doc. dr. sc. Marina Tranfić Bakić- član
3. Izv. prof. dr. sc. Mila Radan - mentor

Datum obrane: 26.10.2022.

Rad je u tiskanom i elektroničkom (pdf format) obliku pohranjen u Knjižnici Kemijско-tehnološkog fakulteta Split, Ruđera Boškovića 35.

BASIC DOCUMENTATION CARD

DIPLOMA THESIS

University of Split
Faculty of Chemistry and Technology Split
Graduate study in Chemistry

Scientific area: Natural Sciences

Scientific field: Chemistry

Thesis subject was approved by Faculty Council of Faculty of Chemistry and Technology, session no. XV.

Mentor: Mila Radan, PhD, associate prof.

Comentor:

METABOLOMIC ANALYSIS OF IMPACT OF THIENO[2,3-b]PYRIDINE COMPOUND ON MDA-MB-231 AND MCF-7 BREAST CANCER CELLS

Luka Meić, 157

Abstract:

The term metabolomics refers to the systematic identification and quantification of all metabolites in certain organisms or biological samples.

Currently, two complementary approaches are used in metabolic research: metabolic profiling and metabolic fingerprinting.

Separation and high-resolution techniques for metabolite identification have been used in metabolomic analyses. After the raw spectra have been obtained using analytical techniques, preprocessing of the raw data is required before statistical analysis of the data. Statistical analyses can be univariate or multivariate.

The aim of this study was to investigate the influence of compound (*E*)-3-amino-5-(3-(3-bromophenyl)acryloyl)-*N*-(3-chloro-2-methylphenyl)-6-methylthieno[2,3-*b*]pyridine-2-carboxamide (hereinafter thieno[2,3-*b*]pyridine) on the metabolism of two types of breast cancer cells MCF-7 and MDA-MB-231.

Agilent MassHunter Qualitative Analysis software was used for peak definition, RT time alignment, and peak annotation and integration.

MetaboAnalyst, a platform dedicated to metabolomic data analysis, was used to analyze differentially expressed metabolites across the two cell lines.

The results of the analysis showed that the treatment of cells with thieno[2,3-*b*]pyridine had a great impact on glucose metabolism, that is, energy cycles in the cell, especially on glycolysis, gluconeogenesis, pyruvate metabolism, the occurrence of Warburg effect, and inositol metabolism. Thieno[2,3-*b*]pyridine showed greater cytotoxicity in MDA-MB-231 cells than in MCF-7 cells.

Keywords: metabolomics, GC-MS analysis, MetaboAnalyst, thieno[2,3-*b*]pyridine

Thesis contains: 38 pages, 19 figures, 3 tables, 41 references

Original in Croatian

Defense committee:

- | | |
|--|-------------|
| 1. Ante Prkić – PhD, associate prof. | chairperson |
| 2. Marina Tranfić Bakić - PhD, assistant prof. | member |
| 3. Mila Radan - PhD, associate prof. | supervisor |

Defense date: 26.10.2022.

Printed and electronic (pdf format) version of thesis is deposited in Library of Faculty of Chemistry and Technology Split, Ruđera Boškovića 35.

Završni rad je izrađen u Zavodu za biokemiju, Kemijsko-tehnološkog fakulteta u Splitu pod mentorstvom izv. prof. dr. sc. Mile Radan, u razdoblju od lipnja do listopada 2022. godine.

Veliku zahvalu upućujem izv. prof. dr. sc. Mili Radan na mentorstvu, savjetima i pomoći pri izradi diplomskog rada te joj se moram zahvaliti na uvodu u složen, ali neopisivo intrigantan i dojmljiv svijet biokemije.

Također, moram zahvaliti prijateljima i kolegama što su ovih pet godina studiranja učinili naizgled mnogo kraćim.

Na kraju, najveću zahvalu dugujem svojoj obitelji za cjelokupnu potporu tokom čitavog školovanja.

ZADATAK ZAVRŠNOG RADA

1. Pomoću GC-MS sustava provesti komparativnu metaboličku analizu ekstrakta MDA-MB-231 i MCF-7 stanica raka tretiranih tieno[2,3-b]piridinom.
2. Istražiti računalne programe za predobradu neobrađenih GC-MS podataka i metabolomičku analizu.
3. Interpretirati dobivene rezultate te predložiti promjene metaboličkih puteva kao odgovor stanica raka na primjenu tieno[2,3-b]piridinskog spoja.

SAŽETAK

Pojam metabolomika se odnosi na sustavnu identifikaciju i kvantifikaciju svih metabolita u određenom organizmu ili biološkom uzorku.

Trenutno se koriste dva komplementarna pristupa za metabolička istraživanja: metaboličko profiliranje i metabolički otisak prsta (eng. *fingerprint*).

Za metabolomičke analize koriste se tehnike za separaciju i tehnike visoke rezolucije za identifikaciju metabolita. Nakon što su analitičkim tehnikama dobiveni sirovi spektri, prije same statističke analize podataka, potrebna je predobrada neobrađenih podataka. Statistička analiza može biti univarijatna ili multivarijatna.

Cilj ovog rada je istražiti utjecaj spoja (*E*)-3-amino-5-(3-(3-bromfenil)akriloil)-*N*-(3-kloro-2-metilfenil)-6-metiltieno[2,3-*b*]piridin-2-karboksamida (u nastavku teksta tieno[2,3-*b*]piridin) na metabolizam dvije vrste stanica raka dojke MCF-7 i MDA-MB-231.

Za definiranje pikova, poravnanje RT vremena te označavanje i integraciju pikova korišten je softver Agilent MassHunter Qualitative Analysis.

Platforma za analizu metabolomičkih podataka MetaboAnalyst je korištena za analizu različito eksprimiranih metabolita kroz dvije stanične linije.

Rezultati analize pokazuju da tretman stanica tieno[2,3-*b*]piridinom ima veliki učinak na metabolizam glukoze odnosno energetske cikluse u stanici, a posebno na glikolizu, glukoneogenezu, metabolizam piruvata, pojavu Warburgova efekta i metabolizam inozitola. Tieno[2,3-*b*]piridin ima veću citotoksičnost za MDA-MB-231 u odnosu na MCF-7 staničnu liniju.

Ključne riječi: metabolomika, GC-MS analiza, MetaboAnalyst, tieno[2,3-*b*]piridin

SUMMARY

The term metabolomics refers to the systematic identification and quantification of all metabolites in certain organism or biological sample.

Currently, two complementary approaches are used for metabolic research: metabolic profiling and metabolic fingerprinting.

Separation techniques and high-resolution techniques for metabolite identification are used in metabolomic analyses. After the raw spectra have been obtained by analytical techniques, preprocessing of the raw data is required before the actual statistical analysis of the data. Statistical analysis can be univariate or multivariate.

The aim of this work is to investigate the influence of the compound (*E*)-3-amino-5-(3-(3-bromophenyl)acryloyl)-*N*-(3-chloro-2-methylphenyl)-6-methylthieno[2,3-*b*]pyridine-2-carboxamide (hereinafter thieno[2,3-*b*]pyridine) on the metabolism of two types of breast cancer cells MCF-7 and MDA-MB-231.

Agilent MassHunter Qualitative Analysis software was used for peak definition, RT time alignment, and peak annotation and integration.

MetaboAnalyst, a platform dedicated to metabolomics data analysis, was used to analyze differentially expressed metabolites across two cell lines.

The results of the analysis show that the treatment of cells with thieno[2,3-*b*]pyridine has a great impact on glucose metabolism, i.e. energy cycles in the cell, especially on glycolysis, gluconeogenesis, pyruvate metabolism, occurrence of the Warburg effect and inositol metabolism. Thieno[2,3-*b*]pyridine has greater cytotoxicity on MDA-MB-231 compared with MCF-7 cell line.

Keywords: metabolomics, GC-MS analysis, MetaboAnalyst, thieno[2,3-*b*]pyridine

SADRŽAJ

UVOD	1
1. OPĆI DIO	2
1.1. Omika	2
1.2. Metabolomika	3
1.2.1. Metaboličko profiliranje i metabolički otisak prsta.....	4
1.3. Priprema uzorka u metabolomici	4
1.4. Analitičke tehnike u metabolomičkoj analizi	5
1.4.1. GC-MS tehnika u metabolomičkoj analizi	7
1.5. Predobrada podataka	7
1.5.1. Filtriranje šuma i korekcija bazne linije	7
1.5.2. Spektralna dekonvolucija pikova.....	8
1.5.3. Definiranje, detekcija i poravnanje pikova.....	8
1.5.4. Normalizacija	9
1.6. Statistička analiza u metabolomici	9
1.6.1. Univarijantne metode analize	10
1.6.2. Multivarijantne metode analize	11
1.6.3. Alati za vizualizaciju	12
1.6.4. Alati za strojno učenje	14
1.7. Softverski alati u metabolomici	15
1.7.1. Softverski alati za predobradu podataka.....	15
1.7.2. Baze podataka u metabolomici.....	15
1.7.3. MetaboAnalyst 5.0	16
2. EKSPERIMENTALNI DIO	23
2.1. GC-MS analiza ekstrakta stanica raka	23
2.2. Predobrada GC-MS podataka	24
2.3. Statistička analiza podataka u programu MetaboAnalyst	24
2.4. Analiza obogaćenja u programu MetaboAnalyst	27
3. REZULTATI I RASPRAVA	29
3.1. Identifikacija spojeva	29
3.2. ANOVA	30
3.3. Analiza glavnih komponenti - PCA	30
3.4. Analiza kvantitativnog obogaćivanja	31

4. ZAKLJUČAK.....	34
LITERATURA	35

UVOD

Metabolomika stvara profil malih molekula koje su izvedene iz staničnog metabolizma te izravno odražava ishod složenih mreža biokemijskih reakcija. Na taj način pruža uvid u višestruke aspekte stanične fiziologije. [1]

Za separaciju metabolita najčešće se koriste tekućinska i plinska kromatografija. One se kombiniraju s instrumentima za detekciju poput nuklearne magnetske rezonancije i spektrometrije masa. Međutim, prikupljanje podataka tek je prvi korak. Metabolomika se u konačnici sastoji od integracije instrumenata za analizu, kemije, statistike i računalne znanosti sa saznanjima iz biologije. [1]

Metabolomička analiza provedena je korištenjem plinske kromatografije povezane s spektrometrijom masa i platforme za analizu metabolomičkih podataka MetaboAnalyst.

Cilj ovog diplomskog rada je proučiti djelovanje tieno[2,3-b]piridinskog spoja na metabolizam dviju staničnih linija raka dojke. Također, u radu su prikazane mogućnosti i potencijal ovakve analize kao alata za pronalaženje novih lijekova.

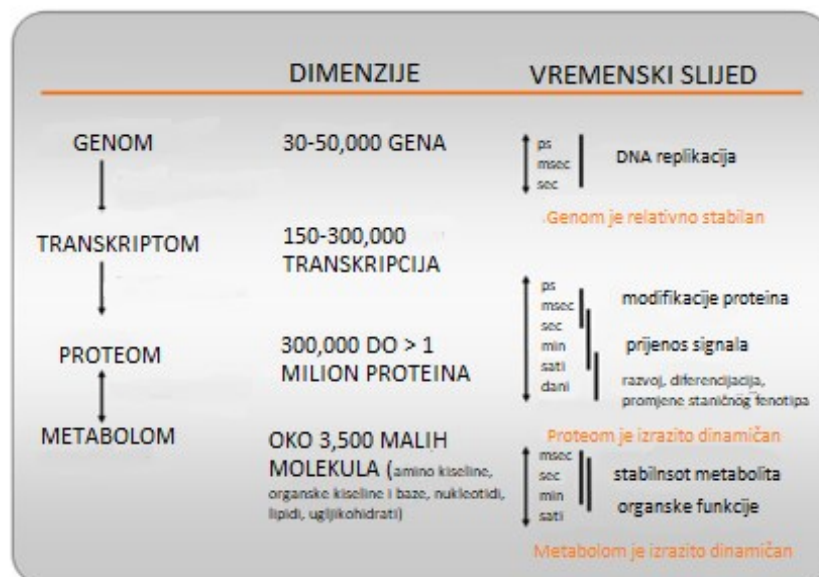
1. OPĆI DIO

1.1. Omika

Pojam "omika" potječe od latinskog sufiksa "ome" što znači masa ili mnogo nečega. Prema tome znanosti čiji naziv završava sufiksom omika uključuju velik broj mjerenja potrebnih za dobivanje krajnjeg rezultata. Ta mjerenja se odnose na kemijske markere koji pokazuju određenu biološku aktivnost. [2]

Pojam "biomarker" označava molekulu koja je karakteristična za određeno biološko stanje u stanici ili organizmu. Početak korištenja kemijskih tvari kao markera za praćenje događaja u organizmu pripisuje se Hipokratu koji je "voćni miris" zadaha opisivao kao razlog ("biomarker") za dijagnosticiranje šećerne bolesti (diabetes). [2]

Cilj "omičkih pristupa" je sveobuhvatno proučavanje bioloških procesa kako bi se prepoznao utjecaj različitih molekula (geni, RNA, proteini i metaboliti), za razliku od zasebnog proučavanja svakog pojedinog procesa. Trenutno se približavamo pokušaju odgovaranja na složena biološka pitanja bez potrebe za pretjeranim poznavanjem cijelog biološkog okvira. Daljim razvojem tehnologije znanstvenici bi ovim pristupom mogli objasniti složena biološka pitanja te dobiti izvrstan alat za pomoć pri dijagnostici bolesti i razvoju lijekova. Ovisno jesu li cilj istraživanja geni, proteini ili metaboliti područje dobiva naziv genomika, proteomika ili metabolomika. Slika 1. pokazuje složenost različitih "omičkih" razina ovisno o kvantitativnoj dimenziji i vremenskoj skali procesa unutar organizma. [2]



Slika 1.: Složenost različitih "omičkih" razina ovisno kvantitativnoj dimenziji i vremenskoj skali procesa unutar organizma [2]

1.2. Metabolomika

Pojam metabolomika se odnosi na sustavnu identifikaciju i kvantifikaciju svih metabolita u određenom organizmu ili biološkom uzorku. Ona istražuje aktivnost i status staničnog i organskog metabolizma na globalnoj ili mrežnoj razini za ocrtavanje krajnjih točaka fiziologije i patofiziologije. To podrazumijeva proučavanje malih molekula, uključujući endogene i egzogene molekule, koji su proizvodi i supstrati kemijskih reakcija unutar bioloških sustava. Ona je krajnja točka "omičke kaskade" i u suštini najbliže opisuje fenotip. Metabolomički eksperiment izravno odražava aktivnost metaboličke mreže koja dovodi do proizvodnje ovih metabolita i daje bitne informacije o biološkom statusu sustava. [1, 3, 4]

Sve veća razlučivost koju pružaju spektroskopija nuklearne magnetske rezonancije (NMR) i spektrometrija mase (MS), zajedno sa snažnim kemometričkim softverima, omogućuju simultano određivanje i usporedbu tisuća kemijskih entiteta, što je dovelo do širenja studije biokemije malih molekula u bakterijama, biljkama i sisavcima. U Tablici 1. prikazani su najvažniji pojmovi vezani uz metabolomiku. [3]

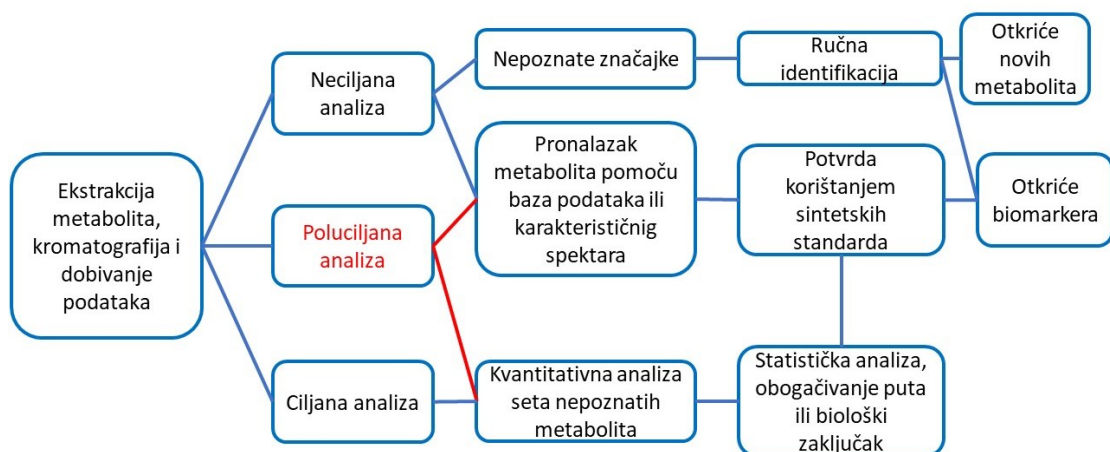
Tablica 1.: Najvažniji pojmovi vezani uz metabolomiku [4]

metaboliti	male molekule koje sudjeluju u općim metaboličkim reakcijama potrebnim za održavanje, rast i normalan rad funkcija stanice
metabolom	kompletan skup metabolita u organizmu
metabolomika	identifikacija i kvantifikacija svih metabolita u biološkom sustavu
metabolomičko profiliranje	kvantitativna analiza skupa metabolita u odabranom biokemijskom putu ili specifičnoj klasi spojeva koja uključuje ciljanu analizu, analizu vrlo ograničenog broja metabolita, npr. pojedinačni analiti kao prekursori ili produkti biokemijske reakcije
metabolički otisak prsta	neciljani sveobuhvatni pristup provjere za klasificiranje uzoraka baziranih na obrascima metabolita ili "otiscima prstiju" koji se mijenjaju kao odgovor na bolesti, okolišne ili genetske poremećaje s krajnjim ciljem identifikacije diskriminirajućih metabolita
metabolomički otisak stopala	analiza otisaka izvanstaničnih metabolita u staničnoj kulturi medija kao odraz izlučivanja ili unosa metabolita od strane stanica

1.2.1. Metaboličko profiliranje i metabolički otisak prsta

Metabolomika nije definirana određenim eksperimentom. Trenutno se koriste dva komplementarna pristupa za metabolička istraživanja: metaboličko profiliranje i metabolički otisak prsta (eng. *fingerprint*). Metaboličko profiliranje se još naziva i ciljanom metabolomikom (eng. *targeted metabolomics*), a metabolički otisak prsta neciljanom metabolomikom (eng. *untargeted metabolomics*). [5]

U ciljanom pristupu razvijaju se kvantitativne analitičke metode za određivanje metabolita u određenom metaboličkom putu ili klasi uzoraka. Određuje se apsolutna koncentracija metabolita uz testiranje poznate hipoteze što zahtjeva određeno dublje razumijevanje metabolizma. Ovakav pristup donosi nepristranu informaciju koja može dati bolji uvid u biokemijske puteve i fiziološke interakcije. Za razliku do ciljanog pristupa, neciljanom pristupu nije svrha identificirati svaku sastavnicu (metabolit) zasebno. Ovim pristupom se uspoređuju relativne razlike u sveukupnom uzorku odnosno otisku metabolita u stanici ili organizmu koji se mijenjaju kao odgovor na razne bolesti ili izloženost toksinima. Kombinacija ova dva pristupa naziva se polu ciljani pristup u kojem su poznate koncentracije tvari, ali hipoteza nije postavljena. Slika 2. sažeto prikazuje ciljanu, poluciljanu i neciljanu metabolomičku analizu. [1, 5]



Slika 2.: Ciljana, poluciljana i neciljana statistička analiza [1]

1.3. Priprema uzorka u metabolomici

Metabolomička istraživanja se mogu provoditi na različitim biološkim materijalima koji mogu biti ljudskog, životinjskog ili biljnog porijekla. Najčešće se koriste krv, urin, tkiva, slina ili stanice. Priprema uzorka se razlikuje ovisno o metodi i vrsti metabolita koji se

analiziraju, a prvi je i osnovni korak u metabolomičkoj analizi. Ona je ključna za kvalitetu i pouzdanost dobivenih rezultata. Sama priprema uzorka za analizu općenito podrazumijeva tri glavna koraka: prikupljanje uzoraka, ekstrakciju metabolita i pripremu ekstrakta za analitički uređaj. [6]

Prikupljanje i skladištenje uzoraka su vrlo važni koraci pripreme uzorka jer kod njih najčešće dolazi do pogrešaka koje utječu na cijelo istraživanje. Uzorci se moraju prikupljati na jednak način i skladištiti pod istim uvjetima. Temperatura znatno utječe na stabilnost metabolita pa je skladištenje vrlo važno. Skladištenje se mora provoditi na način da se minimalizira mogućnost mijenjane koncentracije metabolita zbog utjecaja temperature i drugih čimbenika. [6]

Ovisno o analitičkoj metodi i kemijskom karakteru metabolita odabire se pogodno otapalo za ekstrakciju. Važno je održavati procese što je moguće jednostavnijima i bržima za obradu velikog broja uzoraka, a u isto vrijeme spriječiti bilo kakvu mogućnost degradacije tijekom ekstrakcije. Otapala se biraju prema njihovoj polarnosti, selektivnosti i toksičnosti. Polarnost otapala igra ključnu ulogu u ekstrakciji jer vrsta metabolita koji se ekstrahiraju zavisi od polarnosti. Najčešća korištena otapala za ekstrakciju su metanol, aceton, acetonitril ili vodene smjese ovih otapala. Metanol je poželjno otapalo jer može otapati polarne i nepolarne spojeve. [6]

Za GC-MS analizu svi uzorci moraju biti hlapljivi. Ako uzorci nisu hlapljivi potrebna je derivatizacija. Osim što se njome povećava hlapljivost metabolita koji sadrže polarne funkcionalne skupine kao što su karboksilne kiseline, ketoni i alkohol, derivatizacijom se sprječava degradacija molekula na visokim temperaturama. Nažalost, ovo je još jedan izvor varijacija u GC-MS analizi koji utječe na ponovljivost istraživanja. Prije derivatizacije uzorak je potrebno osušiti. Za sušenje se najčešće primjenjuje vakuumsko centrifugiranje ili liofilizacija. Tri najčešća derivatizacijska agensa su: *N,O*-bis(trimetilsilil)trifloroacetamid (BSTFA), *N*-metil-*N*-(trimetilsilil)trifloroacetamid (MSTFA) i *N-tert*-butildimetilsilil-*N*-metiltrifloroacetamid (MTBSTFA). BSTFA i MSTFA tvore trimetilsilil (TMS) derivate metabolita, dok MTBSTFA tvori *tert*-butildimetilsilil (TBDMS) derivate. [6]

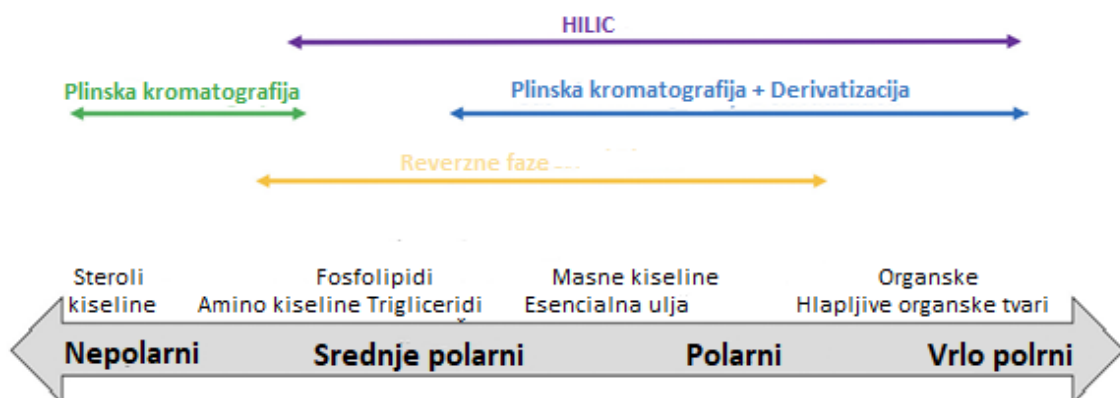
1.4. Analitičke tehnike u metabolomičkoj analizi

Za metabolomičke analize koriste se tehnike za separaciju i tehnike visoke rezolucije za identifikaciju metabolita. Za separaciju se najčešće koriste: plinska kromatografija (eng.

Gass Chromatohraphy - GC), tekućinska kromatohrafija visoke djelotvornosti (eng. *High Performance Liquid Chromatography* - HPLC) i tekućinska kromatografija ultravisoke djelotvornosti (eng. *Ultra Performance Liquid Chromatography* - UPLC). Za identifikaciju se najčešće koriste nuklearna magnetska rezonancija (NMR) i spektroskopija masa (MS). Korištenjem kemometričkih statističkih alata poput analize glavnih komponenti (eng. *Principal Component Analysis* - PCA) i metode parcijalnih projekcija najmanjih kvadrata (eng. *Partial Last Squares* - PLS) dobiva se integrirana slika oba endogenog metabolizma i metabolizma ksenobiotika. Svaka tehnika ima svoje prednosti i nedostatke i niti jedna nije pogodna za analizu svih vrsta metabolita. [7]

NMR ima mnoge prednosti, ali ima relativno malu osjetljivost u usporedbi s MS-om pa koncentracije potencijalnih biomarkera mogu biti ispod granice detekcije. GC-MS analiza zahtijeva derivatizaciju uzorka za stvaranje hlapljivih spojeva. Nehlapljive spojeve koji se ne derivatiziraju i velike ili termolabilne spojeve nije moguće odrediti GC-MS analizom. HPLC povezan s masenom spektrometrijom je prikladan za analizu labilnih i nehlapljivih polarnih i nepolarnih spojeva, bez potrebe za derivatizacijom spojeva prije analize. UPLC koristi porozne čestice unutarnjeg promjera manjeg od 2 µm, pa povezivanje s MS-om rezultira poboljšanom rezolucijom i povećanom osjetljivošću u usporedbi s konvencionalnim HPLC kolonama (HILIC ili C18 kolone [6]). NMR, GC-MS, LC-MS su trenutno prevladavajuće tehnike iako nijedna od njih ne može zadovoljiti sve zahtjeve metabolomike za određivanje svih metabolita. [7]

Na Slici 3. prikazan je odabir analitičke tehnika ovisno o polarnosti spojeva.



Slika 3.: Odabir analitičke tehnika ovisno o polarnosti spojeva. [6] HILIC - Hidrofilna interakcijska tekućinska kromatografska kolona

1.4.1. GC-MS tehnika u metabolomičkoj analizi

U ovom radu usredotočit ćemo se na GC-MS analizu pri određivanju metabolita.

Trenutno je GC-MS analiza prvi izbor za neciljanu analizu u metabolomici, a posebno za hidrofilne metabolite. Pruža vrlo učinkovite i ponovljive analize. Za odvijanje procesa na GC koloni potrebna je derivatizacija uzorka da bi nastali hlapljivi spojevi. Ne hlapljivi spojevi nisu derivatizirani i ne mogu biti detektirani GC-MS analizom. To je ujedno i najveće ograničenje ove tehnike i uvelike joj sužava područje primjenjivosti. Koristeći ovaj pristup, hlapljivi metaboliti mogu se izravno odvojiti i kvantificirati što omogućuje istovremeno profiliranje nekoliko stotina spojeva uključujući organske kiseline, većinu aminokiselina, šećere, šećerne alkohole, aromatske amine i masne kiseline. [7]

Najveća prednost GC-MS analize je postojanje velikog broja online dostupnih baza masenih spektara poput NIST baze podataka ili Golm Metabolome Database – GMD koje uvelike olakšavaju identifikaciju metabolita. [8]

1.5. Predobrada podataka

Nakon što su analitičkim tehnikama dobiveni sirovi spektri, prije same statističke analize podataka potrebna je predobrada neobrađenih podataka. Spektri tipično pokazuju razlike u obliku, širini i položaju pikova. Te razlike nastaju zbog utjecaja pozadinske buke, razlike u uzorcima ili pogreške instrumenta. Cilj predobrade podataka je ispraviti te razlike za bolju kvantifikaciju metabolita i lakšu mogućnost usporedbe različitih uzoraka. Ona obično uključuje filtriranje šuma, korekciju bazne linije, normalizaciju, poravnanje, detekciju i kvantifikaciju pikova te spektralnu dekonvoluciju. [9]

Dostupni su mnogi programi koji sadrže alate za obradu podataka prikladne u metabolomičkim istraživanjima poput XCMS, MZmine, MetAlign, MetaboAnalyst i mnogih drugih. [6]

1.5.1. Filtriranje šuma i korekcija bazne linije

Filtriranje šuma odvaja signal od pozadine koja potječe od kemijske matrice ili interferencije instrumenta. Njime se uklanja šum mjerenja ili distorzija bazne linije. Za korekciju bazne linije ručno se odabiru dva kraja pika, a zatim se primjenjuje linearna aproksimacija po djelićima pika kako bi se prilagodila zakrivljenosti bazne linije. Ovaj postupak je dugotrajan, a njegova točnost uvelike ovisi o vještinama analitičara. Stoga su razvijeni brojni algoritmi za bolju korekciju bazne linije poput ATEB (eng. *Automatic*

Two-side Exponential Baseline Correction Algorithm) i airPLS (eng. *Iteratively Reweighted Penalized Least Squares*). [10]

1.5.2. Spektralna dekonvolucija pikova

Kod kromatografskih analiza može se dogoditi da dvije ili više komponenti eluiraju na sličnim retencijskim vremenima pa im se pikovi preklapaju i na kromatogramu izgledaju kao jedan pik. [11]

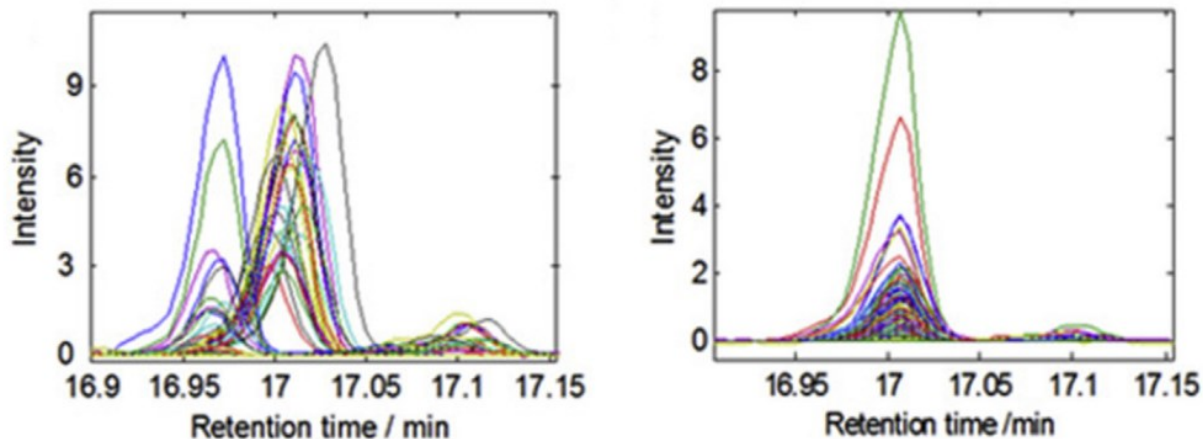
Spektralna dekonvolucija koristi se za odvajanje dviju ili više takvih koeluirajućih komponenti u masenom spektru. Nekoliko je javno dostupnih algoritama za dekonvoluciju. Najpoznatiji algoritam za dekonvoluciju u GC-MS analizi je AMDIS-a (eng. *Automated Mass Spectral Deconvolution and Identification System*). [12]

1.5.3. Definiranje, detekcija i poravnanje pikova

Nakon procesa dekonvolucije potrebno je definirati i integrirati pikove koji odgovaraju određenom spoju. Ovaj korak za cilj ima pružiti što točnije kvantitativne informacije o koncentracijama iona. [13]

Razvijene su tri strategije za detekciju pikova. Prva metoda zasebne detekcije pikova po RT i m/z vrijednostima identificira pikove s intenzitetima iznad definiranog praga za RT i m/z , neovisno jedno o drugome. Ako oba kriterija prelaze određene definirane granične vrijednosti (eng. *thresholds*), pik je identificiran. Druga metoda izdvaja kromatogram za svaki pojedinačni ion i neovisno ga obrađuje. U toj metodi svaki pojedinačni pik kromatograma predstavlja određeni mali raspon m/z vrijednosti. Trećom strategijom se dvo- ili trodimenzionalni model uzorka izotopa prilagođava sirovom signalu. Svi ovi pristupi stvaraju kvantificirane skupove podataka (liste pikova) pojedinačnih metabolita za kasniju multivarijatnu analizu skupova podataka. [12]

U metabolomici često se analizira veći broj uzoraka. Uslijed malih razlika u temperaturi i tlaku unutar kolone, matricama uzoraka ili ručnom ubrizgavanju u kolonu može doći do blagog pomaka retencijskog vremena (RT) na kromatografskoj koloni. Potrebno je poravnanje RT vrijednosti kako bi se zasebni uzorci mogli međusobno uspoređivati. Na Slici 4. prikazana je ilustracija poravnanja RT vrijednosti. [14]



Slika 4.: Ilustracija RT poravnanja [14]

1.5.4. Normalizacija

Normalizacija se primjenjuje kako bi se smanjile sustavne varijacije među uzorcima. Na taj način se postiže da uzorci budu međusobno usporedivi. [9]

Najčešći način za normalizaciju podataka je dodatak iste količine internog standarda u sve uzorke. Na temelju omjera prema internom standardu izračunava se relativna zastupljenost neke tvari u uzorku. [11]

Za biofluide, a posebice uzorke urina, koncentracije metabolita uvelike ovise o čimbenicima koji općenito nisu od interesa za metabolomičke studije. Tako primjerice količina konzumirane vode značajno utječe na koncentraciju metabolita u biofluidima. Postoje tri često korištena pristupa za normalizaciju kojima se otklanja taj neželjeni utjecaj. Prvi pristup je totalna normalizacija signala, koji izračunava ukupan intenzitet cijelog spektra i uzima ga za konstantnu vrijednost. Druga metoda je normalizacija duljine vektora. Svaki spektar koji sadrži mnogo varijabli se može promatrati kao točka u visokodimenzionalnom prostoru. Ovom metodom se euklidska udaljenost u višedimenzionalnom prostoru uzima za konstantu. Posljednjim pristupom se cijeli spektar dijeli s određenim vršnim intenzitetom, najčešće najvećim vrhom. [12]

1.6. Statistička analiza u metabolomici

Statistička analiza može biti univarijatna ili multivarijatna, a obje metode imaju svoje prednosti i nedostatke. [15]

Univarijantnu analizu je vrlo jednostavno implementirati no ona ne uzima u obzir međusobne odnose između koncentracija metabolita. Metaboliti mogu biti neovisni ili ovisni, ali su međusobno povezani biokemijskim putevima i mogu biti u međuodnosu s drugim metabolitima u sustavu. Ova metoda je ograničena na promatranje samo jedne varijable (metabolita). [15]

Za razliku od univarijantne analize, multivarijantna analiza promatra više varijabli istovremeno. Ovakva velika dimenzionalnost podataka često izaziva probleme. Dolazi do tzv. maskiranja metabolita i šuma pa se nevažne varijable mogu činiti značajnima kada to doista i nije slučaj. Te probleme se pokušava riješiti kombinacijom ovih dvaju pristupa. [15]

Većina ovih testova treba uzeti u obzir određene pretpostavke koje se mogu naći u bilo kojem statističkom udžbeniku, na primjer da su podaci prethodno prikladno obrađeni, normalizirani i skalirani na jediničnu ili približno jediničnu varijancu. [15]

1.6.1. Univarijantne metode analize

Najčešće korištene univarijantne tehnike analize u metabolomici su t-test i analiza varijance (ANOVA). p-vrijednosti najčešće se procjenjuju univarijantnim metodama, parametarskim pristupom ili permutacijskim testovima. [16]

Problem testiranja više hipoteza rezultira većom šansom za dobivanjem krivog rezultata čak i pri niskim p-vrijednostima. Korekcija mjere lažnih otkrića (eng. *False Discovery Rate* - FDR) koristi se za procjenu šanse dobivanja krivog rezultata. Q-vrijednost za svaku vrijednost procjenjuje koji je minimalni FDR pri kojem se ta vrijednost može smatrati značajnom. Postoji više metoda za procjenu ili kontrolu FDR-a. Neki algoritmi, kao što je analiza značajnosti mikronizova (eng. *Significance Analysis of Microarray* - SAM) mogu izravno procijeniti q-vrijednosti korištenjem pristupa ponovnog uzorkovanja. [9] SAM identificira gene sa statistički značajnim promjenama u ekspresiji asimilacijom skupa genski specifičnih t-testova. [16]

T-test se najčešće koristi kada se ispituje je li srednja vrijednost populacije (μ) jednaka nekoj unaprijed određenoj vrijednosti (a) na temelju slučajnog uzorka za koji se standardna devijacija (σ) procjenjuje iz uzorka umjesto da je poznata. Test se također može primijeniti na razliku između dvije srednje vrijednosti, opet s nepoznatim standardnim odstupanjima i za testove proporcija ili kako bi se utvrdilo razlikuju li se koeficijenti u regresijskim modelima značajno od nule. [17]

Jednofaktorska ili jednosmjerna analiza varijance jedan je od najčešćih i najjednostavnijih oblika univarijantne analize. Ta analiza u suštini predstavlja statističku usporedbu srednjih vrijednosti. Predstavljena je jednostavnom jednadžbom:

$$y_{ij} = \mu + T_i + e_{ij}$$

Jednadžba govori da je promatrana ili mjerena vrijednost (y) u grupama (i,j) linearna kombinacija ukupne srednje vrijednosti (μ), nekog djelovanja ili grupnog efekta (T) i neke neočekivane slučajne varijacije ili pogreške (e_{ij}). [17]

1.6.2. Multivarijantne metode analize

Multivarijantne tehnike analize koje se najčešće koriste su analiza glavnih komponenti - PCA, analiza parcijalnih najmanjih kvadrata (eng. *Partial Least Squares Analysis* - PLS) i više faktorska ANOVA. Ove metode se prvenstveno koriste za prepoznavanje uzoraka i smanjenje dimenzionalnosti podataka. Na točnost ovih modela može utjecati prekomjerno prilagođavanje zbog pretpostavki u vezi s linearnošću, neovisnošću, homoskedastičnošću i unimodalnosti. One nemaju adekvatan prediktivni kapacitet kada se primjenjuju na velike i visokodimenzionalne skupove metabolomičkih podataka jer se samim smanjenjem dimenzionalnosti gubi određeni dio informacije. [18]

Cilj multivarijantne redukcije dimenzionalnosti je pronaći K po P matricu (\mathbf{A}) koja optimalno transformira početnu matricu (\mathbf{X}) u novu matricu (\mathbf{T}) koja se sastoji od P dimenzija:

$$\mathbf{T} = \mathbf{X}\mathbf{A}$$

Dakle, svaki red nove matrice (\mathbf{T}) je transformacija odgovarajućeg reda početne matrice (\mathbf{X}). Pa tako i -ti red matrice \mathbf{X} postaje vektor-stupac x_i , a odgovarajući red matrice \mathbf{T} postaje vektor-stupac t_i . To pokazuje da je matrica \mathbf{A}^T linearna transformacija prostora ulaznih podataka (podatci iz matrica \mathbf{X}) do izlaznog prostora matrice \mathbf{T} (podatci iz matrice \mathbf{T}), koji se naziva prostorom 'rezultata':

$$\mathbf{T}_i = \mathbf{A}^T x_i$$

U slučaju kad je P manji od K dimenzionalnost će se smanjiti (tj. dobivena matrica će imati manje dimenzije od ulazne). Ovo je ključna karakteristika multivarijantne analize u metabolomici. Optimalna transformacija matrice ovisi o odabranom logaritmu transformacije (PCA, PLS, ...). [19]

Analiza glavnih komponenti (PCA) je najraširenija multivarijatna analiza u kemometriji općenito. Cilj joj je linearnom transformacijom prevesti početnu informaciju u što je moguće manje dimenzija koje i dalje sadržavaju glavne karakteristike početne informacije. Pojednostavljeno, PCA traži jednodimenzionalno područje koje ima najveću varijancu. Ovo područje se naziva prva glavna komponenta (PC1). Zatim traži područje koje sadrži najveću moguću preostalu varijancu. Ovo područje naziva se sekunda glavna komponenta (PC2). Proces se ponavlja dok se ne obuhvati cijelo područje varijance, a analitičar u obzir uzima onoliko glavnih komponenti koje sadrže zadovoljavajući postotak sveukupne varijance. Svaka glavna komponenta je ortogonalna (okomita) na sve ostale. [12]

Analiza parcijalnih najmanjih kvadrata (PLS) koristi linearnu regresiju za pronalaženje najboljeg linearnog prediktora neke varijable (Y) na temelju zavisne varijable (X). Te varijable se određuju na način da se maksimizira kovarijanca (koliko se dvije varijable mijenjaju zajedno) između X i Y varijabli. Zbog toga omogućuje analizu podataka kod koje je velik broj varijabli snažno koreliran. [19, 20]

PLS za izračunavanje koristi informacije iz dvaju skupova (X i Y), dok PCR za računanje koristi isključivo informacije iz jednog (X). Time se X i Y dekomponiraju na skorove i opterećenja. PLS možemo promatrati kao PCA koja se provodi na X i Y, pri tome se glavne komponente dodatno rotiraju u prostoru čime se postiže maksimalna kovarijanca između X i Y. Poput PCA, PLS također traži novu matricu (T), koja određuju varijable modela. PLS se može udruživati i sa ostalim analitičkim tehnikama da se dobiju još složeniji prediktivni modeli, poput diskriminantne analize parcijalnih najmanjih kvadrata (eng. *Partial Least-Squares Discriminant Analysis* - PLS-DA) ili ortogonalne diskriminantne analize parcijalnih najmanjih kvadrata (eng. *Orthogonal Partial Least-squares Discriminant Analysis* - OPLS-DA). [20]

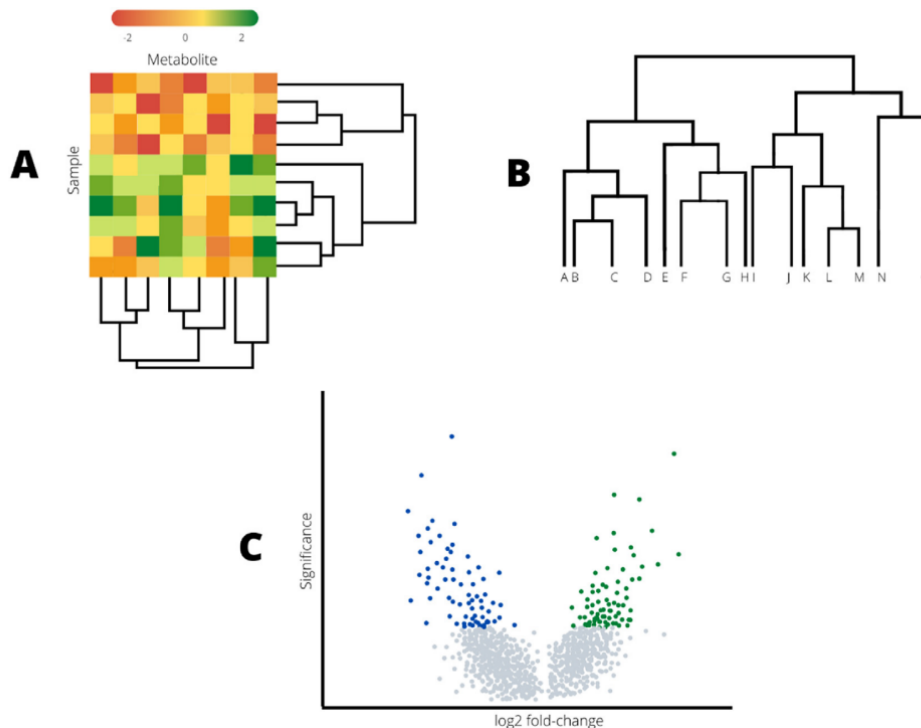
1.6.3. Alati za vizualizaciju

Osim klasičnih PCA i PLS-DA statističkih analiza koje vizualno predstavljaju rezultate u metabolomici se koriste i drugi alati za vizualizaciju poput: grafičkog prikaza vulkana (eng. *volcano plot*), toplinskih karti (eng. *heatmap*), grupiranja (eng. *cluster analysis*) ili samoorganiziranih mapa (eng. *Self Organized Maps* – SOM). Ovi alati pomažu da se sažeto predstave razlike između referentnog ili kontrolnog skupa podataka i testnog skupa podataka, također spadaju u multivarijantne metode analize. [18]

Grafički prikaz vulkana je dijagram raspršenosti koji pomoću FDR-a i omjera promjene metabolita (engl. *Fold change* - omjera koncentracije metabolita u kontrolnoj skupini i uzorku) pomaže u identifikaciji statistički značajnih metabolita. Toplinske karte koriste boju za predstavljanje intenziteta promatranih signala (boje signala odgovaraju njihovim intenzitetima). [18]

Postoje dvije vrste algoritama za grupiranje. Nehijerarhijski algoritam grupiranja jednostavno dijeli podatke u grupe. Najjednostavniji primjer nehijerarhijskog grupiranja je K-means grupiranje. Srednja vrijednost uzoraka u grupi je centar te grupe, a uzorci se svrstavaju u grupe na način da svaki uzorak odlazi u grupu s bližim centrom. Broj grupa (K) dolazi iz podataka, pa ako uzorke želimo podijeliti na bolesne i zdrave imamo dvije skupine. Drugi algoritam grupiranja je hijerarhijsko grupiranje koji promatra i grupira uzorke kao zasebne grupe. Te zasebne grupe se spajaju na temelju sličnosti sve dok na kraju svi uzorci ne ostanu u istoj grupi. Rezultat je prikazan kao dendrogram odnosno drvoliki dijagram koji ocrta uzajamne odnose među grupama. [12]

Samo organizirane mape (SOM) koriste pravokutne ili heksagonalne koordinatne mreže određene veličine. Metoda se temelji na učenju iz popunjavanja mreže podacima reprezentativnog uzorka i vizualizaciji podataka koji se raspoređuju u mrežu na temelju naučenih obrazaca. [12]

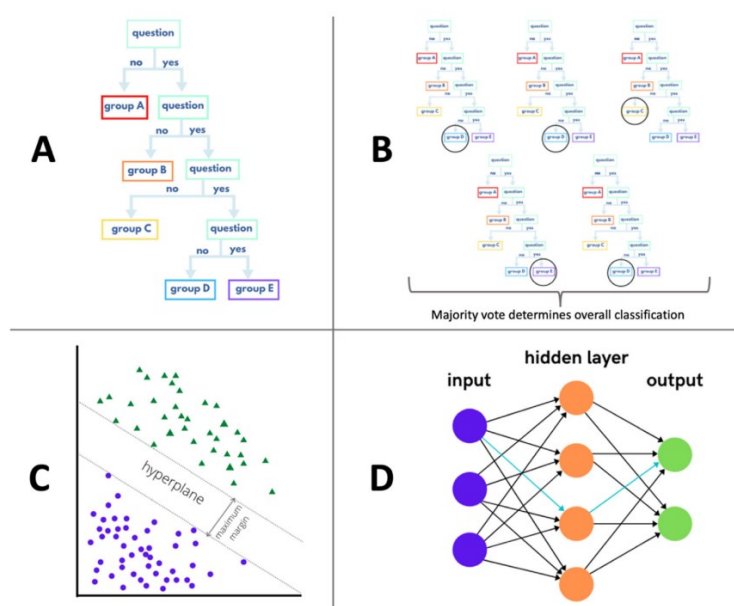


Slika 5.: A toplinska karta, B hijerarhijsko grupiranje, C grafički prikaz vulkana [18]

1.6.4. Alati za strojno učenje

Najpoznatiji alata za strojno učenje su: metoda potpornih vektora (eng. *Support Vector Machines* - SVM), klasifikacijska i regresijska stabla (eng. *Classification and Regression Trees* - CART), umjetne živčane mreže (eng. artificial neural networks – ANN) i slučajne šume (eng. *Random Forests*). [18]

SVM se koristi za klasifikaciju i regresijsku analizu. Kod ove metode spektri su vektori, a kombinacijom spektara cilj je pronaći minimalni skup vektora koji se nazivaju potporni vektori. Ti potporni vektori definiraju rubne uvjete i model. CART je algoritam koji se koristi za izradu modela predviđanja iz velike količine podataka, a rezultirajući model je predstavljen grafički kao stablo odlučivanja. Sažeto razmatra doprinose varijabli, ali je nepouzdan kod mnogovarijantnih skupova podataka. ANN je model koji je dizajniran da simulira funkcije niske razine živčanog sustava. Model je sastavljen od zbirke povezanih nelinearnih obradnih elemenata. Mreža automatski generira identifikacijske karakteristike iz obrađenih podataka za trening. ANN dobro rješava složene odnose, ali su dobiveni rezultati dosta teški za interpretirati te je potrebno mnogo vremena za osposobiti (utrenirati) mrežu. Nasumična šuma gradi zbirku stabala odlučivanja za klasifikaciju metabolomskih podataka, svako stablo glasuje, a šuma zatim određuje ukupni poredak. Stabla odlučivanja su algoritmi koji se koriste za predviđanje modela iz velike količine podataka. Na Slici 6. dani su shematski prikazi ovih modela. [18]



Slika 6.: Slika prikazuje CART (A), nasumične šume (B), SVM (C) i ANN (D). [18]

1.7. Softverski alati u metabolomici

Za izvođenje metabolomičke analize neophodni su snažni softverski alati zbog goleme količine i raznovrsnosti podataka koje je potrebno obraditi. Metabolomička analiza zahtjeva obradu podataka na više načina. Potrebna je sposobnost obrade sirovih spektralnih podataka i statistička analiza za pronalazak značajnih metabolita. Usporedba podataka s bazama metabolita za identifikaciju metabolita još je jedan važan dio analize. Uz sve to potrebna je bioinformatička analiza te vizualizacija molekularnih interakcijskih mreža. Na kraju važnu ulogu ima integracija i analiza multiomičkih podataka (povezivanje s ostalim poljima: genomikom, proteomikom...). Važno je napomenuti da ne sadrže svi softveri sve gore navedene mogućnosti obrade podataka pa se u analizi često koriste i više softvera, svaki za određeni dio analize. [21]

1.7.1. Softverski alati za predobradu podataka

Jednom prikupljeni neobrađeni podaci (kromatogrami, spektri ili NMR podaci) obrađuju se softverskim alatima za olakšavanje kvantifikacije spojeva. Ti alati mogu biti komercijalni ili javno dostupni. [22]

Neki komercijalni softveri specifični za instrumente su Bruker Data Analysis i Bruker Profile Analysis za Bruker, MassHunter za Agilent, MarkerLynx za Waters. Najpoznatiji javno dostupni softveri su XCMS, MZmine i MET-IDEA koji mogu rukovati podacima iz različitih instrumenata. [23]

1.7.2. Baze podataka u metabolomici

Kod neciljane metabolomičke analize, za identifikaciju metabolita iz spektra, potrebne su različite baze podataka. Također, softverski alati za izvođenje analize obogaćivanja i analize metaboličkih puteva zahtijevaju baze podataka na temelju kojih mapiraju značajne metabolite u poznate biokemijske putove. [22]

Znanstvenici u drugim područjima se uglavnom oslanjaju na komercijalne baze podataka za analizu (NIST Standardna referentna MS baza i Aldrich Spectral Viewer NMR baza) ili razvijaju prilagođene, interne baze podataka s ograničenim ili nikakvim javnim pristupom. Nasuprot tome u metabolomici se često postupa na drugi način. Prihvaćene su baze podataka s otvorenim pristupom te razmjena podataka svih aspekata eksperimentalnog procesa: od praćenja uzorka, algoritama za analizu podataka i na kraju objavljivanja metapodataka. [24]

U posljednje vrijeme razvijene su mnoge spektralne baze podataka otvorenog pristupa za masene spektre i NMR spektre poput: Golm Metabolome Database - GMD (GC-MS), METLIN (MS/MS), MassBank (GC-MS i HPLC-MS), NMRShiftDB (NMR). Također su razvijene i baze podataka koje nude usporedbu s više spektralnih parametara (npr. MS spektri i vrijednosti NMR pomaka) za točniju identifikaciju poput SDBS i HMDB baze. [24]

Na kraju izrađene su i baze podataka koje ne sadrže spektre već smještaju metabolite u određene skupine poput KEGG i SMPB baze. Najčešće su im u fokusu metabolički putevi, bolesti ili utjecaj lijekova. [25, 26]

Kako bi istraživanja bila javno dostupna, Nacionalni metabolomički repozitorij (eng. *National Metabolomics Data Repository* - NMDR), Metabolomics Common Funda smješten na Kalifornijskom sveučilištu u San Diegu razvio je Metabolomics Workbench. Metabolomics Workbench je nacionalni i međunarodni repozitorij za metabolomičke podatke i metapodatke. [27]

1.7.3. MetaboAnalyst 5.0

MetaboAnalyst je sveobuhvatna platforma namijenjena analizi metabolomičkih podataka putem javno dostupnog korisničkog sučelja na internetu (<https://www.metaboanalyst.ca/>). Kako je navedeno na stranici, MetaboAnalyst je proteklog desetljeća postao najkorištenija platforma za obradu metabolomičkih podataka s preko 300.000 korisnika. MetaboAnalyst verzije od 1.0 do 3.0 uglavnom su se bazirale na općoj statističkoj i funkcionalnoj analizi ciljanih metabolomskih podataka. S verzijom 4.0, postupno se usmjerava na rješavanje složenije bioinformatike te rješavanje statističkih potreba iz globalnih metabolomskih podataka, uključujući obradu sirovih spektra, funkcionalnu analizu i integraciju s drugim omičkim podacima. Trenutačni verzija MetaboAnalyst 5.0 podržava obradu sirovih MS spektra, sveobuhvatnu normalizaciju podataka, statističku analizu, funkcionalnu analizu, meta-analizu te integraciju i analizu multiomičkih podataka. Cilj je omogućiti visokoučinkovitu ciljanu i ne ciljanu metabolomičku analizu te prevesti sirove podatke u biološki značajan odgovor. Veći nedostatak MetaboAnalysta je to što trenutno ne podržava obradu sirovih spektra iz GC-MS ili MS/MS analiza, koje se također često koriste u globalnoj metabolomici. [28, 29]

U Tablici 2. prikazana je usporedba raznih verzija MetaboAnalysta sa ostalim sličnim softverskim platformama za metabolomiku.

Tablica 2.: Usporedba Raznih verzija MetaboAnalysta s drugim sličnim softverima za metabolomiku. [30]

	MetaboAnalyst			XCMSOnline	W4M	3Omics	NOREVA
	5.0	4.0	3.0				
Obrada surovih spektara							
Optimizacija parametara	+++	-	-	+	++	-	-
Podržani algoritmi	+++	-	-	++	++	-	-
Ponovljivost analize	+++	-	-	-	+	-	-
Anotacija spojeva	+	+	-	+++	++	-	-
Statistička analiza							
Univarijantna	+++	++	++	+	+	-	+
Multivarijantna	+++	++	+	+	+++	-	++
Klasteriranje	+++	+++	++	+	+	-	-
Analiza statističke snage	√	√	√	-	-	-	-
Vremenski - serijska analiza	√	√	√	-	-	-	√
Analiza biomarkera	√	√	√	-	-	-	-
Meta - analiza biomarkera	√	√	-	-	-	-	-
Funkcionalna analiza							
Funkcionalna analiza (MS pikovi)	+++	++	-	++	-	-	-
Sanaliza obogaćenja	+++	+	+	-	-	++	-
Analiza puteva	+++	++	+	-	-	++	-
Funkcionalna - meta analiza	+++	-	-	++	-	-	-
Integrativna analiza							
Analiza zajedničkog puta	+++	+	+	-	-	+++	-
Mreže bazirane na saznanjima	++	++	-	-	-	++	-
Mreže bazirane na putevima	++	-	-	-	-	-	-
Ostalo							
Normalizacija podataka	++	+	+	-	+	-	+++
Procjena vrijednosti koje nedostaju				-	-	-	
Spajanje tehničkih replikata	√	-	-	-	-	-	-

Oznake: '√' podržava opciju, '-' nema opciju i '+' kvantitativnija procjena (više '+', bolje podržana opcija)

- XCMS online: <https://xcmsonline.scripps.edu/>.
- Workflow4Metabolomics (W4M): <https://workflow4metabolomics.usegalaxy.fr/>.
- 3Omics: <https://3omics.cmdm.tw/>.
- NOREVA: <http://idrblab.cn/noreva/>.

Web sučelje MetaboAnalysta temeljeno je na JavaServer Faces tehnologiji pomoću PrimeFaces (v11.0) biblioteke. Programski jezik JavaScript je upotrijebljen za interaktivnu vizualizaciju. Za pozadinske izračune upotrebljavaju se funkcije napisane u programu R. [31] MetaboAnalystR paket je također javo dostupan za instalaciju, namijenjen je korisnicima upoznatim s radom u R programu. Ovaj paket ima primjenu u analizi sirovih spektara i analizi serije podataka. [32]

MetaboAnalyst 5.0 nudi 12 modula (softverskih alata) za obradu metabolomičkih podataka. Alati služe za statističku analizu, analizu biomarkera, analizu metaboličkih putova i mreža, analizu obogaćenja (eng. *enrichment analysis*), cjelokupnu predobradu

sirovih LC-MS podataka, analizu MS pikova te analizu statističke snage. Za predobradu sirovih LC-MS podataka podaci za unos moraju biti u mzML, mzXML ili mzData formatima, dok je kod ostalih alata za analizu dovoljno za unos podataka imati pripremljene tablice u .txt ili .csv formatu. [28]

1.7.3.1. Statistička analiza

MetaboAnalyst 5.0 nudi tri alata za statističku analizu: klasičnu statističku univarijantnu i multivarijantnu analizu, statističku analizu uz tablicu metapodataka, te statističku meta-analizu.

Klasična statistička analiza nudi najčešće korištene statističke metode i metode strojnog učenja. Za univarijantnu analizu dostupni su: omjer promjene metabolita, t-test, dijagram vulkana, ANOVA, korelacijska analiza, analiza značajnosti mikronizova (SAM). Za multivarijantnu analizu dostupni su: analiza glavnih komponenti (PCA), diskriminantna analiza parcijalnih najmanjih kvadrata (PLS-DA) i ortogonalna diskriminantna analiza parcijalnih najmanjih kvadrata (OPLS-DA), dendrogram, toplinske karte, K-means, samoorganizirane mape (SOM), slučajne šume i metoda potpornih vektora (SVM). [28]

Drugi alat za statističku analizu dostupan je ukoliko su prilikom istraživanja prikupljene karakteristike fenotipa (npr. dob, spol) i drugi eksperimentalni čimbenici te formirani u dodatnu tablicu meta podataka. Pomoću tih dodatnih podataka moguća je vizualizacija i izračunavanje povezanosti između fenotipova i metabolomskih značajki s obzirom na druge eksperimentalne čimbenike ili druge varijacije. [28, 33]

Statistička meta-analiza se koristi u koliko se želi proizvesti statističku analizu za više setova podataka koji su prikupljeni pod usporedivim uvjetima kroz više zasebnih istraživanja. Ova analiza se koristi za identifikaciju robusnih biomarkera (spojeva ili pikova) koji se pojavljuju više usporedivih tj. sličnih istraživanja. [28]

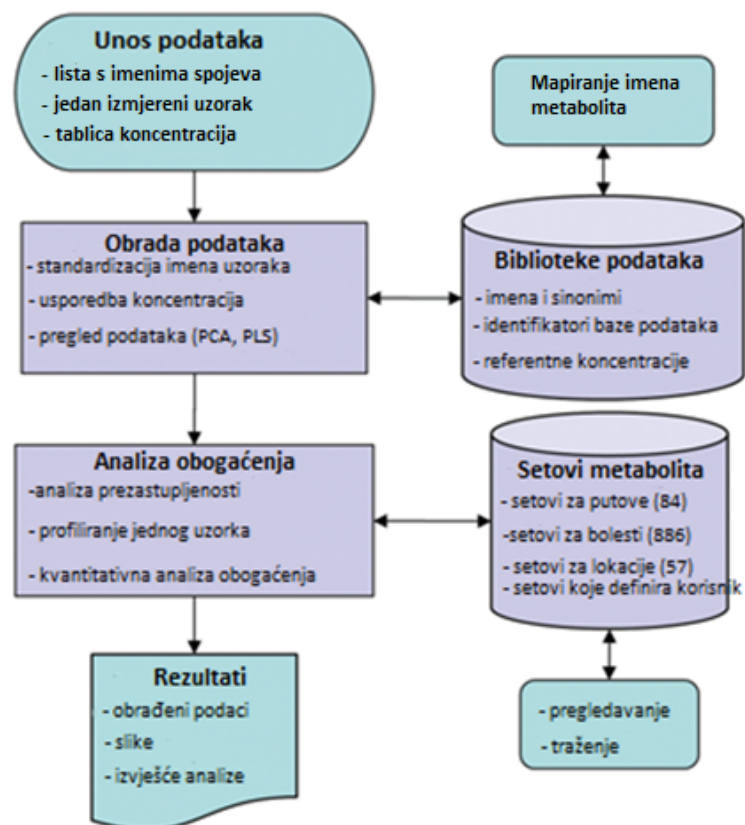
1.7.3.2. Analiza biomarkera

MetaboAnalyst 5.0 koristi krivulje osjetljivosti (eng. *Receiver Operating Characteristic* - ROC krivulje) za identifikaciju potencijalnih biomarkera i procjenu njihove učinkovitosti. Pruža klasičnu univarijantnu i multivarijantnu analizu ROC krivulje. [28] ROC krivulja je graf osjetljivosti u odnosu na lažno pozitivnu stopu (1-specifičnost) za sve moguće vrijednosti graničnih točaka. [34]

1.7.3.3. Analiza obogaćivanja (enrichment analysis)

MetaboAnalyst provodi analizu obogaćivanja skupa metabolita (eng. *Metabolite Set Enrichment Analysis* - MSEA). [28]

MSEA je web aplikacija dizajnirana za pomoć pri identifikaciji i tumačenju obrazaca promjena koncentracije metabolita u biološki značajnom kontekstu. MSEA pomoću popratnih baza podataka omogućuje kreiranje popisa izmijenjenih (modificiranih) metabolita iz biotekućine ili uzorka tkiva. Ta informacija se koristi za predlaganje biološkog puta ili bolesti za dalja istraživanja. Trenutno se koristi pet pažljivo prikupljenih biblioteke metabolita: biblioteke bazirane na putovima, bolestima (krv, urin, cerebrospinalna tekućina) i mjestu. Za izradu biblioteka korištene su SMPDB i HMDB baze podataka te podatci prikupljeni iz brojnih radova. Ovisno o tipu podataka za unos, MSEA nudi tri vrste analize obogaćivanja: analiza prezastupljenosti (eng. *Overrepresentation Analysis* – ORA), profiliranje jednog uzorka (eng. *Single Sample Profiling* – SSP) i kvantitativan analiza obogaćivanja (eng. *Quantitative Enrichment Analysis* – QEA). ORA se koristi za procjenu, kada se želi utvrditi je li određeni skup metabolita zastupljeniji od očekivanja unutar zadane složene liste. Kako podatak za unos koristi se samo popis spojeva. SSP se koristi pri metabolomičkim analizama ljudskih biotekućina (krv, urin ili cerebrospinalna tekućina) u kojima su normalni rasponi koncentracija metabolita poznati. SSP određuje jesu li određene koncentracije metabolita u danom uzorku značajno više ili niže od njihovih normalnih raspona. U ovom slučaju podatci se unose kao tablica od dva stupca s imenima spojeva i njihovim koncentracijama. QEA vrši analizu obogaćivanja izravno iz neobrađenih podataka o koncentraciji i ne zahtijeva popis značajno promijenjenih komponenti. Za ovu analizu je potrebno imati podatke kao tablicu koncentracije koja sadrži podatke o koncentraciji metabolita iz više uzoraka. Uz QEA obogaćeni setovi metabolita mogu se identificirati kada je samo nekoliko spojeva značajno promijenjeno ili kada je više spojeva tek neznatno (ali dosljedno) promijenjeno. Na Slici 7. prikazan je tijek rada u programu MSEA. [35]



Slika 7.: MSEA se sastoji od četiri koraka: unos, obrada podataka, analiza podataka i preuzimanje podataka. [35]

1.7.3.4. Analiza metaboličkih putova

Analiza metaboličkih putova integrira analizu obogaćivanja i analizu topologije puta te vizualizaciju za 26 modelnih organizama (čovjek, miš, štakor, krava, kokoš, E.coli, ...) [28]

Analiza obogaćivanja puta nudi analizu prezastupljenosti (ORA) i kvantitativnu analizu obogaćivanja (QEA). Analiza topologije puta govori o važnosti spoja unutar određene metaboličke mreže. Važnost se određuje mjerenjem centralnosti određenog spoja u danoj metabolomskoj mreži. Centralnost se izražava na dva načina: međucentralnost (eng. *Cetweenness Centality*) i stupanj centralnosti (eng. *Degree Centrality*). Međucentralnost predstavlja broj veza čvora od interesa s drugim čvorovima, a stupanj centralnosti predstavlja broj najkraćih puteva koji prolaze kroz čvor od interesa. Utjecaj puta se izračunava kao zbroj mjere važnosti odgovarajućih metabolita normaliziranih zbrojem mjere važnosti svih metabolita u svakom putu. [36, 37]

MetaboAnalyst uzima metaboličke putove iz KEGG baze podataka i vizualizira ih kao mreže kemijskih spojevi, s metabolitima kao čvorovima i reakcijama kao spojnica. [37]

1.7.3.5. Zajednička analiza puta (joint-pathway analysis)

Pomoću zajedničke analize puta MetaboAnalyst pruža integriranu (multiomičku) analizu gena i metabolita od interesa u kontekstu metaboličkih putova. Trenutno su podržava podatke o metabolomici 25 modeliranih organizama uključujući čovjeka, miša i štakora. [28]

1.7.3.6. Istraživanje mreža (network explorer)

Pomoću ovog modula je moguće istraživanje mreža na dva načina. Prvi način je učitavanje popisa metabolita, gena ili KEGG ortologa, a zatim vizualno istraživanje njihovih odnosa u različitim biološkim mrežama. U drugom načinu učitavaju se tablice podatak za izvođenje nepristrane rijetke djelomične korelacijske analize mreže (eng. *Debiased Sparse Partial Correlation - DSPC*) i vizualno analiziraju. [28]

Vizualno istraživanje se koristi za dobivanje novih uvida ili pomoć pri razvoju novih hipoteza. Ono nadopunjuje zajedničku analizu puta na način da omogućava identifikaciju veza koje prelaze granice konvencionalnih putova, a također omogućuju globalni pristup na funkcionalne promjene koje možda nisu očite kada se ispituju pojedinačni putovi. Trenutačno je podržano pet tipova bioloških mreža uključujući globalnu metaboličku mrežu KEGG, mrežu interakcije gena i metabolita, mrežu interakcije bolesti i metabolita, mrežu interakcije metabolita i metabolita te mrežu interakcije bolest, metabolita i gena. [38] DSPC algoritam stvara mreže korištenjem grafičkog LASSO modela za izračunavanje parcijalnih korelacijskih koeficijenata i P-vrijednosti za svaki par značajki u skupu podataka. Rezultat se vizualno prikazuje kao interaktivna mreža čvorova (spojeva) na način da veličina čvora odgovara stupnjevima (broj spojeva s ostalim čvorovima), a debljina spoja ta dva čvora ovisi o njihovoj korelaciji. [30]

1.7.3.7. Obrada LC-MS spektara

MetaboAnalyst 5.0 nudi obradu sirovih LC-MS spektara. To uključuje definiranje, detekciju i poravnanje pikova pomoću automatski optimiziranih parametara programa XCMS i CAMERA. [28] Podaci moraju biti u otvorenim mzML, mzXML, netCDF ili

mzData formatima. Kad je obrada podataka dovršena moguć je vizualni pregled rezultata pomoću interaktivnog 3D PCA dijagramu, kromatograma ukupne ionske struje (eng. *Total Ion Chromatogram* - TIC), kromatogram baznog vrha (eng. *Base Peak Intensity* - BPI), rezultati korekcije RT-a itd. Klikom na bilo koju značajku od interesa moguć je prikaz odgovarajućeg kromatograma ekstrahiranih iona (eng. *Extracted Ion Chromatogram* - EIC). Dobivene rezultate moguće je preuzeti ili ih odmah preusmjeriti na druge module te provesti statističku ili funkcionalnu analizu. [30]

1.7.3.8. Funkcionalna analiza MS pikova

Ovaj modul, korištenjem LC-MS podataka visoke rezolucije, izvodi analizu obogaćivanja metaboličkih putova i vizualizaciju na temelju mummichog algoritma (algoritam kojem je cilj predvidjeti biološke aktivnost u mreži metabolita izravno iz tablice MS pikova, zaobilazeći identifikaciju metabolita [39]). [40]

Meta-analiza neciljanih MS podataka proširuje funkcionalnu analizu MS pikova tako što pomaže kod smanjenja pristranost koju pojedinačne analize mogu imati prema specifičnim protokolima za obradu uzoraka ili LC-MS instrumentima. Cilj f meta-analize je identificirati funkcionalne profile kroz druga neovisna istraživanja ili udruživanjem pikova iste analize provedene komplementarnim instrumentima. [28]

1.7.3.9. Analiza statističke snage

U ovom modulu moguće je učitati podatke pilot istraživanja ili iz drugih sličnih istraživanja, a program zatim izračuna minimalni broj uzoraka koji su potrebni za otkrivanje statistički značajne razlike između dviju populacija. Stupanj pouzdanosti s kojim su razlike otkrivene određuje sam korisnik. [40]

2. EKSPERIMENTALNI DIO

Cilj ovog rada je istražiti utjecaj spoja (*E*)-3-amino-5-(3-(3-bromfenil)akrilil)-*N*-(3-kloro-2-metilfenil)-6-metiltieno[2,3-*b*]piridin-2-karboksamida (u nastavku teksta tieno[2,3-*b*]piridin) na metabolizam dvije vrste stanica raka dojke MCF-7 i MDA-MB-231.

2.1. GC-MS analiza ekstrakta stanica raka

Za analizu je korišten Agilent 8890 GC sustava povezan s triple quad spektrometarskim sustavom MS 7000D GC/TQ. Korištena je kolona HP-5 MS (30 m × 0,25 mm × 0,25 μm, Agilent, Santa Clara, CA, SAD).

Temperaturni program je u početku postavljen na temperaturu od 60 °C te je ta temperatura održavana 2 minute. Zatim je temperatura povećana do 210 °C brzinom od 10 °C/min, nakon čega slijedi povećanje do 240 °C brzinom od 5 °C/min te povećanje do 315 °C brzinom od 25 °C/min. Na kraju je temperatura od 315 °C održavana 3 minute.



Slika 8.: Agilent 8890 GC [41]

2.2. Predobrada GC-MS podataka

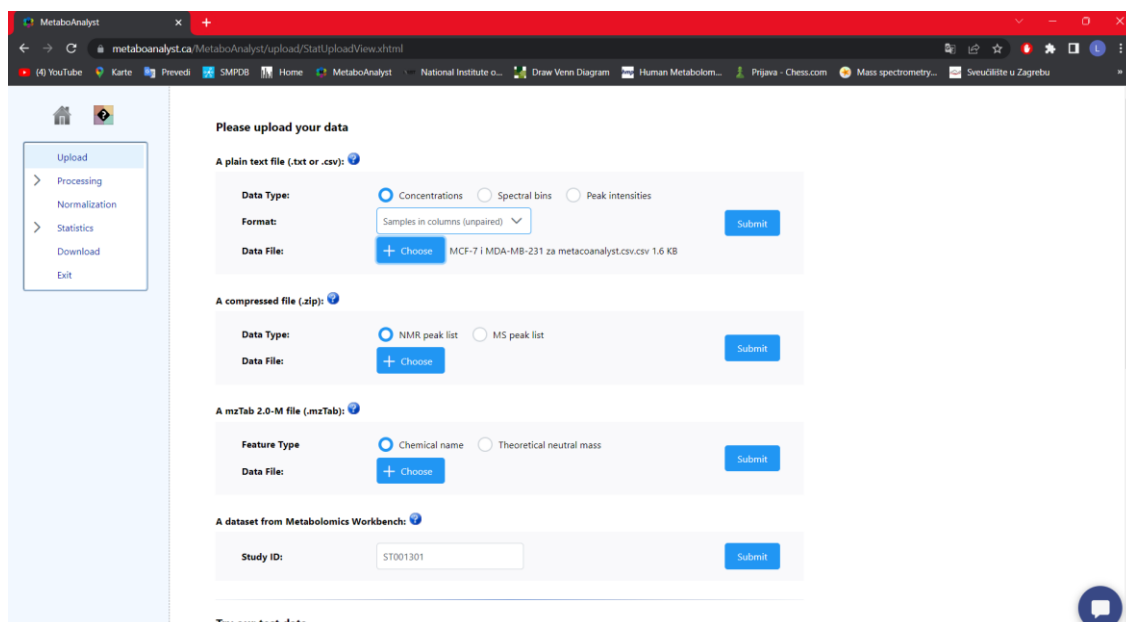
Za definiranje pikova, poravnanje RT vremena te označavanje i integraciju pikova korišten je softver Agilent MassHunter Qualitative Analysis.

Za identifikaciju metabolita korištena je NIST baza podataka. Vrijednost intenziteta za svaki metabolit je normalizirana pomoću ribitola koji je korišten kao interni standard.

2.3. Statistička analiza podataka u programu MetaboAnalyst

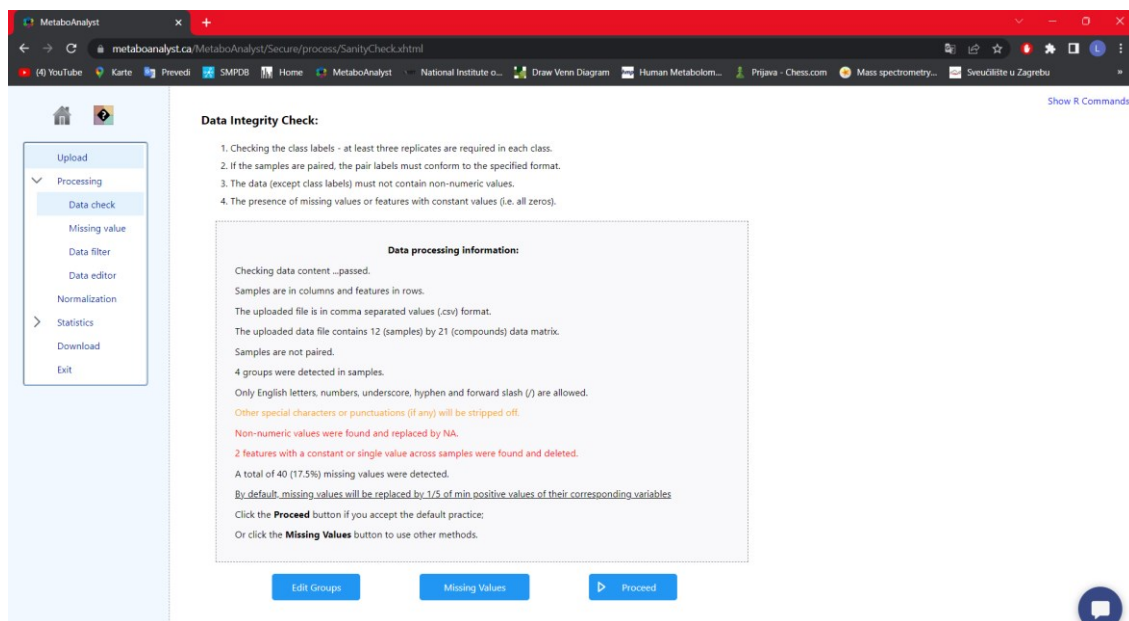
Platforma za analizu metabolomičkih podataka MetaboAnalyst je korištena za analizu različito eksprimiranih metabolita kroz dvije stanične linije.

1.) Podaci su učitani u .csv formatu, kao nespareni uzorci poredani u stupce.



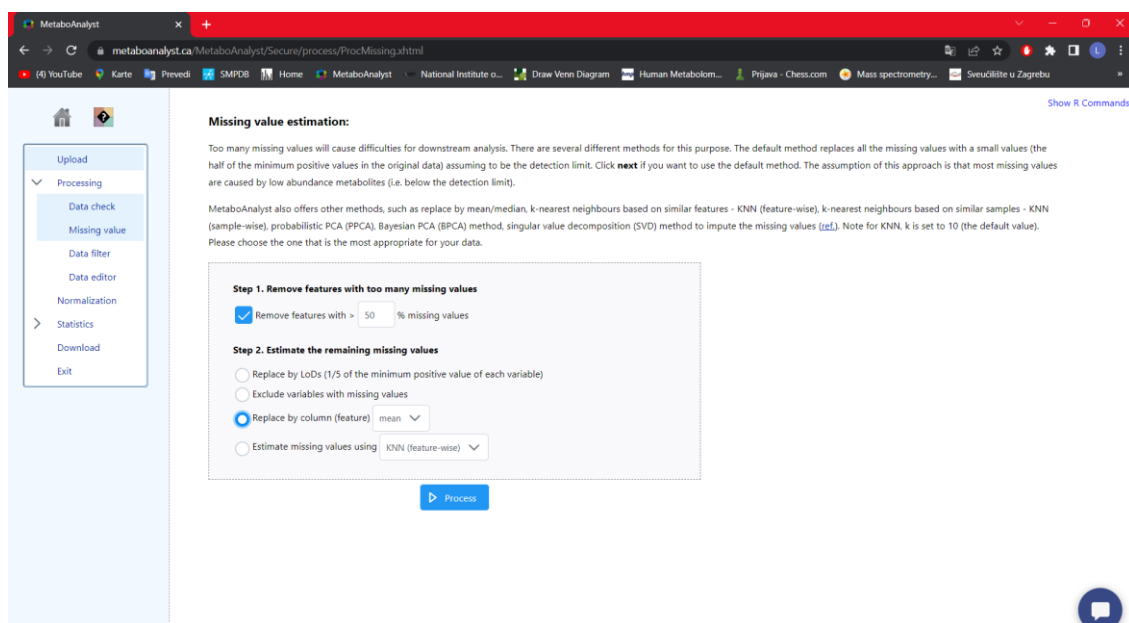
Slika 9.: Učitavanje podataka

2.) Nakon unosa podataka pritiskom na tipku **Submit** sustav preusmjerava do dijela za provjeru unesenih podataka.



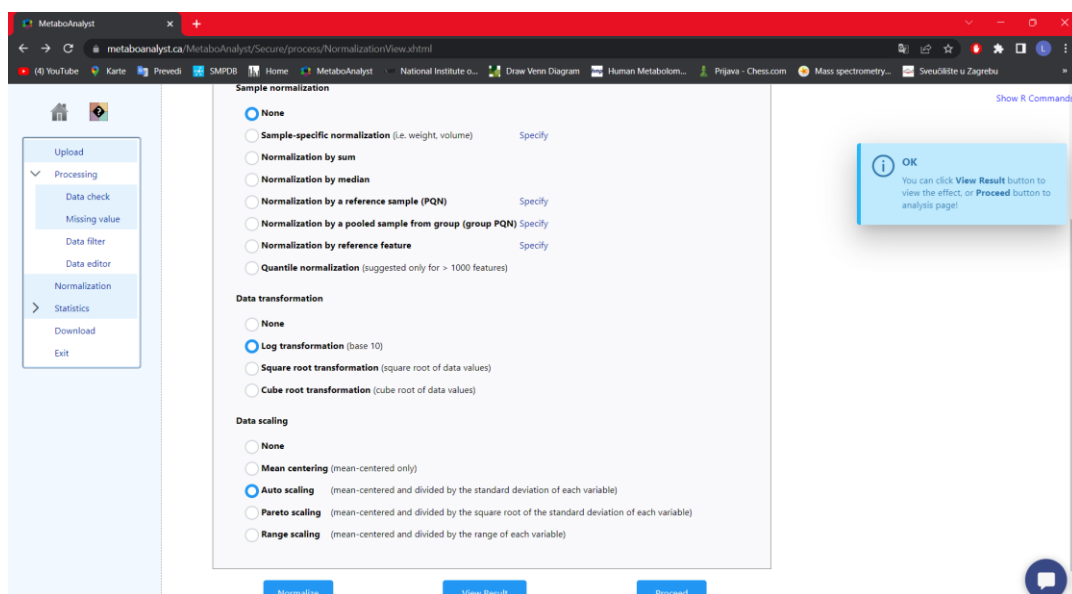
Slika 10.: Provjera unesenih podataka

Klikom na tipku **Missing Values** postavljen je način procjene vrijednosti koje nedostaju. Podaci kojima nedostaje više od 50% vrijednosti izuzeti su iz analize, ostale vrijednosti koje nedostaju procijenjene su na srednju vrijednost ostalih mjerenja.



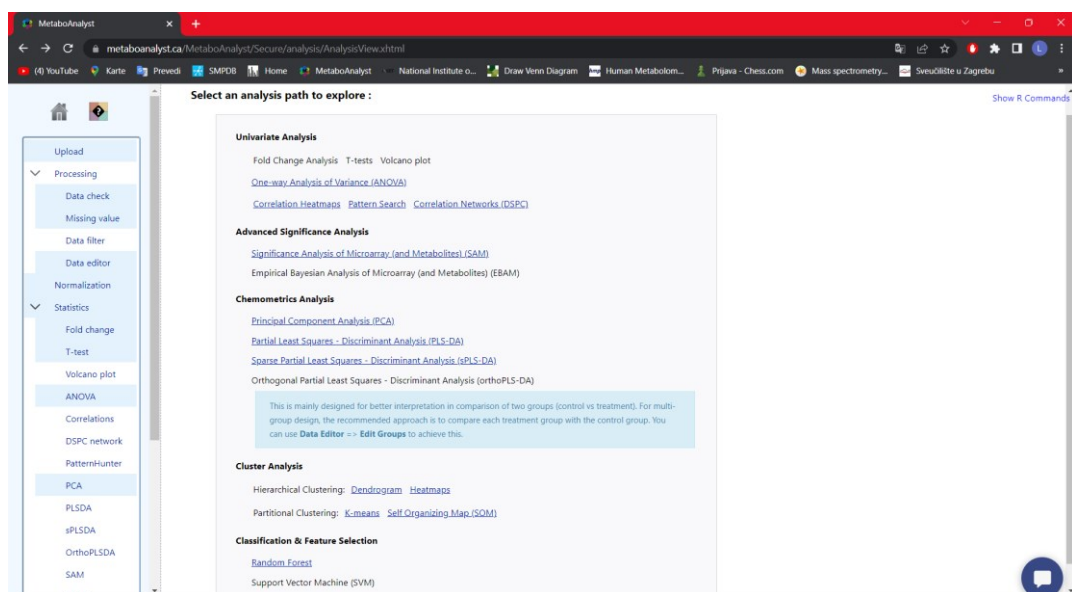
Slika 11.: Procjena vrijednosti koje nedostaju

3.) Određivanje parametara za normalizaciju, transformaciju i skaliranje podataka. Kako je normalizacija već prethodno provedena pomoću internog standarda, normalizaciju u aplikaciji nismo koristili te smo taj parametar postavili na **None**. Transformacija podataka je postavljena na logaritamsku transformaciju. Skaliranje podataka postavljeno je na automatsko skaliranje.



Slika 12.: Normalizacija, transformacija i skaliranje podataka

4.) Klikom na ikonu **Normalize**, zatim **Proceed** sustav preusmjerava na panel u kojem je moguće odabrati potrebne analize.



Slika 13.: Statistička analiza podataka

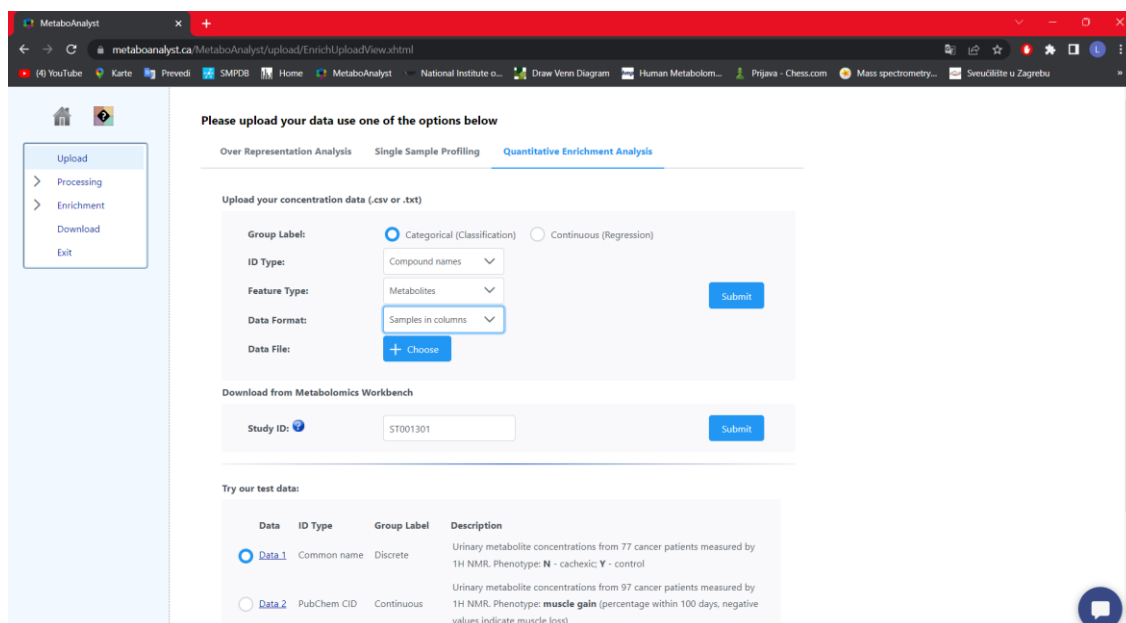
Analiza statističke značajnosti provedena je t-testom. Jednosmjerna ANOVA korištena je za usporedbu učinka tieno[2,3-b]piridina na obje stanične linije. Analiza glavnih komponenti (PCA) korištena je kao metoda klasteriranja.

2.4. Analiza obogaćenja u programu MetaboAnalyst

Analiza obogaćivanja skupa metabolita (MSEA) korištena je za uspostavljanje veze između metaboličkih otisaka i promjena koncentracije metabolita.

Postupak učitavanja podataka sličan je postupku za statističku analizu.

Korišten je modul za kvantitativnu analizu obogaćivanja. Prilikom unosa podataka naznačeno je da su korišteni podatci o metabolitima.



Please upload your data use one of the options below

Over Representation Analysis Single Sample Profiling **Quantitative Enrichment Analysis**

Upload your concentration data (.csv or .txt)

Group Label: Categorical (Classification) Continuous (Regression)

ID Type:

Feature Type:

Data Format:

Data File:

Download from Metabolomics Workbench

Study ID:

Try our test data:

Data	ID Type	Group Label	Description
<input checked="" type="radio"/> Data 1	Common name	Discrete	Urinary metabolite concentrations from 77 cancer patients measured by 1H NMR. Phenotype: N - cachexic; Y - control
<input type="radio"/> Data 2	PubChem CID	Continuous	Urinary metabolite concentrations from 97 cancer patients measured by 1H NMR. Phenotype: muscle gain (percentage within 100 days, negative values indicate muscle loss)

Slika 14.: Unos podataka za kvantitativnu analizu obogaćivanja

Ostatak postupka je isti kao kod statističke analize uz još dva dodatna koraka. U prvom se vrši provjera metabolita koje je sustav prepoznao. Promatrani su samo spojevi navedeni u Bazi podataka ljudskog metaboloma (HMDB 4.0).

Name/ID Standardization:

- For enrichment analysis, only well-annotated HMDB compounds (i.e. those in our pathway libraries & metabolite sets) will be mapped. For general-purpose name mapping, use **Compound ID Conversion** tool in **Other Utilities** module;
- Greek alphabets are not recognized, they should be replaced by English names (i.e. alpha, beta);
- Query names in normal white indicate exact match - marked by "1" in the downloaded file;
- Query names highlighted indicate **no exact or unique match** - marked by "0" in the downloaded file;
- For **compound name**, you should click the **View** link to perform **approximate search** and manually select the correct match if found;
- For **KEGG ID**, it is possible to have multiple hits, you should click the **View** link to manually select the correct match if found;

Query	Hit	HMDB	PubChem	KEGG	Details
Pyruvate	Pyruvic acid	HMDB0000243	1060	C00022	
Lactic acid	L-Lactic acid	HMDB0000190	61503	C00186	
Oxalic acid	Oxalic acid	HMDB0002329	971	C00209	
Serine	Serine	HMDB0062263	5951	C00716	
Threonine	L-Threonine	HMDB000167	6288	C00188	
Phosphate	Phosphate	HMDB0001429	57424078	C00009	
Glycine	Glycine	HMDB0000123	750	C00037	
Pyroglutamic acid	Pyroglutamic acid	HMDB0000267	7405	C01879	
Proline	L-Proline	HMDB0000162	145742	C00148	
Erythrose	Erythrose	HMDB0002649	546072	C01796	
Hydroxycinnamic acid	4-Hydroxycinnamic acid	HMDB0002035	637542	C00811	

Slika 15.: Provjera prepoznatih metabolita

U drugom koraku bira se baza podataka metaboličkih puteva. Korištena je SMPDB baza podataka.

Parameter Setting

Unlike transcriptomics which allows comprehensive gene expression profiling, targeted metabolomics usually covers only a small percentage of metabolome (the actual coverage is platform/protocol specific). This means that metabolites (defined in our current pathways or metabolite sets) do not have equal probabilities of being measured in your studies, and the **enriched functions are the results from both platform/protocol specific effects and biological perturbations**. Since the primary interest is to detect the latter, we highly recommend **uploading a reference metabolome** containing all measurable metabolites from your platform to eliminate the former effects.

Please select a metabolite set library

Pathway based	<input checked="" type="radio"/> SMPDB	99 metabolite sets based on normal human metabolic pathways.
	<input type="radio"/> KEGG	84 metabolite sets based on KEGG human metabolic pathways (Oct. 2019).
	<input type="radio"/> Drug related	461 metabolite sets based on drug pathways from SMPDB.
Disease signatures	<input type="radio"/> Blood	344 metabolite sets reported in human blood.
	<input type="radio"/> Urine	384 metabolite sets reported in human urine.
	<input type="radio"/> CSF	166 metabolite sets reported in human cerebral spinal fluid (CSF).
	<input type="radio"/> Feces	44 metabolite sets reported in human feces.
Chemical structures	<input type="radio"/> Super-class	35 super chemical class metabolite sets or lipid sets
	<input type="radio"/> Main-class	464 main chemical class metabolite sets or lipid sets
	<input type="radio"/> Sub-class	1072 sub chemical class metabolite sets or lipid sets
Other types	<input type="radio"/> SNPs	4,598 metabolite sets based on their associations with SNPs loci.
	<input type="radio"/> Predicted	912 metabolite sets predicted to change in the case of dysfunctional enzymes.
	<input type="radio"/> Locations	73 metabolite sets based on organ, tissue, and subcellular localizations.

Slika 16.: Odabir Baze podataka za analizu metaboličkih puteva

3. REZULTATI I RASPRAVA

3.1. Identifikacija spojeva

Metaboličkim profiliranjem pomoću GC-MS analize identificiran je 21 spoj, što je prikazano u Tablici 3. Promatrani su samo spojevi navedeni u Bazi podataka ljudskog metaboloma (HMDB 4.0). Podaci kojima nedostaje više od 50% vrijednosti su isključeni iz analize. Cilj je bio odrediti učinak tieno[2,3-b]piridina na svaki metabolit i pronaći metabolite koji su značajno izmijenjeni kod tretiranih u odnosu na kontrolne stanice.

Tablica 3.: Lista identificiranih spojeva u MDA-MB-231 i MCF-7 staničnim linijama

Br.	RT (min)	Metabolit	MDA-MB-231		MCF-7	
			P vrij.	Fold change	P vrij.	Fold change
1	6.1	Pirogroždana kiselina	0.04*	2.19	0.15	0.76
2	6.3	L-Mliječna kiselina	<0.001*	0.53	0.03*	1.50
3	7.5	Oksalna kiselina	0.8	1.09	0.1	0.56
4	9.2	Serin				
5	9.4	Treonin				
6	9.5	Fosfat	0.03*	0.71	0.5	0.91
7	9.9	Glicin	0.28	0.44	0.3	0.63
8	11.5	D-Alanin				
9	12.4	Piroglutaminska acid				
10	12.8	L-Prolin	0.1	0.75	0.76	0.89
11	16.3	Eritroza	0.1	0.69		
12	16.4	4-Hidroksicinaminska kiselina	0.08	1.22	0.1	1.09
13	16.9	D-Fruktoza	0.3	0.52		
14	17.2	D-Glukoza	0.5	0.96	0.4	0.88
15	17.4	Manoza	0.9	1.07	0.7	0.96
16	18.5	Palmitinska kiselina	0.6	0.8	0.4	0.8
17	19.5	Mio-inozitol	0.03*	0.71	0.3	0.76
18	19.9	Oktadekanol	0.4	0.99	0.4	1.07
19	21.0	Stearinska kiselina	0.3	0.71	0.5	1.42
20	25.8	Gliceril monopalmitat	0.05*	0.73	0.9	0.99
21	28.5	Gliceril monostearat	0.04*	0.74	0.02*	1.2

Test statističke značajnosti pokazuje da su rezultati za MDA-MB-231 stanice značajniji od rezultata za MCF-7 stanice.

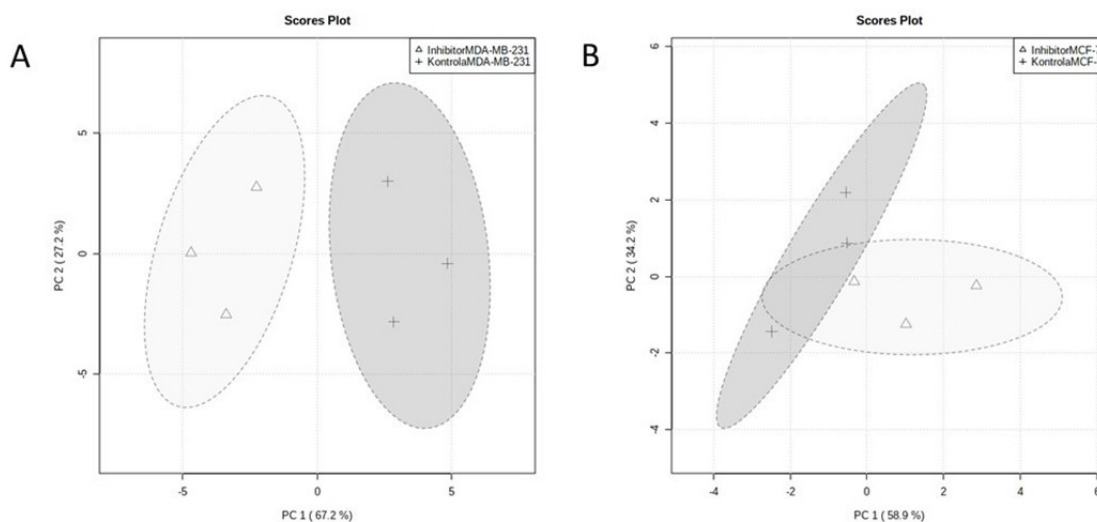
U MDA-MB 231 stanicama šest metabolita pokazalo je značajne razlike u odnosu na kontrolne stanice ($p < 0,05$), dok je u MCF-7 stanicama razlika značajna za samo dva metabolita.

3.2. ANOVA

Jednosmjerna analiza varijance (ANOVA) je korištena za određivanje statističke značajnosti među stanicama MDA-MB-231 i MCF-7. Piruvat, mio-inozitol i laktat označeni su kao značajni metaboliti korištenjem zadanog praga ($p < 0,05$) za jednosmjernu ANOVA i post hoc Tukeyjev test.

3.3. Analiza glavnih komponenti - PCA

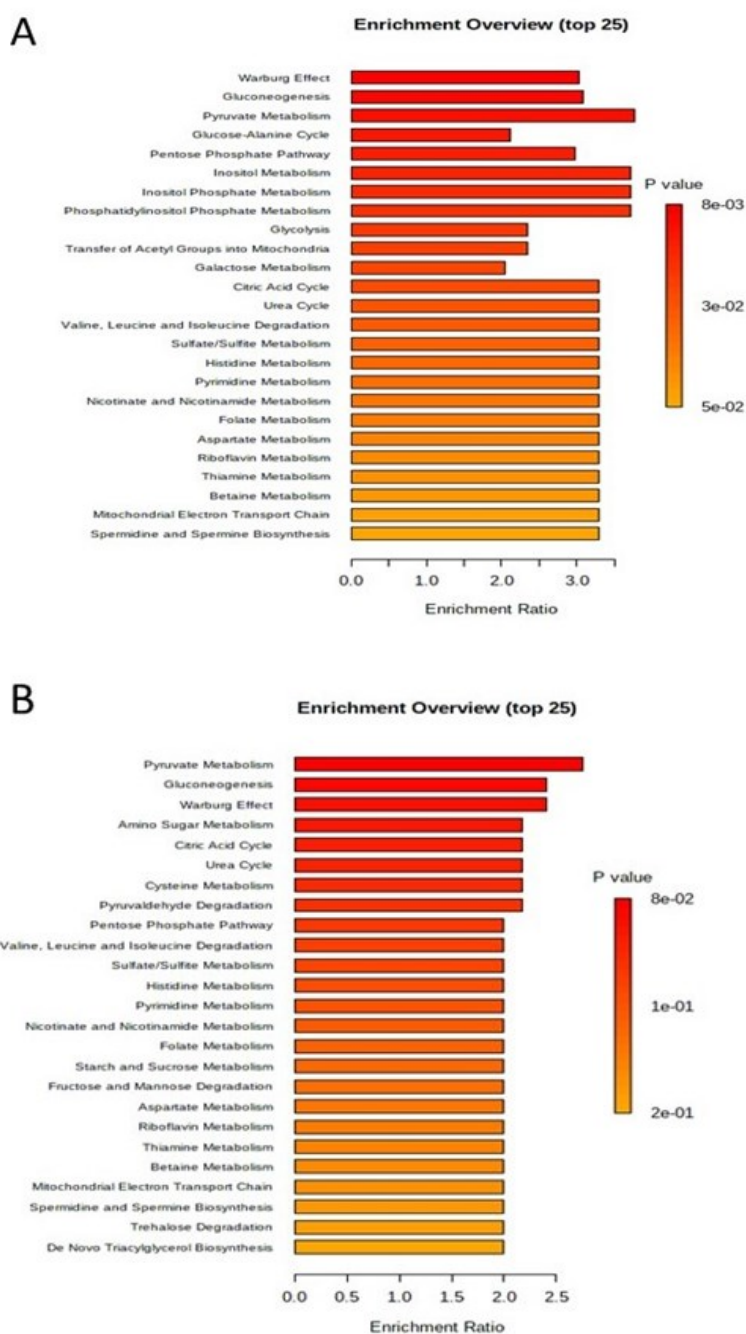
Rezultati PCA su pokazali da tretman tieno[2,3-b]piridinom uzrokuje metaboličke promjene u obje stanične linije. PCA dijagrami jasno pokazuju grupiranje tretiranih skupina u odnosu na kontrolne. PC1 čini više od 50% varijacije u podacima u obje linije.



Slika 17.: Analiz glavnih komponenti metaboličkih profila MDA-MB231 (A) i MCF- 7 (B) staničnih linija nakon tretiranja s tieno[2,3-b]piridinom u trajanju od 48h

3.4. Analiza kvantitativnog obogaćivanja

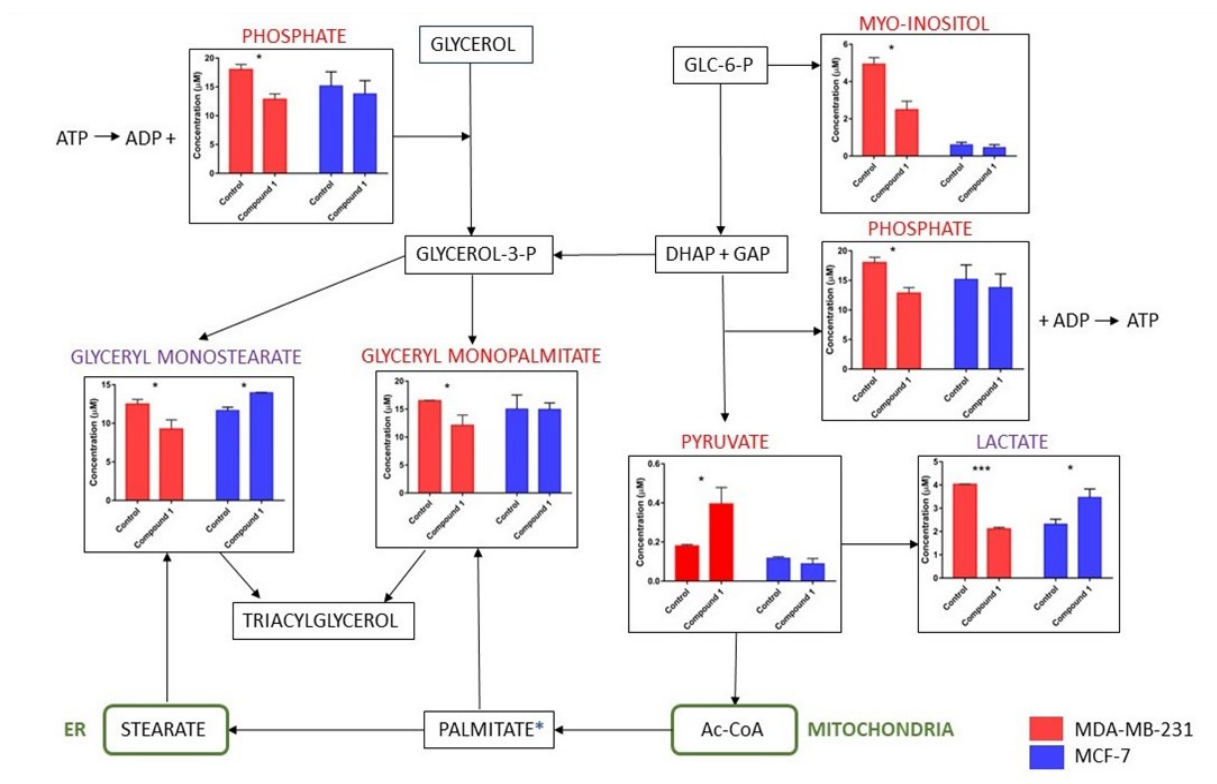
Analiza kvantitativnog obogaćivanja korištena je za identifikaciju obrazaca koncentracija metabolita i uvid u moguće biološke mehanizme. Rezultati se smatraju statistički značajnim u koliko je $p < 0,05$, taj statistički značaj nije utvrđen za MCF-7 stanice. Rezultati analize prikazani su na Slici 18.



Slika 18.: Analiza obogaćivanja seta metabolita MDA-MB231(A) i MCF-7(B) stanica nakon tretiranja s tieno[2,3-b]piridinom u trajanju od 48h

Rezultati analize pokazuju da tretman stanica tieno[2,3-b]piridinom ima veliki učinak na metabolizam glukoze odnosno energetske cikluse u stanici, a posebno na glikolizu, glukoneogenezu, metabolizam piruvata, pojavu Warburgova efekta i metabolizam inozitola.

Rezultati prikazani na Slici 19. pokazuju različito djelovanje spoja na dvije različite stanične linije raka. Tieno[2,3-b]piridin povećava koncentraciju piruvata, a smanjuje razine laktata, fosfata, mio-inozitola, gliceril monopalmitat i gliceril monostearata u MDA-MB 231 stanicama. Kod MCF-7 stanica uočen je obratan učinak za laktat i glicerilmonostearat, čije su koncentracije kod tretiranih stanica povećane u odnosu na kontrolne stanice.



Slika 19.: Relativne koncentracije metaboliti uključenih u glikolitički i triacilglicerolni put u MDA-MB-231 stanice (crveno) i MCF-7 stanice (plavo) nakon tretmana sa spojem 1 u trajanju od 48 h.

Podaci su izraženi kao srednja vrijednost \pm σ tri neovisna eksperimenta. * p < 0,05; *** p < 0,001; ljubičasti tekst označava koncentracije metabolita koji su promijenjeni nakon tretmana u obje stanične linije

Tieno[2,3-b]piridin je pokazao veću citotoksičnost za MDA-MB-231 stanice u odnosu na MCF-7 staničnu liniju. Povećana koncentracija piruvata, a smanjenje koncentracije laktata kod MDA-MB-231 stanica ukazuje na pomak prema aerobnoj glikolizi. Kod MCF-7 stanica nasuprot tome, primijećen je suprotan učinak. Povećanje koncentracije laktata, a smanjenje piruvata ukazuje na pomak prema anaerobnoj glikolizi.

4. ZAKLJUČAK

U ovom radu istraživana je utjecaj novosintetiziranog tieno[2,3-b]piridinskog spoja na metabolizam dviju staničnih linija raka dojke MCF-7 i MDA-MB-231.

Testom statističke značajnosti utvrđeno je da su rezultati za MDA-MB-231 stanice značajniji od rezultata za MCF-7 stanice.

Iz rezultata PCA analize jasno se vide grupiranja tretiranih u odnosu na kontrolne stanice, iz čega je vidljivo da se metaboličke promjene odvijaju u obje stanične linije.

Kvantitativnom analizom obogaćivanja nije utvrđen statistički značaj za MCF-7 stanice, pa za razliku od MDA-MB-231 stanica te podatke ne možemo smatrati statistički relevantnim.

Tieno[2,3-b]piridin povećava koncentraciju piruvata, a smanjuje razine laktata, fosfata, mio-inozitola, gliceril monopalmitata i gliceril monostearata u MDA-MB 231 stanicama. Kod MCF-7 stanica uočen je suprotan učinak za laktat i gliceril-monostearat, čije su koncentracije kod tretiranih stanica povećane u odnosu na kontrolne stanice.

Kod obiju stanica tieno[2,3-b]piridin ima veliki učinak na energetske cikluse u stanici. Najveći utjecaj utvrđen je za metaboličke putove: glikolizu, glukoneogenezu, metabolizam piruvata, pojavu Warburgova efekta i metabolizam inozitola.

Iz svega prikazanog jasno se može zaključiti da tieno[2,3-b]piridin ima veću citotoksičnost za MDA-MB-231 u odnosu na MCF-7 staničnu liniju. Povećana koncentracija piruvata, a smanjenje koncentracije laktata kod MDA-MB-231 stanica ukazuje na pomak prema aerobnoj glikolizi. Kod MCF-7 stanica nasuprot tome, primijećen je suprotan učinak. Povećanje koncentracije laktata, a smanjenje piruvata ukazuje na pomak prema anaerobnoj glikolizi.

LITERATURA

- [1] Liu, X. i Locasale, J. W., »Metabolomics: A Primer,« *Trends in Biochemical Sciences*, br. 42(4), p. 274–284, 2017.
- [2] Lay Jr, J. O., Liyanage, R., Borgmann, S. i Wilkins, C. L., »Problems with the “omics.”,« *TrAC Trends in Analytical Chemistry*, br. 25(11), pp. 1046-1056, 2006.
- [3] Idle, J. R. i Gonzalez, F. J., »Metabolomics,« *Cell Metabolism*, br. 6(5), p. 348–351, 2007.
- [4] Dettmer, K., Aronov, P. A. i Hammock, B. D., »Mass spectrometry-based metabolomics,« *Mass Spectrometry Reviews*, br. 26(1), p. 51–78, 2006.
- [5] Dettmer, K. i Hammock, B. D., »Metabolomics-a new exciting field within the "omics" sciences,« *Environmental Health Perspective*, br. 112(7), pp. A396-A397, 2004.
- [6] Zeki, Ö. C., Eylem, C. C., Reçber, T., Kır, S. i Nemitlu, E., »Integration of GC-MS and LC-MS For Untargeted Metabolomics Profiling,« *Journal of Pharmaceutical and Biomedical Analysis*, br. 190, p. 113509, 2020.
- [7] Zhang, A., Sun, H., Wang, P., Han, Y., i Wang, X., »Modern analytical techniques in metabolomics analysis,« *The Analyst*, br. 137(2), p. 293–300, 2012.
- [8] Shulaev, V., »Metabolomics technology and bioinformatics,« *Briefings in Bioinformatics*, br. 7(2), p. 128–139, 2006.
- [9] Ren, S., Hinzman, A. A., Kang, E. L., Szczesniak, R. D i Lu, L. J., »Computational and statistical analysis of metabolomics data,« *Metabolomics*, br. 11(6), p. 1492–1513, 2015.
- [10] Yi, L., Dong, N., Yun, Y., Deng, B., Ren, D., Liu, S. i Liang, Y., »Chemometric methods in data processing of mass spectrometry-based metabolomics: A review,« br. 914, p. 17–34, 2016.
- [11] Colby, B. N., »Spectral deconvolution for overlapping GC/MS components,« *Journal of the American Society for Mass Spectrometry*, br. 3(5), p. 558–562, 1992.

- [12] Blekherman, G., Laubenbacher, R., Cortes, D. F., Mendes, P., Torti, F. M., Akman, S., ... i Shulaev, V., »Bioinformatics tools for cancer metabolomics,« *Metabolomics*, br. 7(3), pp. 329-343, 2011.
- [13] Katajamaa, M. i Orešič, M., »Data processing for mass spectrometry-based metabolomics,« *Journal of chromatography A*, br. 1158(1-2), pp. 318-328, 2007.
- [14] Feizi, N., Hashemi-Nasab, F. S., Golpelihi, F., Saburouh, N. i Parastar, H. »Recent trends in application of chemometric methods for GC-MS and GC×GC-MS-based metabolomic studies,« *TrAC Trends in Analytical Chemistry*, br. 138, p. 116239, 2021.
- [15] Wilson, P. B. i Grootveld, M., *Computational Techniques for Analytical Chemistry and Bioanalysis* (vol. 20), Royal Society of Chemistry, 2020.
- [16] Tusher, V. G., Tibshirani, R. i Chu, G., »Significance analysis of microarrays applied to the ionizing radiation response,« *Proceedings of the National Academy of Sciences*, br. 98(9), pp. 5116-5121, 2001.
- [17] M. J. de Smith, *Statistical analysis handbook*, The Winchelsea Press, 2018.
- [18] Paul, A. i de Boves Harrington, P., »Chemometric applications in metabolomic studies using chromatography-mass spectrometry,« br. 135:116165, 2021.
- [19] Worley, B. i Powers, R., »Multivariate Analysis in Metabolomics,« *Current Metabolomics*, br. 1(1), p. 92–107, 2012.
- [20] Geladi, P., & Kowalski, B. R., »Partial least-squares regression: a tutorial,« *Analytica chimica acta*, br. 185, pp. 1-17, 1986.
- [21] Chen, Y., Li, E. M. i Xu, L. Y., »Guide to Metabolomics Analysis: A Bioinformatics Workflow,« *Metabolites*, br. 12(4), p. 357, 2022.
- [22] Cambiaghi, A., Ferrario, M. i Masseroli, M., »Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration,« *Briefings in bioinformatics*, br. 18(3), pp. 498-510, 2017.
- [23] Garg, N., Kaponov, C. A., Lim, Y. W., Koyama, N., Vermeij, M. J., Conrad, D., ... i Dorrestein, P. C., »Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures,« *International journal of mass spectrometry*, br. 377, pp. 719-727, 2015.

- [24] Johnson, S. R i Lange, B. M., »Open-access metabolomics databases for natural product research: present capabilities and future potential,« *Frontiers in bioengineering and biotechnology*, br. 3, p. 22, 2015.
- [25] Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. i Morishima, K., »KEGG: new perspectives on genomes, pathways, diseases and drugs,« *Nucleic acids research*, pp. D353-D361, 2017.
- [26] Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., ... i Scalbert, A., »HMDB 4.0: the human metabolome database for 2018,« *Nucleic acids research*, br. 46(D1), pp. D608-D617, 2018.
- [27] [Mrežno]. <https://www.metabolomicsworkbench.org/about/index.php>. [Pokušaj pristupa 29. 09. 2022.].
- [28] [Mrežno]. <https://www.metaboanalyst.ca/>. [Pokušaj pristupa 10. 09. 2022.].
- [29] Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N. i Xia, J., »Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data,« *Nature Protocols*, pp. 1-27, 2022.
- [30] Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., ... i Xia, J., »MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights,« *Nucleic Acids Research*, br. 49(W1), pp. w388-w396, 2021.
- [31] [Mrežno]. <https://www.metaboanalyst.ca/docs/About.xhtml>. [Pokušaj pristupa 12. 09. 2022.].
- [32] [Mrežno]. <https://www.metaboanalyst.ca/docs/RTutorial.xhtml>. [Pokušaj pristupa 12. 09. 2022.].
- [33] [Mrežno]. <https://www.metaboanalyst.ca/docs/Format.xhtml>. [Pokušaj pristupa 11. 09. 2022.].
- [34] Faraggi, D., Reiser, B., i Schisterman, E. F., »ROC curve analysis for biomarkers based on pooled assessments,« *Statistics in Medicine*, br. 22(15) , pp. 2515-2527, 2003.
- [35] Xia, J. i Wishart, D. S., »MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data,« *Nucleic acids research*, br. 38(2), pp. w71-w77, 2010.

- [36] J. a. W. D. Xia, »MetPA: a web-based metabolomics tool for pathway analysis and visualization,« *Bioinformatics*, br. 26, pp. 2342-2344, 2010.
- [37] Xia, J. i Wishart, D. S., »Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst,« *Nature Protocols*, br. 6(6), pp. 743-760, 2011.
- [38] Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., ... i Xia, J., »MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis,« *Nucleic acids research*, br. 46(1), pp. 486-494., 2018.
- [39] Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., ... i Pulendran, B., »Predicting network activity from high throughput metabolomics,« *PLoS computational biology*, br. 9(7), p. e1003123, 2013.
- [40] [Mrežno]. <https://www.metaboanalyst.ca/MetaboAnalyst/ModuleView.xhtml>. [Pokušaj pristupa 25. 09. 2022.].
- [41] [Mrežno]. <https://www.lqa.com/instrument/agilent-7000d-triple-quad-gc-ms-instrument/>. [Pokušaj pristupa 03. 10. 2022.].