

Analiza rezultata sekvenciranja i povezanost konzistencije i učestalosti stolice sa sastavom crijevne mikrobiote

Eranović, Ivana

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Food Technology and Biotechnology / Sveučilište u Zagrebu, Prehrambeno-biotehnološki fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:159:137052>

Rights / Prava: [Attribution-NonCommercial-NoDerivs 3.0 Unported / Imenovanje-Nekomercijalno-Bez prerada 3.0](#)

Download date / Datum preuzimanja: **2024-11-06**



Repository / Repozitorij:

[Repository of the Faculty of Food Technology and Biotechnology](#)



SVEUČILIŠTE U ZAGREBU
PREHRAMBENO-BIOTEHNOLOŠKI FAKULTET

DIPLOMSKI RAD

Zagreb, srpanj 2023.

Ivana Eranović

ANALIZA REZULTATA
SEKVENCIRANJA I POVEZANOST
KONZISTENCIJE I UČESTALOSTI
STOLICE SA SASTAVOM
CRIJEVNE MIKROBIOTE

Rad je izrađen u Laboratoriju za bioinformatiku na Zavodu za biokemijsko inženjerstvo Prehrambeno-biotehnološkog fakulteta Sveučilišta u Zagrebu pod mentorstvom izv. prof. dr. sc. Jurice Žučka.

Rad je izrađen u sklopu projekta pod nazivom „Pilot projekt crijevna mikrobiota radno - aktivne populacije RH“. Ovaj projekt rezultat je suradnje znanstvenika s Instituta za antropologiju pod vodstvom dr. sc. Luke Bočkora sa znanstvenicima Prehrambeno-biotehnološkog fakulteta Sveučilišta u Zagrebu pod vodstvom izv. prof. dr. sc. Jurice Žučka s ciljem određivanja sastava crijevne mikrobiote u radno-aktivnoj populaciji stanovništva RH na preliminarnom uzorku. Istraživanje je odobrilo Etičko povjerenstvo Instituta za antropologiju.

ZAHVALA

Iskrenu zahvalu, veću od cijelog metagenoma, dugujem svom mentoru, izv. prof. dr. sc. Jurici Žučku, na strpljenju i razumijevanju tijekom izrade ovog rada kao i na svim savjetima i konstruktivnim kritikama iz kojih sam puno naučila.

Veliko hvala i mojim divnim prijateljicama iz WhatsApp grupa Venus, Bakalari i Ferali što ste uvijek bile tu za mene, a osobito hvala Kiki na svakom satu provedenom u knjižnici.

I najveće hvala mojim roditeljima i sestri te ostatku obitelji na svojoj ljubavi i podršci koju su mi pružili, a mojoj najdražoj majci hvala što je uvijek vjerovala u mene!

TEMELJNA DOKUMENTACIJSKA KARTICA

Diplomski rad

Sveučilište u Zagrebu
Prehrambeno-biotehnološki fakultet
Zavod za biokemijsko inženjerstvo
Laboratorij za bioinformatiku

Znanstveno područje: Biotehničke znanosti
Znanstveno polje: Nutricionizam

Diplomski sveučilišni studij: Nutricionizam

ANALIZA REZULTATA SEKVENCIRANJA I POVEZANOST KONZISTENCIJE I UČESTALOSTI STOLICE SA SASTAVOM CRIJEVNE MIKROBIOTE

Ivana Eranović, univ. bacc. nutr. 0058211585

Sažetak: Crijevna mikrobiota vrlo je dinamičan sustav, a na njegov sastav značajno utječu različiti obrasci prehrane, ali i mnogi drugi čimbenici. Zanimljivo je utjecaja pojedinih čimbenika na sastav mikrobiote može značajno utjecati na interpretaciju rezultata dobivenih istraživanjima. Učestalost i konzistencija stolice su mjere intestinalnog tranzitnog vremena – čimbenika koji determinira sadržaj vode i nutrijenata u crijevu te time oblikuje uvjete mikrobnog staništa. Cilj ovoga rada bio je povezati učestalost i konzistenciju stolice sa sastavom crijevne mikrobiote. Istraživanje je provedeno na 59 ispitanika od kojih su prikupljeni uzorci stolice iz kojih je određivan sastav mikrobiote sekvenciranjem 16S rRNA gena na Illumina MiSeq uređaju, a informacije o učestalosti i konzistenciji stolice dobiveni su ispunjavanjem upitnika. Za analizu podataka dobivenih sekvenciranjem korišten je program QIIME2 te DADA2 metoda pročišćavanja podataka. Prema rezultatima ove analize nije uočena značajna povezanost sastava crijevne mikrobiote s učestalosti i/ili konzistencijom stolice na promatranoj skupini ispitanika.

Ključne riječi: *mikrobiota, stolica, QIIME2, varijante sekvence amplikona (ASV)*

Rad sadrži: 50 stranica, 14 slika, 5 tablica, 46 literaturnih navoda, 1 prilog

Jezik izvornika: hrvatski

Rad je u tiskanom i elektroničkom (pdf format) obliku pohranjen u: Knjižnica Prehrambeno-biotehnološkog fakulteta, Kačićeva 23, Zagreb

Mentor: izv. prof. dr. sc. Jurica Žučko

Stručno povjerenstvo za ocjenu i obranu:

1. izv. prof. dr. sc. Ivana Rumora Samarin (predsjednica)
2. izv. prof. dr. sc. Jurica Žučko (mentor)
3. izv. prof. dr. sc. Andreja Leboš Pavunc (članica)
4. prof. dr. sc. Igor Slivac (zamjenski član)

Datum obrane: 17. srpnja 2023.

BASIC DOCUMENTATION CARD

Graduate Thesis

University of Zagreb
Faculty of Food Technology and Biotechnology
Department of Biochemical Engineering
Laboratory for Bioinformatics

Scientific area: Biotechnical Sciences

Scientific field: Nutrition

Graduate university study programme: Nutrition

ANALYSIS OF SEQUENCING RESULTS AND ASSOCIATION BETWEEN STOOL CONSISTENCY AND FREQUENCY AND GUT MICROBIOTA COMPOSITION

Ivana Eranović, univ. bacc. nutr. 0058211585

Abstract: Gut microbiota is a very dynamic system, and its composition is strongly affected by individual's diet pattern, as well as numerous other factors. Disregarding the impact of these factors on microbiota composition can significantly affect the interpretation of results obtained in the study. Stool consistency and frequency are used as proxies for intestinal transit time – a factor determining water and nutrient content in the intestine and, consequently, shaping the conditions of the microbial habitat. The aim of this study was to observe connection between stool consistency and frequency and gut microbiota composition. Study was conducted on 59 subjects from whom stool samples were collected and analysed by sequencing 16S rRNA gene on Illumina MiSeq platform to determine microbiota composition. Data considering stool consistency and frequency was collected via online questionnaire. QIIME2 pipeline was used for analysis of sequencing data including DADA2 method for denoising step. The results of this study showed no significant association between stool consistency and/or frequency and gut microbiota composition in the observed group of subjects.

Keywords: *microbiota, stool, QIIME2, amplicon sequence variants (ASV)*

Thesis contains: 50 pages, 14 figures, 5 tables, 46 references, 1 supplement

Original in: Croatian

Graduate Thesis in printed and electronic (pdf format) form is deposited in: The Library of the Faculty of Food Technology and Biotechnology, Kačićeva 23, Zagreb

Mentor: Jurica Žučko, PhD, Associate professor

Reviewers:

1. Ivana Rumora Samarin, PhD, Associate professor (president)
2. Jurica Žučko, PhD Associate professor (mentor)
3. Andreja Leboš Pavunc, PhD, Associate professor (member)
4. Igor Slivac, PhD, Full professor (substitute)

Thesis defended: July 17th, 2023

Sadržaj

1. UVOD.....	1
2. TEORIJSKI DIO	2
2.1. CRIJEVNA MIKROBIOTA	2
2.1.1. Sastav crijevne mikrobiote i mikrobioma	2
2.1.2. Uloga crijevne mikrobiote	4
2.1.3. Povezanost sastava crijevne mikrobiote s učestalosti i konzistencijom stolice . . 5	
2.2. ANALIZA SASTAVA CRIJEVNE MIKROBIOTE	7
2.2.1. Metode za određivanje sastava crijevne mikrobiote	7
2.2.2. Sekvenciranje	10
2.2.3. Bioinformatička analiza	12
3. EKSPERIMENTALNI DIO	16
3.1. ISPITANICI.....	16
3.2. METODE RADA.....	16
3.2.1. Dijetetičke metode	16
3.2.2. Genomičke metode	17
3.2.3. Bioinformatičke metode	17
4. REZULTATI I RASPRAVA	25
4.1. ISPITANICI.....	26
4.2. REZULTATI PROCJENE KVALITETE I PROČIŠĆAVANJA	28
4.2.1. Metoda spojenih krajeva	28
4.2.2. Metoda jednog kraja	31
4.3. TAKSONOMSKA ZASTUPLJENOST MIKROORGANIZAMA PO GRUPAMA	34
4.4. ALFA RAZNOLIKOST	37
4.5. BETA RAZNOLIKOST	41
5. ZAKLJUČCI.....	44
6. LITERATURA.....	45

PRILOZI:

PRILOG 1. UPITNIK PILOT PROJEKTA „CRIJEVNA MIKROBIOTA RADNO-AKTIVNE POPULACIJE RH“

1. UVOD

U proteklom je desetljeću postalo jasno kako crijevni mikrobiom ima značajnu ulogu kako u održavanju homeostaze domaćina tako i u patofiziološkim procesima specifičnima za razne bolesti. Rapidan razvoj modernih metoda sekvenciranja i računalnih alata za analizu nedvojbeno su u najvećoj mjeri doprinijeli da danas imamo širok spektar spoznaja o sastavu i ulogama crijevnih mikrobiota. Međutim, u analizama mikrobioma još postoje mnoga ograničenja koja mogu značajno utjecati na zaključke provedenih studija i na daljnje širenje spoznaja o prirodi crijevnog mikrobioma. Osim daljnjeg razvoja metoda sekvenciranja, potrebno je i dalje usavršavati i razvijati računalne programe za obradu podataka kao i statističke modele (Galloway-Peña i Hanson, 2020).

Također, kod provedbe studija o sastavu crijevnih mikrobiota posebnu pažnju je potrebno posvetiti samom dizajnu studije i uzeti u obzir sve potencijalne čimbenike koji mogu utjecati na sastav crijevnih mikrobiota, a kojih ima mnogo. Neki od njih su dob, indeks tjelesne mase, način poroda, različiti obrasci prehrane, korištenje lijekova, pušenje, konzumacija alkohola, tjelesna aktivnost i konzistencija stolice (Schmidt i sur., 2018).

Provedena su istraživanja koja pokazuju kako je sastav crijevnih mikrobiota povezan s konzistencijom stolice (Vork i sur., 2021; Vandeputte i sur., 2016; Tigchelaar i sur., 2015;) i s učestalosti stolice (Kwon i sur., 2019; Hadizadeh i sur., 2017). Falony i sur. (2016) usporedbom nekolicine čimbenika zapažaju kako upravo konzistencija stolice najviše doprinosi detektiranoj varijaciji u sastavu mikrobiota.

Cilj ovoga rada bio je istražiti postoji li značajna povezanost sastava crijevnih mikrobiota s učestalosti i konzistencijom stolice, odnosno trebaju li se učestalost i/ili konzistencija stolice smatrati značajnim zbunjujućim varijablama (engl. *confounding variable*) u studijama sastava crijevnih mikrobiota. S time na umu, korištenjem programa QIIME2, provedena je bioinformatička analiza podataka dobivenih sekvenciranjem 16S rRNA gena iz uzoraka stolice, a rezultati su korelirani s odgovorima ispitanika iz upitnika o prehrambenim i životnim navikama.

2. TEORIJSKI DIO

2.1. CRIJEVNA MIKROBIOTA

2.1.1. Sastav crijevne mikrobiote i mikrobioma

Termin mikrobiota podrazumijeva skup svih mikroorganizama koji koloniziraju određeni ekološki sustav (npr. koža, usna šupljina, crijeva), a mikrobiom predstavlja cjelokupnu genetičku informaciju svih prisutnih mikroorganizama koji sačinjavaju mikrobiotu. U ljudskom tijelu, debelo je crijevo nastanjeno najgušćom i najraznolikijom mikrobnom zajednicom. Procjenjuje se kako ga nastanjuje oko 10^{12} mikrobnih stanica. Prema trenutnim spoznajama najveći udio čini domena bakterija, a predstavnici arheja i eukariotski mikroorganizmi manje su zastupljeni, dok su virusi još uvijek neistražena domena (Pargin i sur., 2023).

Promatrajući taksonomske razine, najzastupljenija koljena bakterija su *Bacteroidetes* i *Firmicutes*, slijede *Proteobacteria*, *Actinobacteria*, *Fusobacteria*, *Terenicutes* i *Verrucomicrobia*. Navedena koljena čine do 90 % ukupne mikrobne zajednice. Najzastupljeniji rodovi su *Bacteroides*, *Clostridium*, *Peptococcus*, *Bifidobacterium*, *Eubacterium*, *Ruminococcus*, *Faecalibacterium* te *Peptostreptococcus*, ali je njihov udio vrlo varijabilan. Na razini vrste i soja dolazi do još izraženije varijabilnosti u sastavu mikrobiote među pojedincima (Gomaa, 2020; Alonso i Guarner, 2013).

Bogatstvo mikrobioma kolona, prema Human Microbiome Project-u, procjenjuje se na više od 5 milijuna neredundantnih gena koji kodiraju za grupe proteina uključene u oko 20 000 bioloških funkcija. To je broj ukupno otkrivenih gena, a zaključak MetaHIT projekta je kako na individualnoj razini prosjek gena u gastrointestinalnom traktu iznosi oko 536 000, što bi odgovaralo taksonomskoj raznolikosti od barem 160 različitih bakterijskih vrsta (The Human Microbiome Project Consortium, 2012; Qin i sur., 2010).

2.1.1.1. Optimalni sastav mikrobiote

Trenutno ne postoji znanstveni konsenzus o optimalnom sastavu crijevne mikrobiote u zdravih ljudi. Novija istraživanja zaključuju kako taksonomski vrlo različiti sastavi crijevne mikrobiote ispitanika imaju slične metaboličke aktivnosti, odnosno sličan profil prisutnih gena. Dakle, raznolikost sastava mikrobiote ne podrazumijeva nužno i raznolikost u sastavu mikrobioma (Frame i sur., 2020). U prilog tome govore studije dvije obitelji istog bakterijskog koljena,

Lachnospiraceae i *Ruminococcaceae*, koje obje imaju ulogu u probavi vlakana i sintezi butirata, a njihova zastupljenost međusobno negativno korelira. Sličan je odnos zabilježen i kod 2 roda korisnih bakterija *Bifidobacterium* i *Faecalibacterium* koje također pokazuju trend međusobne negativne korelacije. Stoga se pretpostavlja kako fokus budućih istraživanja treba usmjeriti ka funkcionalnim metagenomičkim studijama kako bi se bolje spoznale same aktivnosti crijevne mikrobiote, a ne samo njen sastav (Manor i sur., 2020). Manor i sur. (2020) zaključuju kako vjerojatno ne postoji jedinstven sastav zdrave mikrobiote, nego spektar različitih sastava koji su definirani skupovima korisnih bakterija i koji je karakterističan za svakog pojedinca. Iako je točan sastav zdrave mikrobiote možda i nemoguće odrediti, opće je prihvaćeno kako zdravu crijevnu mikrobiotu karakterizira bogatstvo i raznolikost prisutnih vrsta dok se disbioza u sastavu crijevne mikrobiote može definirati kao promjena sastava koja je uzrok ili posljedica nekog patološkog stanja (Rinninella i sur., 2019).

2.1.1.2. Heterogenost sastava mikrobiote

Promatrajući sastav mikrobne zajednice kroz različite dijelove probavne cijevi mogu se uočiti značajne razlike i zaključiti kako je sastav mikrobiote kroz cijeli probavni sustav vrlo heterogen. Razlog tomu su različiti uvjeti koji prevladavaju u različitim dijelovima probavne cijevi. Tako na primjer tanko crijevo ima niži pH i kraće tranzitno vrijeme, više razine kisika, antimikrobnih peptida i jednostavnih ugljikohidrata nego debelo crijevo. Sve navedeno doprinosi tome da je tanko crijevo kolonizirano fakultativnim anaerobima koji mogu tolerirati efekte žučnih kiselina i antimikrobnih peptida i koji kao izvor energije koriste jednostavne ugljikohidrate, kao na primjer *Proteobacteria* i *Lactobacillales*. S druge strane, debelo je crijevo kolonizirano anaerobnim bakterijama koje energiju dobivaju fermentacijom rezistentnih polisaharida (prehrambenih vlakana) i mukusa. Zbog opisanih uvjeta koji prevladavaju u tankom crijevu, gustoća i raznolikost bakterijskih vrsta koje ga nastanjuju značajno je manja nego u debelome crijevu.

Osim razlika kroz dužinu intestinalnog trakta, važno je naglasiti uočene razlike u sastavu mikrobiote kada se gleda poprečni presjek crijeva (od središta lumena prema epitelu). Epitel kolona oblažu 2 sloja mukoze, unutarnji gusti sloj koji je vezan za epitel i vanjski rahli sloj (Donaldson i sur., 2016). U vanjskom, rahlijem mukoznom sloju povećana je gustoća bakterija koje fermentiraju mucin poput *Akkermansia muciniphila* i nekih vrsta iz roda *Bacteroides*. Bliže

epitelu, povećava se prisutnost aerotolerantnih koljena poput *Proteobacteria* i *Actinobacteriae* (Tropini i sur., 2017).

Određivanje sastava mikrobiote u uzorcima stolice dominantna je metoda u većini studija. Uzimanje uzoraka je relativno jednostavno, a sastav mikrobiote u stolici daje zadovoljavajuću aproksimaciju sastava mikrobiote u crijevu. Međutim, kako bi se bolje razumjela uloga mikrobiote u fiziološkim i patofiziološkim procesima pažnju treba usmjeriti i na sastav mikrobiote mukoznog sloja koji oblaže epitel debeloga crijeva. U usporedbi sa sastavom mikrobiote u lumenu crijeva, mukozni je sloj koloniziran manje raznolikom, ali jedinstvenom mikrobnom zajednicom. Uzimanjem uzoraka mukoznog sloja biopsijom na različitim dijelovima debeloga crijeva ustanovljeno je kako su gotovo sve identificirane vrste prisutne na svim lokacijama, ali u različitim omjerima. Usporedbom fekalne i mukozne mikrobiote zdravih ispitanika zamijećena je dobra korelacija među identificiranim rodovima, a iznimka je bila u rodu *Bacteroides* koji je bio zastupljeniji u uzorcima mukoznog sloja (Vaga i sur., 2020). Studije koje uključuju ispitanike s upalnim bolestima crijeva navode značajniju razliku u sastavu mikrobiote izolirane iz ova dva izvora (Čipčić Paljetak i sur., 2022; Parthasarathy i sur., 2016; Zoetendal i sur., 2002).

2.1.2. Uloga crijevne mikrobiote

Prisutnost i aktivnost određenih gena definira biološku funkciju mikrobiote (Qin i sur., 2010). Najbolje objašnjena uloga crijevne mikrobiote jest metabolizam nutrijenata. Mikrobiota vrši fermentaciju neprobavljenih ugljikohidrata i oligosaharida, a kao nusprodukt te fermentacije nastaju kratko-lančane masne kiseline poput butirata, propionata i acetata, koje su korisne za zdravlje domaćina. Mikrobiota također sudjeluje u metabolizmu aminokiselina na način da konvertira neke aminokiseline u male signalne molekule ili antimikrobne peptide. Primjeri takvih metaboličkih putova posredovanih djelovanjem mikrobiote su prevođenje *L*-histidina u histamin koji ima ulogu u imunosnom odgovoru te prevođenje glutamata u gama-aminomaslačnu kiselinu (GABA), glavni inhibicijski neurotransmiter u središnjem živčanom sustavu. Nadalje, mikrobiota sudjeluje u sintezi vitamina K, nekih komponenti vitamina B skupine te konjugirane linolne kiseline koja pozitivno djeluje na zdravlje, a osobito značajnu ulogu u regulaciji dijabetesa i lipidnog profila. Neke bakterije iz sastava mikrobiote sudjeluju u metaboličkoj transformaciji polifenola. Uklanjanjem šećera koji je vezan na polifenolnu komponentu, polifenoli se prevode u aktivnu formu koja ima antimikrobne, antioksidativne i

mnoge druge blagotvorne učinke za domaćina (Rowland i sur., 2018).

Osim uloge u metabolizmu nutrijenata, mikrobiota također sudjeluje u metabolizmu raznih ksenobiotika i lijekova. Nakon interakcije s mikrobiotom smanjuje se štetnost nekih ksenobiotika, a neki lijekovi pokazuju poboljšanu djelotvornost. Također, mikrobiota doprinosi zaštiti domaćina od infekcija patogenim mikroorganizmima kroz različite mehanizme djelovanja (Jandhyala i sur., 2015).

Od velikog je interesa istražiti i ulogu mikrobiote u različitim patološkim stanjima. Kod mnogih je bolesti primijećena disbioza u sastavu crijevne mikrobiote, iako se za mnoge ne može ustvrditi je li ta disbioza uzrok ili posljedica bolesti. Do promjena u sastavu mikrobiote dolazi kod neurodegenerativnih bolesti poput Parkinsonove i Alzheimerove bolesti, kod kardiovaskularnih bolesti, osobito kod hipertenzije i ateroskleroze, zatim kod metaboličkih bolesti poput pretilosti, dijabetesa, nealkoholne masne bolesti jetre te kod gastrointestinalnih bolesti kao što su upalne bolesti crijeva i kolorektalni karcinomi (Chen i sur., 2021).

2.1.3. Povezanost sastava crijevne mikrobiote s učestalosti i konzistencijom stolice

2.1.3.1. *Intestinalno tranzitno vrijeme*

Intestinalno tranzitno vrijeme parametar je koji se rijetko proučava te se često zanemaruje u studijama intestinalne mikrobiote iako dokazi ukazuju na to kako je upravo tranzitno vrijeme jedan od glavnih čimbenika koji utječe na oblikovanje sastava i aktivnosti crijevne mikrobiote. Tranzitno vrijeme determinira sadržaj nutrijenata i vode u intestinalnom traktu te samim time oblikuje uvjete u mikrobnom staništu. Može se određivati direktnom metodom, ingestijom specijalno dizajnirane kapsule, ali ta je metoda nepraktična i invazivna. Mnogo jednostavnije i praktičnije metode koje se koriste za aproksimaciju intestinalnog tranzitnog vremena su određivanje konzistencije stolice, frekvencije stolice te mjerenje sadržaja vode u stolici (Procházková i sur., 2023; Vandeputte i sur., 2016).

U zdravih je ispitanika izmjeren medijan intestinalnog tranzitnog vremena izravnom metodom. Tranzitno vrijeme tankoga crijeva iznosilo je 4:47 sati, a kolona 21:06 sati. Pritom je tranzitno vrijeme uzlaznog, transverznog, silaznog i rektosigmoidalnog dijela kolona iznosilo redom 5:41, 4:53, 1:54 i 2:18 sati. Medijan ukupnog tranzitnog vremena kroz tanko i debelo crijevo zajedno izmjeren je na oko 30 sati (Nandhra i sur., 2020). Normalni interval tranzitnog vremena tankog crijeva iznosi 2 – 6 sati, a kolona 10 – 59 sati, dok je interval ukupnog intestinalnog tranzitnog vremena 10 – 73 sata (Lee i sur., 2014). Kod pojedinih bolesti intestinalno tranzitno

vrijeme može se uvelike razlikovati. Do najvećih oscilacija tranzitnoga vremena dolazi upravo u kolonu te one mogu iznositi 0,1 – 46 sati za proksimalni dio kolona, 0,3 – 80 sati za distalni dio kolona te 1 – 134 sata za rektosigmoidalni dio kolona (Procházková i sur., 2023).

2.1.3.2. Učestalost i konzistencija stolice i sastav mikrobiote

Nekoliko je studija dokazalo statistički značajnu negativnu korelaciju između alfa raznolikosti mikrobiote i učestalosti stolice. Kod ispitanika koji su imali veći broj stolica dnevno raznolikost crijevne mikrobiote bila je manja. U grupama s manjim brojem stolica zabilježen je veći omjer *Bacteroidetes* : *Firmicutes* te niža prisutnost roda *Bifidobacterium*. Također, detektirane su porodice i rodovi bakterija karakteristični za grupe s najvećim te za grupe s najmanjim zabilježenim brojem stolica. Iako je broj studija koje povezuju učestalost stolice sa sastavom mikrobiote malen, svi dostupni radovi pokazuju korelaciju između dva ispitivana čimbenika, ali na temelju do sada dostupnih dokaza ne mogu se iznositi zaključci o uzrocima primijećene korelacije (Kwon i sur., 2019; Hadizadeh i sur., 2017).

Konzistencija stolice također se koristi kao marker intestinalnog tranzitnog vremena. Za razliku od frekvencije, konzistencija stolice mnogo je bolje istražen čimbenik. Za određivanje konzistencije stolice koristi se Bristolska skala stolice (engl. *Bristol Stool Scale*) koja je podijeljena na 6 tipova, tip 1 predstavlja najtvrdju, a tip 6 najmekšu stolicu. Studije provedene na ovu temu utvrđuju statistički značajnu negativnu korelaciju između konzistencije stolice i raznolikosti mikrobiote. Dakle, mekša stolica, koja je povezana s kraćim tranzitnim vremenom, povezana je s manjom raznolikosti mikrobiote. Primijećena je i nekolicina korelacija u brojnosti pojedinih taksonomskih razina s konzistencijom stolice. Neki od primjera su omjer *Bacteroidetes* : *Firmicutes* koji je pozitivno korelirao s konzistencijom stolice, a značajno veća prisutnost metanogene porodice *Methanobacteriaceae* zabilježena je kod tvrdih stolica (Vork i sur., 2021; Vandeputte i sur., 2016; Tigchelaar i sur., 2015).

Generalni zaključak studija provedenih na temu usporedbe sastava mikrobiote s frekvencijom i konzistencijom stolice jest kako bi se ove parametre, kao mjeru tranzitnog vremena, trebalo uključiti u buduće studije na način da se razmotri njihov utjecaj kao potencijalnih zbunjujućih varijabli (engl. *confounding variable*) jer je neosporno kako imaju ulogu u oblikovanju sastava crijevne mikrobiote, a vrlo su lako mjerljivi čimbenici.

2.2. ANALIZA SASTAVA CRIJEVNE MIKROBIOTE

2.2.1. Metode za određivanje sastava crijevne mikrobiote

Metode određivanja sastava i uloge mikrobiote mogu se podijeliti u 4 grupe: tradicionalni pristup koji se bazira na kultivaciji prisutnih mikroorganizama u uzorku, trenutno dominantna metoda sekvenciranja DNA bazirana na stvaranju amplikona, moderna metagenomska metoda koja podrazumijeva sekvenciranje cijelih genoma te metoda metatranskriptomike (Athanasopoulou i sur., 2023).

2.2.1.1. Kulturomika

Razvojem tehnologija, od najranijih metoda detekcije sastava mikrobiote koje podrazumijevaju kultivaciju i detekciju mikroorganizama na temelju njihovih biokemijskih osobina razvio se tzv. kulturomički pristup. Ova se metoda izvodi na način da se stanice mikroorganizama izoliraju iz uzorka, zatim kultiviraju te se potom vrši analiza sastava metodama Sanger sekvenciranja ili masene spektrometrije. Iako ju karakteriziraju brojni nedostaci uglavnom vezani za kompliciranost postupka kultivacije, ova metoda ima vrlo visoku osjetljivost u usporedbi s modernijim metodama te može detektirati veći broj različitih bakterijskih vrsta. Metode sekvenciranja amplikona i metagenomska metoda mogu propustiti detekciju rijetkih bakterija koje nisu zastupljene u referentnim bazama podataka ili koje su prisutne u malim količinama, odnosno ispod praga filtriranja koji se koristi za pročišćavanja generiranih podataka (Lagier i sur., 2012).

2.2.1.2. Sekvenciranje amplikona

Danas je najkorištenija metoda određivanja sastava mikrobiote sekvenciranje amplikona. Razlog tomu je razvoj tehnologija sekvenciranja nove generacije (engl. *Next Generation Sequencing*, NGS), osobito NGS platformi kao što su Illumina i Ion Torrent. Opći princip rada ove metode jest ciljano umnožavanje određene regije gena lančanom reakcijom polimeraze (engl. *Polymerase Chain Reaction*, PCR) te sekvenciranje dobivenih PCR amplikona. Temeljna ideja leži u tome da gen koji se umnaža i sekvencira ima svojstvo napraviti taksonomsku distinkciju među organizmima u uzorku. Taksonomski marker gen za bakterijske

stanice je 16S rRNA, za eukariote 18S rRNA gen, a za gljive se koristi ITS regija (engl. *Internal Transcribed Spacer*) – ne kodirajuća regija DNA smještena između gena koji kodiraju za malu i veliku ribosomsku podjedinicu (Callahan i sur., 2016).

Protokol rada obuhvaća sljedeće korake: prikupljanje uzoraka, ekstrakciju DNA iz uzoraka, PCR amplifikaciju ciljane regije DNA, stvaranje knjižnice, sekvenciranje amplikona te bioinformatičku analizu.

Odlike korištenja ove metode su dobar omjer cijene i učinkovitosti. Zahtjeva malo biomase iz uzorka jer se koristi metoda PCR umnažanja. NGS tehnologija omogućuje generiranje očitavanja s visokim protokom (engl. *high-throughput technology*) odnosno omogućuje istovremeno sekvenciranje više amplikona. Kako se ne sekvencira cijeli genom, količina generiranih podataka je relativno mala, a njihova bioinformatička analiza relativno jednostavna (Athanasopoulou i sur., 2023; Allaband i sur., 2019).

Glavni nedostaci metode su dobivanje lažno-pozitivnih ili lažno-negativnih rezultata do kojih dolazi zbog pogrešnih očitavanja u samom sekvenciranju i zbog koraka PCR umnažanja. Tijekom PCR umnažanja dolazi do stvaranja kimera – amplikona koji su nastali iz dvije različite sekvence. Iz navedenih razloga dolazi do generiranja „nonsense“ sekvenci koje mogu biti pogrešno klasificirane. Nadalje, taksonomska identifikacija bakterija ovisi o kvaliteti korištene referentne baze s kojom se očitavanja uspoređuju. Osjetljivost metode je, u najboljem slučaju, do razine vrste, ne dobivaju se podatci o prisutnim sojevima bakterija. Također, poznavanje vrste i relativne brojnosti bakterija u uzorku nije dovoljno da bi se na temelju tih saznanja donosili zaključci o ulozi pojedinih bakterija u patološkim zbivanjima ili u održavanju homeostaze (Jo i sur., 2016).

Marker gen 16S rRNA

16S rRNA je gen koji kodira za 16S podjedinicu ribosoma prokariota. Karakteristike tog gena su da je prisutan u svim bakterijama i da ima varijabilne regije na temelju kojih se može odrediti taksonomija. Neki od razloga za široko korištenje 16S rRNA gena u studijama crijevne mikrobiote jest nizak horizontalni prijenos gena i sposobnost razlikovanja vrlo srodnih organizama do razine vrste. Puna dužina gena iznosi oko 1500 bp što je predugačko za sekvenciranje na NGS platformama te se stoga razvio pristup korištenja hipervarijabilnih regija gena (V1 – V9) kao reprezentativnih za cijeli gen. U većini studija koristi se hipervarijabila

regija V4 dužine oko 250 bp jer njena dužina odgovara kapacitetima Illumina platforme. Konzervirane se regije koriste za kreiranje početnica, što je vrlo važan korak u analizi jer odabir početnica može značajno utjecati na rezultat. Generalna preporuka je koristiti iste početnice kao u studiji s kojom se dobiveni rezultati žele uspoređivati (Allaband i sur., 2019; de la Cuesta-Zuluaga i Escobar, 2016).

2.2.1.3. Metagenomika

Pojam metagenomike podrazumijeva sekvenciranje cijelog genoma svih organizama prisutnih u uzorku. Sekvenciranje metagenoma moguće je na NGS i na TGS (engl. *Third Generation Sequencing*) platformama. NGS platforme zahtijevaju fragmentaciju DNA sekvenci dok je na TGS platformama moguće sekvenciranje cijele DNA molekule. Ova metoda omogućuje detekciju svih prisutnih vrsta mikroorganizama u uzorku (ne samo bakterija kao 16S rRNA sekvenciranje), odnosno taksonomsko profiliranje uzorka. K tome, omogućuje identifikaciju funkcionalnih gena koji određuju metaboličku aktivnost mikrobiote, odnosno funkcionalno profiliranje uzorka. Povezivanjem taksonomskog i funkcionalnog profila mikrobne zajednice moguće je ponuditi odgovor na mnogo širi spektar bioloških pitanja. Prednosti metagenomske metode su širi opseg informacija dobivenih analizom, taksonomska rezolucija na razini soja, mala potrebna količina genetskog materijala iz uzorka mjerena u nanogramima. Iako i dalje značajno skuplja od metode sekvenciranja amplikona, cijena ove metode se smanjuje. Nedostatci su kontaminiranost uzorka s DNA domaćina te generiranje velike količine podataka što komplicira bioinformatičku analizu (Athanasopoulou i sur., 2023; Allaband i sur., 2019).

2.2.1.4. Metatranskriptomika

Metatranskriptomika je metoda sekvenciranja cijelog genskog transkripta, odnosno svih mRNA molekula. Metatranskriptomika daje informaciju o transkripcijski aktivnim mikrobnim populacijama te o aktivnim metaboličkim putovima. Korištenjem ove metode moguće je pobliže opisati interakcije između mikrobne zajednice i domaćina te utjecaj konkretne mikrobne aktivnosti na razvoj bolesti ili na održavanje homeostaze. Od prvih studija objavljenih oko 2010. godine, broj istraživanja koja uključuju metatranskriptomiku rapidno raste. Ova je metoda

analize još u razvoju i nailazi na mnoge izazove, ali njena je perspektiva obećavajuća, a razvoj brz (Shakya i sur., 2019).

Uz metatranskriptomiku, kao perspektivne metode za istraživanje uloga crijevne mikrobiote valja spomenuti i metaproteomiku te metabolomiku. Metode koje se trenutno ubrzano razvijaju i može se očekivati njihova šira primjena u skoroj budućnosti (Allaband i sur., 2019).

2.2.2. Sekvenciranje

Sekvenciranje je postupak određivanja redoslijeda monomernih jedinica u polimernim molekulama. Promatrajući molekulu DNA ili RNA, sekvenciranje se može definirati kao laboratorijski postupak određivanja slijeda nukleotida. Ukoliko je molekula od interesa protein ili peptid, sekvenciranje podrazumijeva određivanje redoslijeda aminokiselina. Tehnologije sekvenciranja mogu se podijeliti u 4 grupe: sekvenciranja prve generacije, sekvenciranja nove generacije (NGS) te sekvenciranje treće i četvrte generacije. Prelazak iz jedne generacije tehnologije u sljedeću determiniran je razvojem novih i boljih uređaja za sekvenciranje.

Za tehnologiju sekvenciranja prve generacije smatra se automatizirana Sangerova metoda sekvenciranja. 1977. godine Frederick Sanger, Allan Maxam i Walter Gilbert razvijaju metodu sekvenciranja koja se temeljila na korištenju radioaktivno obilježenih nukleotida. Metoda je s vremenom značajno unaprijeđena te su razvijeni automatizirani uređaji na kojima se provodila analiza. To je dovelo do njene šire primjene sve do 2005. godine kada dolazi do razvoja NGS tehnologija. Analiza na NGS sekvencerima bila je brža jer se temelji na paralelnom čitanju više sekvenci (visokoj protočnosti) te izravnom očitavanju baze u sekvenci bez upotrebe elektroforeze. Daljnjim napretkom, dolazi do razvoja tehnologije sekvenciranja treće generacije, TGS. Ona dodatno ubrzava postupak analize jer se za pripremu uzorka ne koristi postupak PCR amplifikacije. Drugi značajan napredak jest taj što generira duga očitavanja sekvenci. Za razliku od NGS metode kojom je moguće dobiti očitavanja dugačka do nekoliko stotina parova baza, na TGS platformama očitavanja su 1000 puta dulja. Najnovija tehnologija sekvenciranja četvrte generacije ima sposobnost generiranja iznimno dugih očitavanja, od 10^4 do 10^6 baza. Metoda se temelji na korištenju pora nanometarskog promjera. Nanopore su postavljene između dvije komore koje su ispunjene ionima. Gibanjem iona generira se ionska struja, a prolaskom molekule analita kroz poru električni signal se zaustavlja te se temeljem tog prekida signala detektira slijed nukleotida u molekuli. Poznatiji uređaji, odnosno platforme za NGS analize su

Roche 454, Illumina, Ion Torrent, GenapSys, GeneReader i ABI SOLiD. Platforme za sekvenciranje treće generacije su PacBio i HeliScope, a sekvencer četvrte generacije je Oxford Nanopore (Pervez i sur., 2022; Štimac i Martinković, 2021).

2.2.2.1. Illumina

Trenutno najkorištenija platforma za sekvenciranje je Illumina. Tijek rada na uređaju započinje pripremom knjižnice, odnosno zbirke fragmenata koji će se sekvencirati, oni se zatim umnažaju nizom PCR reakcija te se u konačnici očitavaju baze.

Prvi korak u pripremi knjižnice je fragmentiranje DNA izolirane iz uzorka. Slijedi ligacija pri čemu se na oba kraja DNA fragmenata vežu adaptori i barkodovi. Adaptori su oligonukleotidi koji služe za vezanje fragmenta na podlogu, a barkodovi su jedinstvene oligonukleotidne sekvence koje se ligiraju na fragmente DNA kako bi se prilikom analize podataka moglo odrediti iz kojeg uzorka potječe koje očitavanje. Nakon fragmentacije i ligacije adaptora i barkodova provodi se 10 – 15 ciklusa PCR umnažanja. Dobiveni dvolančani fragmenti se zatim denaturiraju u jednolančane te vežu za oligonukleotidne početnice na podlozi, odnosno protočnom članku (engl. *flow cell*), na temelju njihove komplementarnosti s adapterima. Kada su fragmenti imobilizirani na podlogu ponovno se provodi PCR umnažanje. Dvolančani produkti PCR reakcije se denaturiraju, a izvorni fragmenti ispiru. Novosintetizirani lanci se zatim vežu svojim slobodnim krajem za početnicu na podlozi tako da tvore oblik mosta. Slijedi još jedan krug PCR umnažanja čime se dobivaju dvolančani mostovi. Oni se denaturiraju, a novonastali produkti su jednolančane molekule vezane za površinu. Proces PCR umnažanja mostova ponavlja se kako bi se dobio klaster jednolančanih molekula. Stvaranjem istovrsnih klastera koncentriranih na jednom mjestu pojačava se signal koji će baze emitirati tijekom sekvenciranja. Slijedi sam postupak sekvenciranja, odnosno „čitanja“ baza. Illumina koristi postupak sekvenciranja sintezom. Postupak započinje tako da se na reakcijsku podlogu s umnoženim fragmentima dodaju početnice, DNA polimeraza i nukleotidi. Svaki od 4 vrste nukleotida obilježen je fluorescentnim markerom (fluororforom) specifičnim za pojedinu vrstu nukleotida te svi nukleotidi imaju blokiran 3'-OH kraj. Blokiranjem 3'-OH kraja osigurava se da se u svakom ciklusu ugradi samo jedna baza. Polimeraza ugrađuje bazu komplementarnu lancu koji je usidren na podlogu. Ugrađene se baze izlažu laserskom snopu što dovodi do fluorescencije koja se snima kamerom te bilježi. Posljednji korak u prvom ciklusu je dodatak reagensa koji će ukloniti

fluorofor i blokirani 3'-OH kraj čime se omogućuje sljedeći krug elongacije DNA polimerazom te se cijeli postupak ponavlja dok se ne očitaju sve baze (Pervez i sur., 2022; Štimac i Martinković, 2021).

Karakteristični nedostaci ove tehnologije su slabljenje signala na višim pozicijama u sekvenci i defaziranje koje je uzrokovano nepotpunim cijepanjem fluorofora ili blokade s 3'-OH kraja. Dužina očitavanja je do 300 parova baza, a prosječna stopa pogreške na Illumina platformi je 1 – 1,5% (Cao i sur., 2017).

2.2.3. Bioinformatička analiza

Kako bi se nizovi sekvenci dobiveni sekvenciranjem preveli u smislene informacije o taksonomiji ili genskom profilu prisutnih vrsta u uzorku provodi se bioinformatička analiza tijekom koje se generirane sekvence uspoređuju s referentnim baza podataka, koje sadrže informacije o poznatim organizmima, s ciljem dodjeljivanja taksonomije.

Poput razvoja platformi za sekvenciranje, ubrzan razvoj doživjeli su i računalni programi koji se koriste za obradu sirovih podataka. Tako je razvijena nekolicina računalnih alata za analizu sekvenci 16S rRNA marker gena, a neki od njih su QIIME, Mothur, MG-RAST, Genboree, EzTaxon, Phonix2, METAGENassist, MEGANE, VAMPS, SnoWMan, CloVR-16S, RDPipeline, Vegan, ade4 i ape. Neki od njih mogu provesti cjelovitu analizu jer uključuju razne algoritme za kontrolu kvalitete i klasteriranje sekvenci, određivanje taksonomije i izračunavanje mjera raznolikosti uzoraka te vizualizaciju rezultata, dok se drugi moraju kombinirati s drugim alatima (Plummer i sur., 2015).

QIIME2 (*Quantitative Insights Into Microbial Ecology*) je računalni program koji omogućuje cjelovitu analizu podataka generiranih sekvenciranjem na jednostavan i relativno brz način, kroz nekoliko naredbi u programskom jeziku Python. QIIME2 ima zadane parametre koji su podešeni za analizu podataka dobivenih sekvenciranjem 16S rRNA genskih fragmenata kao očitavanja uparenih krajeva (engl. *paired-end reads*) na Illumina platformi. Moguće je analizirati i druge genske fragmente s drugih platformi, ali pritom valja podesiti zadane parametre po potrebi. Za korištenje QIIME2 potrebno ga je instalirati pomoću Anaconda ili Miniconda paketa te pozvati naredbom `conda activate qiime2` čime se aktivira programski paket s trenutno dostupna 22 dodatka (engl. *plugin*) koji sadrže razne metode za obradu podataka sekvenciranja. QIIME2 radi s dvije glavne vrste datoteka: QIIME artefakti (datoteke s nastavkom `.qza`) i QIIME vizualizacije

(datoteke s nastavkom .qzv). QIIME artefakti su izlazne datoteke provedenih naredbi, a QIIME vizualizacije su datoteke koje sadrže rezultate analize koji se trebaju vizualizirati skupa s kodom potrebnim za njihovu vizualizaciju. Generirane .qzv datoteke mogu se otvoriti na internetskoj stranici QIIME2 View (<https://view.qiime2.org/>) (Hall i Beiko, 2018).

2.2.3.1. Tijek rada u QIIME2

Prilikom analize podataka u QIIME2 programu prvo je potrebno utvrditi kvalitetu sirovih podataka koji se analiziraju. Njihova kvaliteta ovisi o eksperimentalnim čimbenicima i uvjetima u svim koracima koji su prethodili bioinformatičkoj analizi. Kako je ranije objašnjeno, zbog same prirode Illumina metode sekvenciranja, za očekivati je kako će kvaliteta baza na višim pozicijama u sekvenci biti lošija. Pozicija baze nakon koje kvaliteta počinje padati determinirat će dužinu skraćivanja sekvenci u koraku pročišćavanja (Hall i Beiko, 2018).

Za pročišćavanje (engl. *denoising*) podataka QIIME2 ima dostupne metode DADA2 (engl. *The Divisive Amplicon Denoising Algorithm*) i Deblur. Njihova glavna uloga je ispravljanje pogrešaka u sekvencama što čini ovaj korak analize vremenski najzahtjevnijim. Razlike u izvedbi između ova dva dodatka su male pa stoga odabir koji od njih koristiti u analizi može biti temeljen na nekim praktičnim značajkama podataka ili prirodi analize. Tako na primjer DADA2 algoritam uključuje spajanje *paired-end* očitavanja pa je pogodniji za korištenje na *paired-end* setovima podataka, dok je Deblur pogodniji za korištenje u meta-analizama jer se koristi unaprijed izračunatom stopom pogreške za razliku od DADA2 koji ju izračunava za svaku pojedinačnu analizu. U ovom je radu korišten DADA2 dodatak.

Sljedeći korak analize je konstruiranje filogenetskog stabla koje je potrebno u kasnijim koracima za izračun mjera filogenetske raznolikosti. U njemu su sadržane informacije o evolucijskoj povezanosti DNA sekvenci. U QIIME2 moguće je konstruirati filogenetsko stablo *de novo* metodom ili metodom insercije fragmenata pri čemu se sekvence iz uzorka uspoređuju s referentnom bazom sekvenci te zatim umeću u filogenetsko stablo. Metoda insercije fragmenata daje pouzdanije rezultate jer su sekvence dobivene iz 16S rRNA prekratke da točno obuhvate veće evolucijske udaljenosti pri konstrukciji *de novo* filogenetskog stabla (Estaki i sur., 2020).

Za određivanje taksonomske pripadnosti dostupan je dodatak *q2-feature-classifier*. On provodi taksonomsku klasifikaciju metodom korištenja klasifikatora strojnog učenja (npr. naivnih Bayesovih klasifikatora) ili metodom konsenzusa taksonomije temeljenom na poravnanju (engl.

alignment-based taxonomy consensus) pri čemu se koriste računalni alati BLAST+ i VSEARCH. Prije korištenja klasifikatore, bazirane na metodama strojnog učenja, je potrebno „istrenirati“ na referentnim sekvencama poznate taksonomije. Mogu se samostalno trenirati za svaku analizu, a postoje i javno dostupni istrenirani klasifikatori pogodni za različite vrste podataka. Na QIIME internetskoj stranici dostupni su klasifikatori istrenirani na Silva i Greengenes taksonomskim bazama (Bokulich i sur., 2018).

Posljednji koraci u analizi su izračun raznolikosti i testiranje statističke značajnosti. Za izračun mjera raznolikosti uzoraka, QIIME2 kao ulazne podatke koristi tablicu sekvenci, filogenetsko stablo i dubinu uzorkovanja koja se manualno određuje za svaku analizu (preporučena vrijednost dubine uzorkovanja jest broj sekvenci u uzorku u kojem je detektirano najmanje sekvenci nakon koraka pročišćavanja). Prema zadanim postavkama, kao mjere alfa-raznolikosti QIIME2 izračunava Faith-ovu filogenetsku raznolikost, Shannon-ovu raznolikost, Pielou-ovu jednakost (engl. *evenness*) te vektor opaženih značajki. Kao mjere beta raznolikosti QIIME2 će izračunati težinsku/netežinsku (engl. *weighted/unweighed*) Uni-Frac raznolikost, Bray-Curtis te Jaccard raznolikost. Moguće je izračunati i mnoge druge mjere raznolikosti pozivanjem specifičnih naredbi.

Za testiranje statističke značajnosti potrebno je odabrati odgovarajući test ovisno o biološkom pitanju, odnosno hipotezi. U QIIME2 dostupno je nekoliko statističkih testova, neki od njih su PREMANOVA, ANOSIM, PERMDISP, Mantel test te ANOVA test koji se koristi kako bi se utvrdilo postoji li statistički značajna razlika u relativnoj zastupljenosti ispitivanih taksonomskih razina među grupama (Estaki i sur., 2020; Hall i Beiko, 2018).

2.2.3.2. DADA2 algoritam

DADA2 je algoritam koji filtrira i reže sekvence, dereplicira, identificira kimere te spaja prednja i stražnja očitavanja (engl. *forward, reverse reads*).

U prvom koraku DADA2 uklanja nukleotide s početka i kraja očitavanja čija kvaliteta nije zadovoljavajuća. Zasebno se zadaje dužina skraćivanja za prednja i stražnja očitavanja. Pritom, kod analize uparenih krajeva (engl. *paired-end*), treba paziti da dužina oba očitavanja nakon skraćivanja bude dovoljna da omogući preklapanje najmanje 20 nukleotida u kasnijem koraku spajanja očitavanja. Zatim se filtriraju očitavanja koja imaju dvosmislene nukleotide te očitavanja koja imaju više od 2 očekivane pogreške. Broj očekivanih pogrešaka, E, definiran je kao srednja

vrijednost broja pogrešaka koje bi se uočile u velikoj zbirci sekvenci gdje je stopa pogreške na svakoj poziciji dana ocjenom kvalitete te pozicije (Edgar i Flyvbjerg, 2015). Dužina rezanja početnih i završnih nukleotida zadaje se ručno, a filter dvosmislenih nukleotida i očekivanih pogreški su zadani parametri, ali se mogu promijeniti ovisno o potrebama analize. Sljedeći korak je dereplikacija – postupak sažimanja podataka na način da se grupiraju sva očitavanja koja kodiraju istu sekvencu. Time se skraćuje vrijeme trajanja daljnjih koraka analize. Zatim DADA2 izračunava stopu pogreške izravno za set podataka na kojem se provodi analiza. Dakle koriste se egzaktna, a ne procijenjena vrijednosti. Stopa pogreške u kombinaciji s ocjenom kvalitete koristi se u glavnom dijelu algoritma za pročišćavanje grešaka. Nakon pročišćavanja pogrešaka, zasebno u prednjim i zasebno u stražnjim očitavanjima, dolazi do spajanja očitavanja te odbacivanja onih očitavanja koja se ne poklapaju. Naposljetku se uklanjaju kimere. Kimere su „lažni“ amplikoni koji nastaju tijekom PCR amplifikacije kada amplifikacija jedne sekvence nije dovršena pa taj nepotpuni amplikon postane početnica za sintezu novog amplikona. U konačnici nastaje hibridni amplikon koji je potekao od 2 različite sekvence. Kao izlazni artefakt DADA2 analize generiraju se ASV (engl. *Amplicon Sequence Variants*) tablice. Iz izlaznih tablica može se očitati koliko je pročišćenih i sparenih sekvenci detektirano u svakom uzorku. Uzorci koji imaju mali broj sekvenci, odnosno nisku pokrivenosti (engl. *sequence depth*) mogu biti isključeni iz daljnje analize (Callahan i sur., 2023).

Princip generiranja ASV-ova smatra se svojevrsnom zamjenom za dosad korištenu metodu ispravljanja pogrešaka konstruiranjem OTU-ova (engl. *Operational Taxonomic Units*). OTU je klaster sekvenci koje se međusobno razlikuju za manje od neke zadane vrijednosti, koja obično iznosi 3 %. Pozitivan aspekt korištenja OTU-a jest taj što se spajanjem sekvenci s 97 %-tnom podudarnošću smanjuje stopa pogrešne interpretacije grešaka u sekvenci kao biološke varijacije, dakle izbjegavaju se lažno pozitivni rezultati. S druge strane, ova metoda ne iskorištava prednosti modernih metoda sekvenciranja tako što onemogućava detekciju sitnih varijacija. Korištenjem ASV-ova omogućena je detekcija najmanjih varijacija uz mali broj lažno pozitivnih rezultata. Dodatno, DADA2 ne koristi referentne sekvence te se stoga može koristiti za pročišćavanje podataka za bilo koji genski lokus (Callahan i sur., 2016).

3. EKSPERIMENTALNI DIO

3.1. ISPITANICI

U ovom istraživanju sudjelovalo je 59 ispitanika u dobi od 18 do 59 godina. 60 % ispitanika bilo je ženskog, a 40 % muškog spola. Rekrutacija se odvijala interno, među djelatnicima Instituta za antropologiju i Prehrambeno-biotehnološkog fakulteta u Zagrebu – institucijama koje su ujedno i provodile istraživanje kao dio pilot projekta „Crijevna mikrobiota radno aktivne populacije u Republici Hrvatskoj“. Projekt se provodio u vremenskom periodu od listopada 2021. do siječnja 2022. godine.

Svi su ispitanici u istraživanju sudjelovali dobrovoljno te su prije sudjelovanja potpisali informirani pristanak za sudjelovanje prilikom čega su detaljno upoznati s ciljem i tijekom istraživanja. Istraživanje je provedeno prema smjernicama za pravilno prikupljanje, obradu i pohranjivanje podataka te su identitet ispitanika, kao i svi dobiveni podatci šifrirani radi zaštite osobnih podataka ispitanika.

Zadatak ispitanika u ovom istraživanju bio je ispuniti upitnik o životnim i prehrambenim navikama te prikupiti i dostaviti uzorak stolice.

Istraživanje je odobreno od strane Etičkog povjerenstva Instituta za antropologiju.

3.2.METODE RADA

Istraživanje je provedeno korištenjem dijetetičkih, genomičkih i bioinformatičkih metoda.

3.2.1. Dijetetičke metode

Dijetetička metoda korištena u istraživanju jest upitnik o prehrambenim navikama. Upitnik se sastojao od 65 pitanja te je, osim pitanja o prehrambenim navikama, sadržavao i pitanja o životnim navikama, antropometrijskim parametrima te neka općenita pitanja. Od posebnog značaja za analizu provedenu u ovom radu bila su pitanja o učestalosti stolice kao i o procjeni konzistencije stolice na temelju Bristolske skale stolice.

Na kraju upitnika uključena je dijetetička metoda 24-satnog prisjećanja na način da su ispitanici morali navesti svu hranu konzumiranu u prethodna dva dana.

Cijeli upitnik ispitanici su ispunjavali samostalno, preko internetske platforme *Google forms*.

3.2.2. Genomičke metode

Za određivanje sastava crijevne mikrobiote ispitanika korištena je genomička metoda sekvenciranja amplikona. Prije početka analize svi prikupljeni uzorci stolice bili su pohranjeni u rashladnim uređajima na -80 °C. DNA iz uzoraka izolirana je korištenjem *QIAmp® PowerFecal®* DNA seta kemikalija i praćenjem *Quick Start* protokola. Za pripremu knjižnice korišten je *Illumina TruSeq DNA library* protokol. Amplikoni su dobiveni nakon provedbe 30 ciklusa PCR umnažanja V3 i V4 regije 16S rRNA gena za što su korištene početnice 341F (5'- CCTAYGGGRBGCASCAG -3') 806R (5'-GGACTACNNGGGTATCTAAT -3'). Provjera PCR amplikona provedena je na 2 %-tnom agaroznom gelu, a pročišćavanje je provedeno uz korištenje kalibriranih *Ampure XP* kuglica. Sekvenciranje je obavljeno prema protokolu za sekvenciranje uparenih krajeva (engl. *paired-end sequencing*) na platformi Illumina MiSeq na Institutu za antropologiju u Zagrebu.

3.2.3. Bioinformatičke metode

Bioinformatička i statistička obrada podataka provedena je korištenjem računalnog programa QIIME2. Nakon provedene analize na Illumina MiSeq uređaju dobiveni su sirovi podatci u obliku demultipleksiranih fastq dokumenata. Za svaki uzorak dobivena su 2 dokumenta – jedan za prednje očitavanje i drugi za stražnje očitavanje. Ime dokumenta za jedan od uzoraka ima oblik IM-1_S26_L001_R1_001.fastq.gz. Pritom IM-1 predstavlja ime uzorka, S26 je identifikator barkoda, L001 je broj trake, R1 predstavlja smjer očitavanja koji je u ovom slučaju prednji, a 001 označava broj seta. Fastq je oznaka vrste dokumenta koji sadrži niz nukleotida i podatke o njihovoj kvaliteti, a nastavak -gz znači da je dokument komprimiran kako bi zauzimao manje memorije. Demultipleksirani sirovi podatci su nizovi sekvenci kojima je uklonjen barkod i koji su nazvani prema ishodišnom uzorku.

3.2.3.1. Analiza uparenih krajeva

Za unos podataka u QIIME2 bilo je potrebno kreirati tzv. *Manifest* dokument, a zatim je

korištena naredba `qiime tools import` za unos podataka.

Manifest dokument je tekstualna datoteka (engl. *tab-separated text file*, tsv) koja u svom sadržaju povezuje identifikacijski broj uzorka s apsolutnom putanjom fastq dokumenta u kojemu se nalazi sekvenca i podatci o kvaliteti sekvence za uzorak. *Manifest* dokument radi distinkciju između prednjeg i stražnjeg očitavanja. U prvom stupcu manifest dokumenta nalaze se identifikacijski brojevi uzorka, u drugom stupcu apsolutna putanja fastq dokumenta koji sadrži prednje očitavanja, a u trećem stupcu nalazi se apsolutna putanja fastq dokumenta koji sadrži stražnja očitavanja.

Za kreiranje *manifest* dokumenta prvo je pomoću naredbe `ls -d $PWD/*> Manifest_file` kreiran tekstualni dokument koji je sadržavao popis apsolutnih putanja svih originalnih fastq dokumenata. Taj je dokument otvoren u programu Microsoft Excel, gdje je uređen na gore opisan način tako da svaki stupac sadrži potrebne podatke. Zatim je Microsoft Excel dokument izvezen u formatu tsv dokumenta. Također, kreirana je i datoteka s meta-podacima pod nazivom „*mapping.txt*“ koja je sadržavala podatke o imenu uzorka i pripadajuće meta-podatke o uzorcima dobivene iz ispunjenih upitnika.

Nakon kreiranja dokumenata, pomoću iduće naredbe podatci su uneseni u program:

```
qiime tools import --type 'SampleData[SequencesWithQuality]' --input-path
Manifest_file.tsv --output-path paired-end-demux.qza --input-format
PairedEndFastqManifestPhred33V2
```

Korišten je format unosa „*PairedEndFastqManifestPhred33V2*“ jer su sirovi podatci generirani korištenjem programskog paketa Illumina 1.8 ili novije verzije.

Nakon unosa podataka u program, sljedeći je korak, prema standardnim operativnim postupcima QIIME platforme, demultipleksiranje. Međutim, taj korak u ovoj analizi nije bilo potrebno izvoditi jer su dobiveni podatci već demultipleksirani. Demultipleksiranje se inače izvodi korištenjem naredbe `qiime demux emp-paired` za *paired-end* metodu analize ili `qiime demux emp-single` za *single-end* metodu analize.

Idući korak jest procjena kvalitete dobivenih očitavanja. Za to je potrebno napraviti vizualizaciju demultipleksiranih uzoraka naredbom:

```
qiime demux summarize --i-data paired-end-demux.qza --o-visualization
paired-end-demux.qzv
```

Zatim slijedi pročišćavanje podataka kojim se nastoji filtrirati sekvence nedovoljne kvalitete i ukloniti greške. Za pročišćavanje je korišten dodatak DADA2 koji je pozvan naredbom:

```
qiime dada2 denoise-paired --i-demultiplexed-seqs paired-end-demux.qza
--p-trim-left-f 17 --p-trim-left-r 21 --p-trunc-len-f 220 --p-trunc-len-r 220
--o-table table.qza --o-representative-sequences rep-seqs.qza
--o-denoising-stats denoising-stats.qza
```

Parametrom `--p-trim-left-f` zadano je da se s prednjih očitavanja reže 17 baza, a parametrom `--p-trim-left-r` zadano je da se sa stražnjih očitavanja reže 21 baza na početku sekvence. Te vrijednosti odgovaraju broju nukleotida u početnicama korištenima za sekvenciranje. Također, sve sekvence skraćene su na duljinu od 220 baza što je zadano kroz parametre `--p-trunc-len-f` i `--p-trunc-len-r`. Ta vrijednost odabrana je na temelju rezultata prethodnog koraka. Iz generiranih grafičkih prikaza bilo je vidljivo kako kvaliteta baza pada ispod vrijednosti 20 za sekvence duže od 220 baza. Izlazni podatci koraka pročišćavanja - tablica značajki (engl. *Feature Table*), sekvence koje su zadovoljile kvalitativne parametre i statistički podatci o pročišćavanju vizualizirani su korištenjem sljedećih naredbi:

```
qiime feature-table summarize --i-table table.qza --o-visualization
table.qzv --m-sample-metadata-file mapping.txt
```

```
qiime feature-table tabulate-seqs --i-data rep-seqs.qza --o-visualization
rep-seqs.qzv
```

```
qiime metadata tabulate --m-input-file denoising-stats.qza
--o-visualization denoising-stats.qzv
```

Rezultati dobiveni pročišćavanjem nisu bili dovoljno dobri za nastavak *paired-end* analize pa je

odlučeno prijeći na drugačiji pristup i podatke obraditi analizom jednog kraja (engl. *single-end*).

3.2.3.2. Analiza jednoga kraja (engl. *single-end*)

Postupak je identičan kao za *paired-end* metodu, jedina je razlika što se u ovom slučaju kao sirovi podatci koriste samo prednja očitavanja, a stražnja očitavanja se zanemaruju. Sirovi podatci uneseni su u program, za što je bilo potrebno kreirati novi *Manifest* dokument, a zatim vizualizirani kako bi se odredila njihova kvaliteta. Potom je proveden postupak pročišćavanja korištenjem DADA2 algoritma:

```
qiime dada2 denoise-single --i-demultiplexed-seqs single-end-demux.qza
--p-trim-left 17 --p-trunc-len 232 --o-representative-sequences
rep-seqs-single-end.qza --o-table table-single-end.qza --o-denoising-stats
stats-single-end.qza
```

Ovoga puta za parametar `--p-trunc-len` određena je vrijednost 232 što je skratilo sva očitavanja na tu dužinu baza. Izlazni podatci su vizualizirani istim naredbama kao u *paired-end* dijelu analize.

Korištenjem sljedeće naredbe kreirano je filogenetsko stablo potrebno za izračun mjera alfa i beta raznolikosti:

```
qiime phylogeny align-to-tree-mafft-fasttree --i-sequences
rep-seqs-single-end.qza --o-alignment aligned-rep-seqs.qza
--o-masked-alignment masked-aligned-rep-seqs.qza --o-tree unrooted-tree.qza
--o-rooted-tree rooted-tree.qza
```

Zatim je, korištenjem dodatka *q2-diversity*, programu zadano da izračuna mjere raznolikosti:

```
qiime diversity core-metrics-phylogenetic --i-phylogeny rooted-tree.qza
--i-table table-single-end.qza --p-sampling-depth 20849
--m-metadata-file mapping.txt --output-dir mjere-raznolikosti
```

U ovoj su naredbi kao ulazni parametri korišteni filogenetsko stablo i tablica značajki te je bilo potrebno odrediti parametar `--p-sampling-depth` koji predstavlja dubinu uzorkovanja. Kako je većina mjera raznolikosti osjetljiva na različite dubine uzorkovanja, ovim se parametrom zadaje broj značajki (jedinstvenih ASV-ova) koji će se biti izuzeti iz svakog uzorka kako bi se dubina uzorkovanja ujednačila. U ovoj analizi zadana je vrijednost 20 849 jer je to najniža frekvencija značajki dobivena nakon pročišćavanja što se može iščitati iz vizualizacije artefakta *table-single-end.qzv*. Odabirom ove vrijednosti niti jedan uzorak nije izostavljen iz daljnjeg izračuna raznolikosti. Izlazni artefakti ove naredbe korišteni su u daljnjem postupku za kreiranje grafičkih prikaza koji daju uvid u sastav mikrobiote detektirane u uzorcima.

Za uvid u alfa raznolikost uzoraka, pomoću sljedeće naredbe generiran je graf alfa-razrijeđenosti (engl. *alpha rarefaction*):

```
qiime diversity alpha-rarefaction --i-table table-single-end.qza
--i-phylogeny rooted-tree.qza --p-max-depth 76568 --m-metadata-file
mapping.txt --o-visualization alpha-rarefaction.qzv
```

Parametar `--p-max-depth` odredit će najvišu dubinu uzorkovanja za kreiranje grafa, a vrijednost 76 568 je odabrana jer je to ujedno i medijalna vrijednost frekvencije značajki iščitana iz tablice značajki.

Sljedeći korak bio je određivanje taksonomije. Korišten je klasifikator treniran na SILVA taksonomskoj bazi u kojem su sekvence skraćene na 215 baza, a izrezan je dio 16S rRNA gena parom primera 314F-805R, upotrebljenim za umnažanje amplikona.

```
qiime feature-classifier classify-sklearn --i-classifier
classifier-silva-138-341F-805R.qza --i-reads rep-seqs-single-end.qza
--o-classification taxonomy.qza
```

Rezultati taksonomske analize vizualizirani su sljedećim naredbama:

```
qiime metadata tabulate --m-input-file taxonomy.qza --o-visualization
taxonomy.qzv
```

```
qiime taxa barplot --i-table table-single-end.qza --i-taxonomy taxonomy.qza
--m-metadata-file mapping.txt --o-visualization taxa-bar-plots.qzv
```

3.2.3.3. Obrada podataka

Za provjeru statističke značajnosti rezultata provedeni su sljedeći koraci. Prvo je provjerena značajnost mjera alfa raznolikosti unutar grupa. Grafički prikazi Faith-ove filogenetske raznolikosti te Pielou-ove jednakosti dobiveni su korištenjem naredbi:

```
qiime diversity alpha-group-significance --i-alpha-diversity
faith_pd_vector.qza --m-metadata-file mapping.txt --o-visualization
faith-pd-group-significance.qzv
```

```
qiime diversity alpha-group-significance --i-alpha-diversity
evenness_vector.qza --m-metadata-file mapping.txt --o-visualization
evenness-group-significance.qzv
```

Za uvid u beta raznolikost, odnosno kako bi se provjerilo postoje li značajne razlike među grupama proveden je PERMANOVA test na netežinskoj UniFrac udaljenosti:

```
qiime diversity beta-group-significance --i-distance-matrix
unweighted_unifrac_distance_matrix.qza --m-metadata-file mapping.txt
--m-metadata-column broj_stolica --o-visualization
unweighted-unifrac-significance.qzv --p-pairwise
```

Parametar `--p-pairwise` koristio se kako bi u izlaznim rezultatima bilo vidljivo između kojih točno grupa je zamijećena međusobna razlika.

Također, korištenjem *Emperor* računalnog alata, iz mjera netežinske UniFrac udaljenosti

generiran je grafički prikaz udaljenosti korištenjem PCoA (engl. *Principle Coordinate Analysis*) metode:

```
qiime emperor plot --i-pcoa unweighted_unifrac_pcoa_results.qza
--m-metadata-file mapping.txt --o-visualization
unweighted-unifrac-emperor.qzv
```

Naposlijetku, provedena je analiza diferencijalnog obilja (engl. *differential abundance analysis*) ANCOM metodom. Za pripremu podataka za ANCOM test nužno je provesti sljedeću naredbu koja korigira tablicu značajki na način da, ukoliko je broj zapažanja određene značajke u uzorku iznosio 0, ovom se naredbom mijenja u 1:

```
qiime composition add-pseudocount --i-table table-single-end.qza
--o-composition-table pseudocount-table-single-end.qza
```

Ovaj je korak potrebno provesti zbog same prirode računanja diferencijalnog obilja koje se temelji na izračunu logaritamskih omjera – kako je $\log(0)$ nedefinirana vrijednost ovom se preinakom podataka omogućuje usporedba svih značajki iz svih uzoraka.

Sljedećom je naredbom provedena ANCOM analiza zasebno za svaku od 3 ispitivane kategorije koje su definirane u meta-podatcima i to u kolonama 'broj_stolica', 'broj_stolica1' i 'konzistencija':

```
qiime composition ancom --i-table pseudocount-table-single-end.qza
--m-metadata-file mapping.txt --m-metadata-column broj_stolica
--o-visualization ancom-broj_stolica.qzv
```

Također, ANCOM analiza provedena je i za svaku ispitivanu kategoriju na nivou roda. Za to je bilo potrebno tablicu značajki prilagoditi sljedećom naredbom:


```
qiime taxa collapse --i-table table-single-end.qza --i-taxonomy
taxonomy.qza --p-level 6 --o-collapsed-table table-level6.qza
```

Nova tablica značajki provedena je kroz naredbe `qiime composition add-pseudocount` i `qiime composition ancom` kako je prethodno opisano.

4. REZULTATI I RASPRAVA

Cilj ovoga rada bio je odrediti taksonomski sastav crijevne mikrobiote ispitanika uključenih u pilot projekt i utvrditi postoji li značajna razlika u sastavu crijevne mikrobiote ispitanika u ovisnosti o učestalosti i konzistenciji stolice.

Sastav crijevne mikrobiote ovisi o mnogim genetski uvjetovanim i okolišnim čimbenicima te o životnim navikama pojedinca. Neki od istraživanih čimbenika su dob, indeks tjelesne mase, način poroda, različiti obrasci prehrane, korištenje lijekova, pušenje, konzumacija alkohola, tjelesna aktivnost i konzistencija stolice (Schmidt i sur., 2018). Svaki od navedenih čimbenika može predstavljati zbunjujuću (engl. *confounding*) varijablu u istraživanju te ih je potrebno razmotriti prilikom dizajna studija na temu sastava crijevne mikrobiote. Vujkovic-Cvijin i sur. (2020) pokazali su kako konzistencija stolice ima snažan učinak na sastav mikrobiote i može utjecati na rezultate analize.

U ovom je radu promatrana povezanost konzistencije i učestalosti stolice na sastav mikrobiote. Informacije o sastavu i konzistenciji prikupljene su kroz upitnik, a sastav mikrobiote određen je iz uzoraka stolice 59 ispitanika metodom sekvenciranja amplikona na Illumina platformi. Dobiveni podaci obrađeni su računalnim programom QIIME2. Za interpretaciju rezultata ispitanici su podijeljeni u grupe s obzirom na broj stolica i s obzirom na konzistenciju stolice te je među tim grupama promatran taksonomski sastav te alfa i beta raznolikost crijevne mikrobiote. Prva podjela ispitanika bila je s obzirom na učestalost stolice – ispitanici su podijeljeni u 5 grupa na temelju prosječnog tjednog broja stolica. Druga je podjela također bila s obzirom na učestalost stolice, ali su ispitanici bili razvrstani u 2 grupe na temelju prosječnog dnevnog broja stolica, dok je treća promatrana podjela bila s obzirom na konzistenciju stolice i ona je dijelila ispitanike u 6 grupa.

4.1. ISPITANICI

Istraživanje je provedeno na 59 ispitanika, od toga je 35 ispitanika bilo ženskog, a 24 muškog spola. Starost ispitanika bila je između 18 i 59 godina. Podatci prikupljeni o ispitanicima koji bi mogli utjecati na sastav mikrobiote bili su način poroda, korištenje antibiotika, konzumacija probiotika, tjelesna aktivnost, učestalost konzumacije alkohola i pušenje.

Antibiotike je, u periodu od 3 mjeseca prije prikupljanja uzoraka stolice, koristilo 8 ispitanika, odnosno 13,6 %. Ti ispitanici nisu isključeni iz analize.

4 ispitanika je kao prosječnu učestalost stolice navelo „više od dva puta dnevno“, njih 19 „jednom do dva puta dnevno“, 28 ispitanika ima prosječnu učestalost stolice „jednom dnevno“, 4 ispitanika „svaki drugi dan“, 2 ispitanika „manje od tri puta tjedno“ te 1 ispitanik „manje od dva puta tjedno“. Za 1 ispitanika nisu bili dostupni podatci o učestalosti stolice.

Prema ovim odgovorima ispitanici su razvrstani u 5 grupa s obzirom na prosječnu tjednu učestalost stolice. Prva grupa obuhvatila je ispitanike koji imaju >14 stolica tjedno te je grupa nazvana „jako_često“, grupa „često“ obuhvaćala je ispitanike sa 7-14 stolica tjedno, grupu „normalno“ čine ispitanici sa 7 stolica tjedno, grupa „rijetko“ obuhvaća ispitanike s 3-4 stolice tjedno, a grupa „jako_rijetko“ ispitanike s 1-2 stolice tjedno. Broj ispitanika u grupama nije proporcionalno raspodijeljen, a prikaz broja ispitanika u svakoj od grupa vidljiv je u tablici 1.

Tablica 1. Raspodjela ispitanika po grupama s obzirom na prosječan tjedni broj stolica

Grupa	Prosječan tjedni broj stolica	Broj ispitanika	%-tak ispitanika
Jako_često	>14	4	7 %
Često	7-14	19	32 %
Normalno	7	28	47 %
Rijetko	3-4	4	7 %
Jako_rijetko	1-2	3	5 %
Nema podataka	/	1	2 %

Druga podjela ispitanika s obzirom na učestalost stolice napravljena je na način da se smanji velika razlika u broju članova pojedine grupe pa su stoga svi ispitanici podijeljeni u 2 grupe s obzirom na prosječan dnevni broj stolica. Grupa pod nazivom „manje“ obuhvatila je ispitanike koji su imali 1 stolicu dnevno ili manje, a grupa „više“ obuhvatila je ispitanike koji su imali 1-2 stolice dnevno ili više. Pritom je broj ispitanika u grupi „manje“ iznosio 35, odnosno 60 %, a u grupu „više“ ušlo je 23 ispitanika, što čini 40 % svih ispitanika.

Treća podjela ispitanika bila je s obzirom na konzistenciju stolice. Ispitanici su prilikom ispunjavanja upitnika procijenili konzistenciju svoje stolice uvidom u Bristolsku skalu stolice te su temeljem tih odgovora podijeljeni u 6 grupa. Bristolska skala stolice inače ima 7 kategorija prema 7 tipova stolice, ali niti jedna ispitanik nije imao tip 6 stolice. Tablica 2 prikazuje raspodjelu broja ispitanika po grupama u ovisnosti o tipu stolice.

Tablica 2. Raspodjela ispitanika po grupama s obzirom na konzistenciju stolice

Grupa	Opis grupe	Broj ispitanika	%-tak ispitanika
Tip 1	Konstipacija	1	2 %
Tip 2	Konstipacija	5	8 %
Tip 3	Normalno	26	44 %
Tip 4	Normalno	22	37 %
Tip 5	Normalno	3	5 %
Tip 7	Dijareja	1	2 %
Nema podataka	/	1	2 %

4.2. REZULTATI PROCJENE KVALITETE I PROČIŠĆAVANJA

4.2.1. Metoda spojenih krajeva

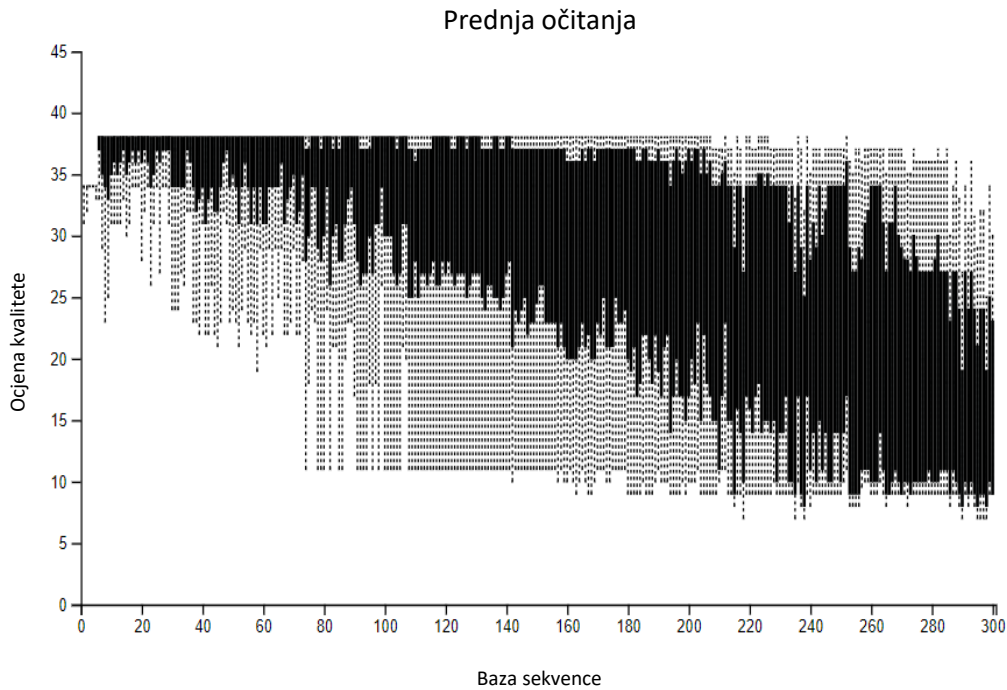
4.2.1.1. Procjena kvalitete podataka

Rezultati vizualizacije sirovih podataka u svrhu procjene kvalitete pokazali su kako je medijalna vrijednost broja očitavanja u uzorcima oko 172 000. Uzorak IM-64 imao je najveći broj očitavanja te on iznosi 305 972, dok je najmanji broj očitavanja, njih 49 649, imao uzorak IM-60. Ostale vrijednosti za prednja i stražnja očitavanja vidljiva su u tablici 3.

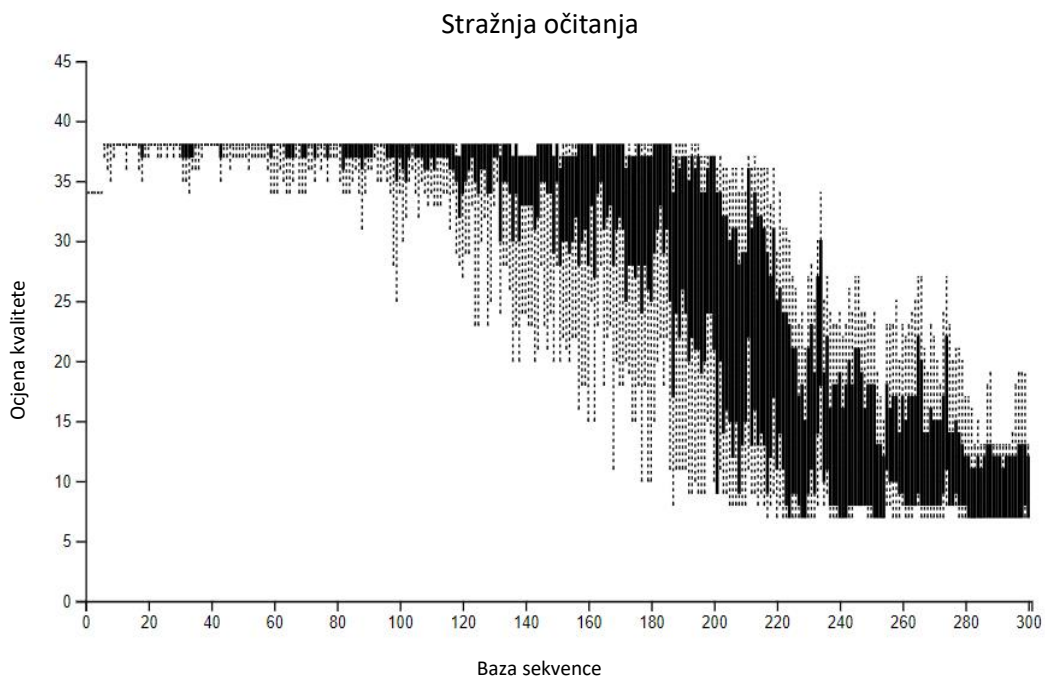
Tablica 3. Broj sekvenci u sirovim uzorcima

	Prednja očitavanja	Stražnja očitavanja
Minimum	49 649	49 649
Medijan	172 129	172 129
Srednja vrijednost	166 024	166 024
Maksimum	305 972	305 972
Ukupni broj sekvenci	9 795 448	9 795 448

Grafički prikaz ocjene kvalitete (engl. *Quality score*) baza u sekvenci vidljiv je na slikama 1 i 2. Prikaz je generiran na način da je nasumično odabrano 10 000 očitavanja od svih 9 795 448 očitavanja detektiranih u uzorcima te je prikazana prosječna kvaliteta baza od 1. do 300. pozicije u sekvenci. Prema tim je vrijednostima odabrana duljina nukleotida na koju će sekvence biti skraćene u procesu pročišćavanja – iznosila je 220. Ta je vrijednost odabrana jer nakon 220. baze medijan ocjene kvalitete baza pada ispod vrijednosti 20.



Slika 1. Grafički prikaz ocjene kvalitete za baze u prednjim očitanjima



Slika 2. Grafički prikaz ocjene kvalitete baza u stražnjim očitanjima

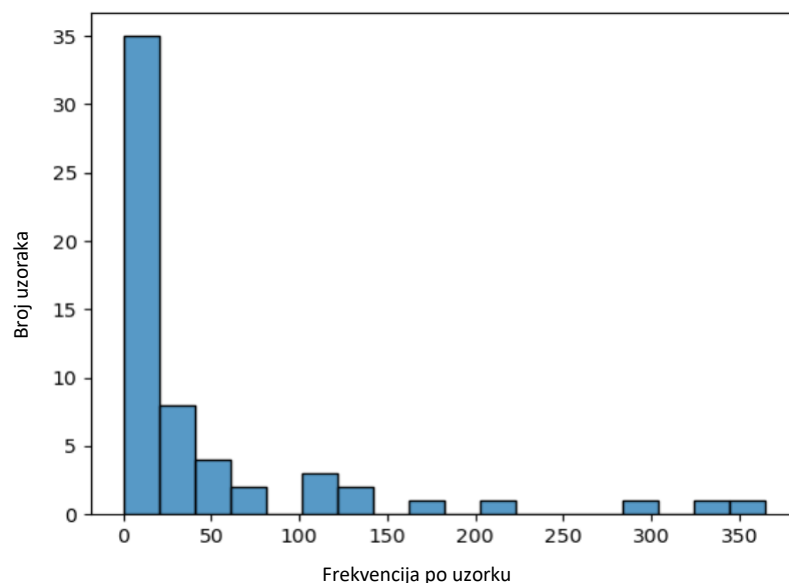
4.2.1.2. Pročišćavanje

Nakon pročišćavanja generirani su statistički podatci o broju filtriranih, pročišćenih, spojenih i ne-kimernih sekvenci. U tablici 4 prikazani su podatci za dio uzoraka te je vidljivo kako je između 40 i 50 % ulaznih sekvenci prošlo filter, dok je postotak sekvenci koje su uspješno spojene i postotak ne-kimernih sekvenci zanemariv, iznosi manje od 1 %.

Tablica 4. Statistički podatci o procesu pročišćavanja sirovih podataka za *paired-end* metodu

Ime uzorka	Ulazne sekvence	Filtrirano	% filtriranog	Pročišćeno	Spojeno	% spojenog	Ne-kimerne sekv.	% ne-kimernih sekv.
IM-1-S26	132938	65256	49.09	64992	4	0	4	0
IM-10-S60	188062	74189	39.45	73962	0	0	0	0
IM-11-S61	193002	90628	46.96	90068	15	0.01	15	0.01
IM-12-S42	139703	62270	44.57	61978	12	0.01	12	0.01
IM-13-S62	183527	76199	41.52	75908	4	0	4	0
IM-14-S43	126651	58863	46.48	58622	13	0.01	13	0.01
IM-15-S23	144829	65929	45.52	65633	16	0.01	16	0.01
IM-16-S63	208316	97418	46.76	97105	54	0.03	54	0.03
IM-17-S44	124062	55832	45	55662	7	0.01	7	0.01
IM-18-S45	127797	59216	46.34	58896	12	0.01	12	0.01
IM-19-S46	137283	56584	41.22	56182	121	0.09	121	0.09
IM-2-S27	107361	42235	39.34	42063	2	0	2	0
IM-21-S47	149985	74126	49.42	73714	3	0	3	0
IM-22-S64	177243	78536	44.31	78233	28	0.02	28	0.02

Iz podataka o sekvencama i iz tablice značajki bilo je vidljivo kako je detektirano samo 68 jedinstvenih sekvenci, odnosno značajki (engl. *feature*) u svih 59 uzoraka. Ukupna frekvencija njihovog pojavljivanja u uzorcima bila je 2759 puta. Na slici 3 prikazan je graf frekvencije po uzorku iz kojega je vidljivo kako je u većini uzoraka, njih 35, frekvencija značajki niža od 25. Pritom u 8 uzoraka nije spojena niti jedan sekvenca.



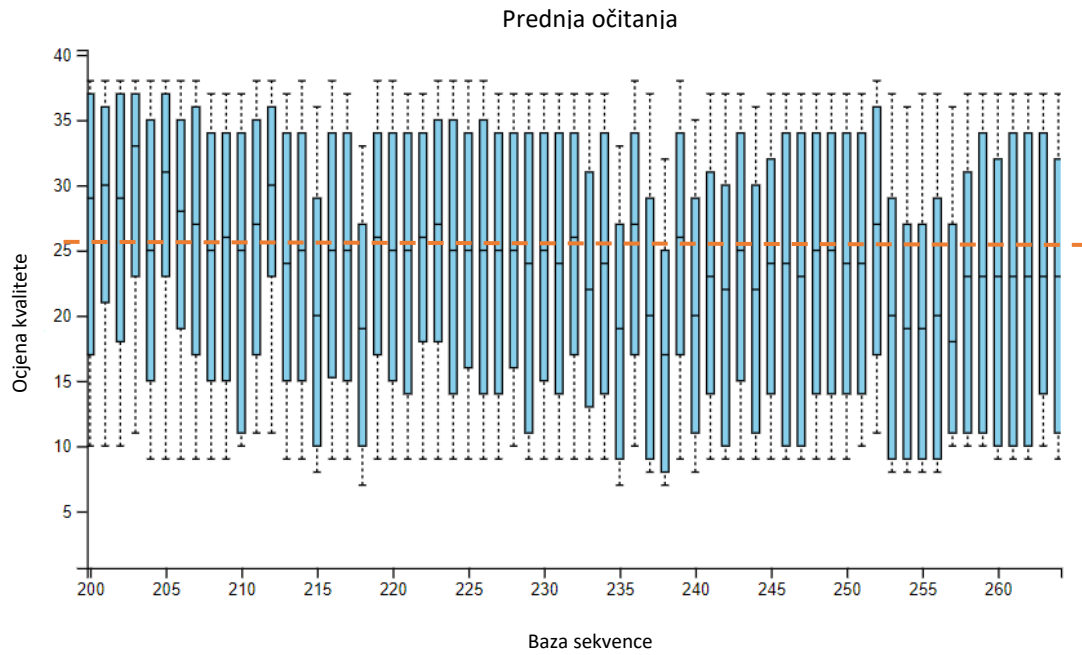
Slika 3. Grafički prikaz frekvencija značajki po uzorku za *paired-end* analizu

Iako je broj očitavanja po uzorku bio visok, očitavanja su bila loše kvalitete te stoga nije bilo moguće provesti uspješno spajanje prednjih i stražnjih očitavanja. U konačnici, nakon pročišćavanja uzorci su imali jako nisku pokrivenost. Nastavak *paired-end* analize nije imao smisla pa je odlučeno prijeći na *single-end* metodu analize.

4.2.2. Metoda jednog kraja

4.2.2.1. Procjena kvalitete podataka

Single-end analiza provedena je na istom setu sirovih podataka, a analizirala su se samo prednja očitavanja. Grafički prikaz ocjene kvalitete baza generiran je na isti način kao za *paired-end* analizu, a na slici 4 prikazan je uvećani graf ocjene kvalitete baza za prednje očitavanje. Vidljivo je kako nakon 232. pozicije medijan ocjene kvalitete počinje oscilirati, a zatim padati ispod vrijednosti 25. Iz tog je razloga odabrana vrijednost 232 za parametar `--p-trunc-len` u procesu pročišćavanja.



Slika 4. Uvećani grafički prikaz ocjene kvalitete baza u prednjim očitanjima

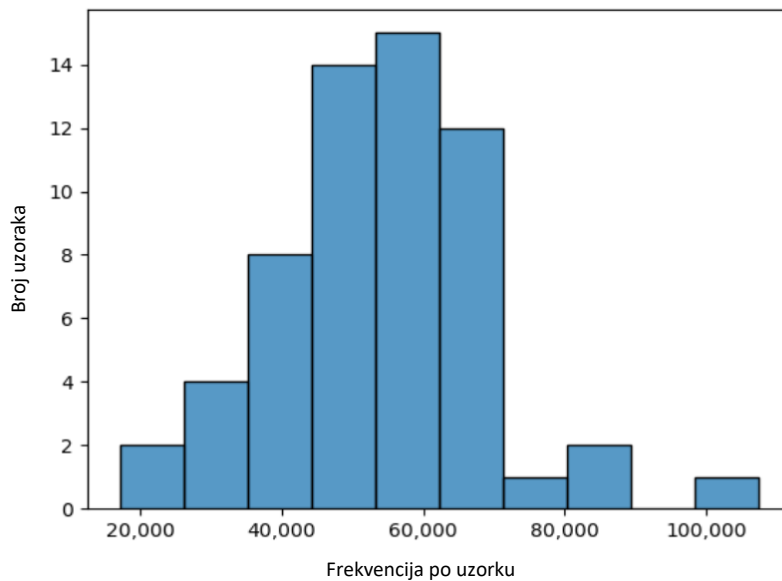
4.2.2.2. Pročišćavanje

Tablica 5 prikazuje statističke podatke procesa pročišćavanja za dio uzoraka iz kojih je vidljivo kako je između 40 i 50 % ulaznih sekvenci prošlo filter te je sličan postotak ulaznih sekvenci prepoznat kao ne-kimerne sekvence.

Tablica 5. Statistički podatci o procesu pročišćavanja sirovih podataka za *single-end* metodu

Ime uzorka	Ulazne sekvence	Filtrirano	% filtriranog	Pročišćeno	Ne-kimerne sekvence	% ne-kimernih sekvenci
IM-1-S26	132938	66289	49.86	65915	63033	47.42
IM-10-S60	188062	80764	42.95	80399	76568	40.71
IM-11-S61	193002	93353	48.37	92766	89484	46.36
IM-12-S42	139703	62338	44.62	62014	61377	43.93
IM-13-S62	183527	79050	43.07	78701	76145	41.49
IM-14-S43	126651	59999	47.37	59669	58920	46.52
IM-15-S23	144829	67391	46.53	66907	66434	45.87
IM-16-S63	208316	104858	50.34	104476	98056	47.07
IM-17-S44	124062	54954	44.3	54802	54124	43.63
IM-18-S45	127797	60972	47.71	60580	59731	46.74
IM-19-S46	137283	61657	44.91	61120	59424	43.29
IM-2-S27	107361	42152	39.26	41922	40447	37.67
IM-21-S47	149985	75430	50.29	75030	72768	48.52
IM-22-S64	177243	86261	48.67	85696	82423	46.5

Iz vizualizacije podataka o sekvencama i iz tablice značajki vidljivo je kako je broj detektiranih značajki u 59 uzoraka iznosio 2 804, a njihova ukupna frekvencija 3 136 274. Na slici 5 vidljiv je grafički prikaz frekvencija značajki po uzorku.

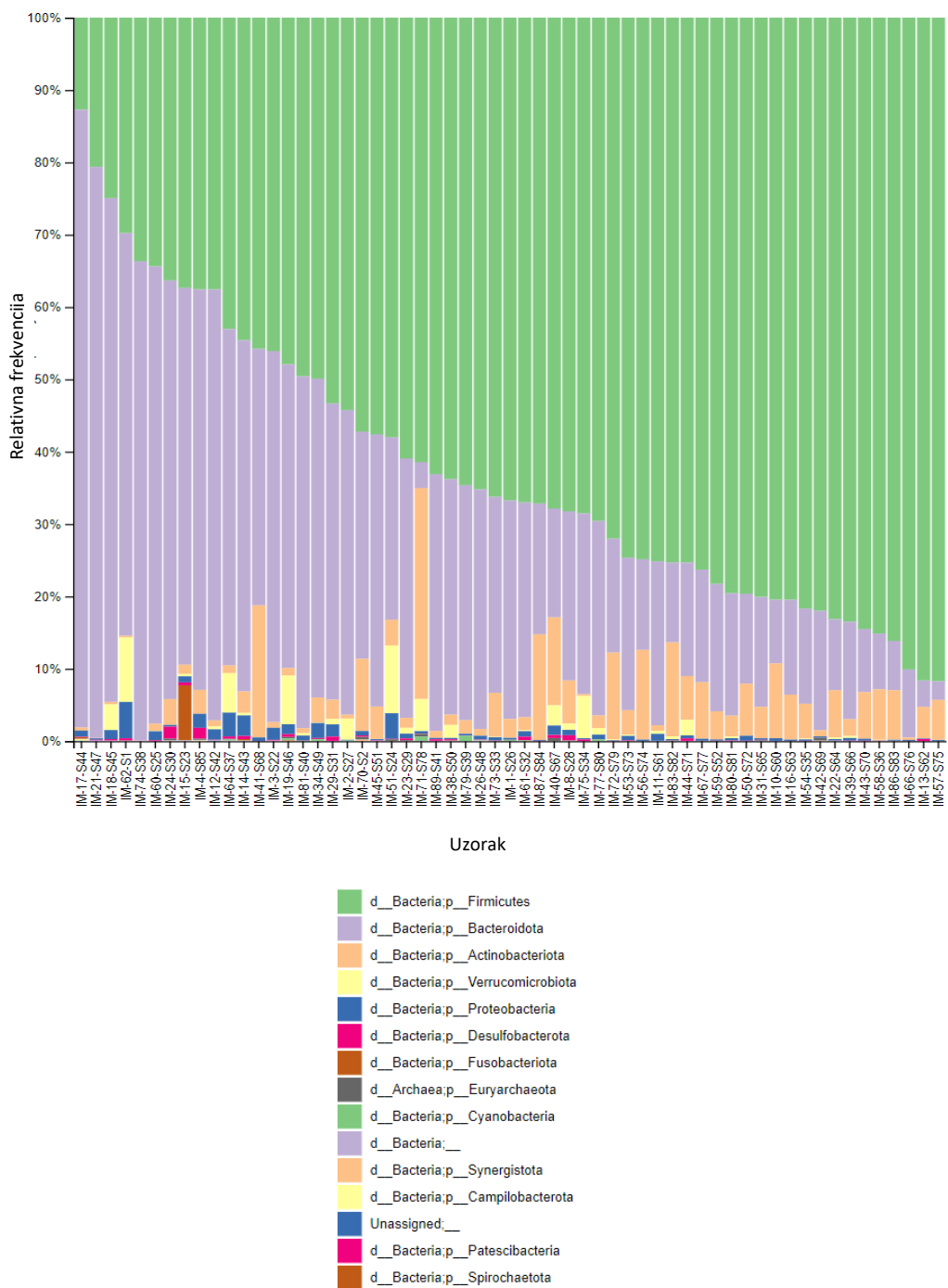


Slika 5. Grafički prikaz frekvencija značajki po uzorku za *single-end* analizu

4.3. TAKSONOMSKA ZASTUPLJENOST MIKROORGANIZAMA PO GRUPAMA

Kao rezultati taksonomske analize u QIIME2 programu generirana je tablica i grafički prikaz u obliku stupčastog dijagrama. Tablica sadrži podatke o taksonomskoj razini svake od 2 409 značajki (sekvenci) koje su analizirane te je svakoj određenoj taksonomiji pridružena pouzdanost dodjele. Vizualno pregledniji izlani artefakt taksonomske analize je stupčasti dijagram u kojem je moguće dobiti uvid u taksonomiju svakog uzorka na taksonomskim razinama od razine 1 koja predstavlja domenu do razine 7 koja predstavlja vrstu. Također podatke je moguće sortirati prema podacima iz meta-podataka korištenih u analizi ili prema taksonomskoj zastupljenosti, padajućim ili rastućim nizom.

Na slici 6 prikazan je stupčasti dijagram taksonomske zastupljenosti bakterija po uzorcima, gledano na taksonomskoj razini koljena. Uzorci su poredani na način da udio koljena *Firmicutes* bude prikazan padajućim nizom.

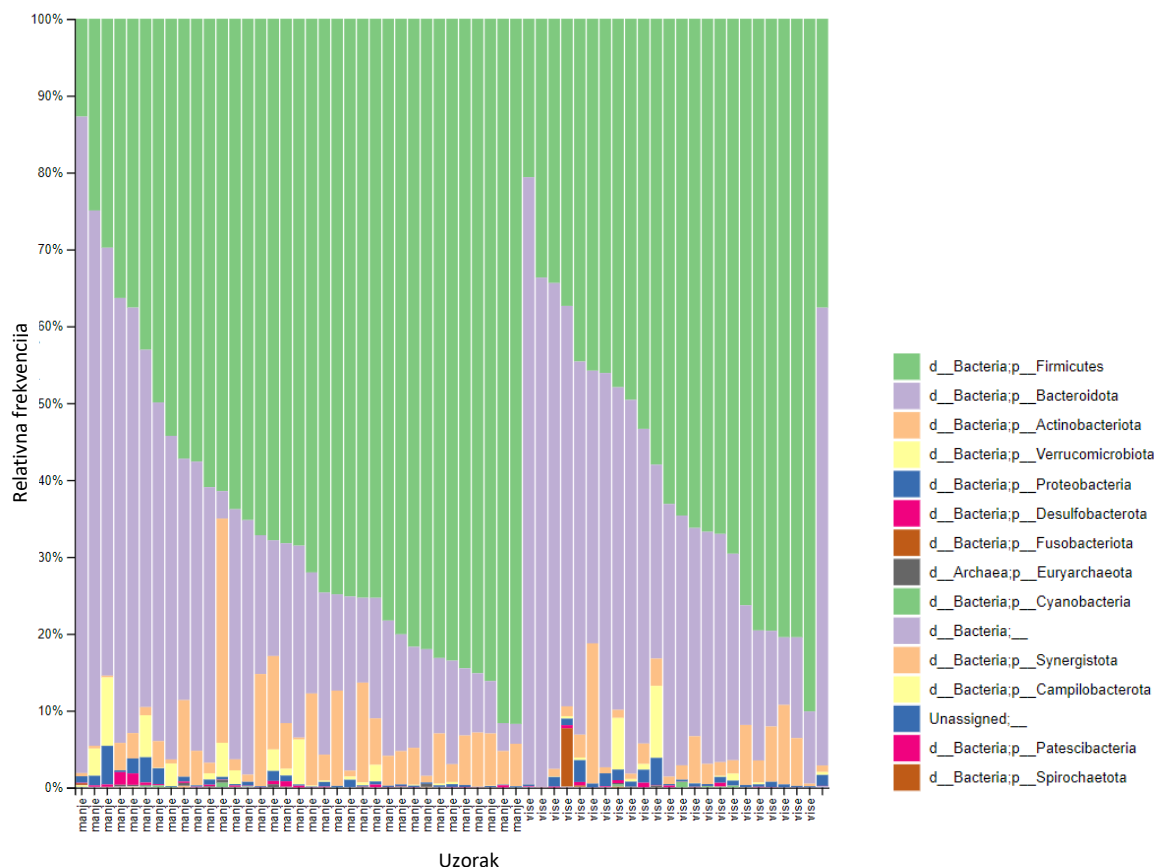


Slika 6. Stupčasti dijagram taksonomske zastupljenosti bakterija na razini koljena po uzorcima

Iz grafičkog prikaza vidljivo je kako najmanju zastupljenost koljena *Firmicutes* ima ispitanik IM-17, a ona iznosi 13 % dok zastupljenost koljena *Bacteroidetes* iznosi 85 %. Ovaj ispitanik, prema učestalosti stolice pripada skupini 'normalno' jer je prijavio 1 stolicu dnevno, a prema konzistenciji stolice pripada grupi s tipom 3 što odgovara normalnoj konzistenciji stolice. Najveću zastupljenost koljena *Firmicutes* ima ispitanik IM-57, a ona iznosi 92 % posto dok kod ovog ispitanika zastupljenost koljena *Bacteroidetes* iznosi 3%. Ispitanik IM-57 i prema učestalosti stolice i prema tipu stolice pripada istim grupama kao ispitanik IM-17. Osim toga, oba su ispitanika ženskoga spola te uzimaju probiotike, dok je ispitanik IM-17 u periodu 3 mjeseca prije prikupljanja uzoraka uzimala antibiotik Augmentin (penicilinski antibiotik – kombinacija amoksicilina i klavulanske kiseline).

Također, može se uočiti kako ispitanik IM-71 ima veću zastupljenost koljena *Actinobacteriota* u odnosu na druge ispitanike, a ona iznosi 29 %, dok je kod ovog ispitanika zastupljenost koljena *Firmicutes* iznosila 61 %, a *Bacteroidetes* 4 %. Ispitanik IM-71 također je ženskoga spola, prema konzistenciji ima tip 1 stolice, a prema učestalost stolice spada u grupu 'rijetko' te ima stolicu svaki drugi dan. Prijavila je korištenje antibiotika Augmentin u periodu 3 mjeseca prije prikupljanja uzorka, a probiotike ne koristi.

Kada se uzorci poredaju prema grupama s manje od jednom ili jednom stolicom dnevno (grupa „manje“) i s jednom do dvije ili više stolica dnevno (grupa „više“), može se uočiti kako u obje grupe postoje ispitanici s visokim i s niskim omjerom *Bacteroidetes:Firmicutes* (slika 7)



Slika 7. Taksonomska zastupljenost bakterija na razini koljena po grupama

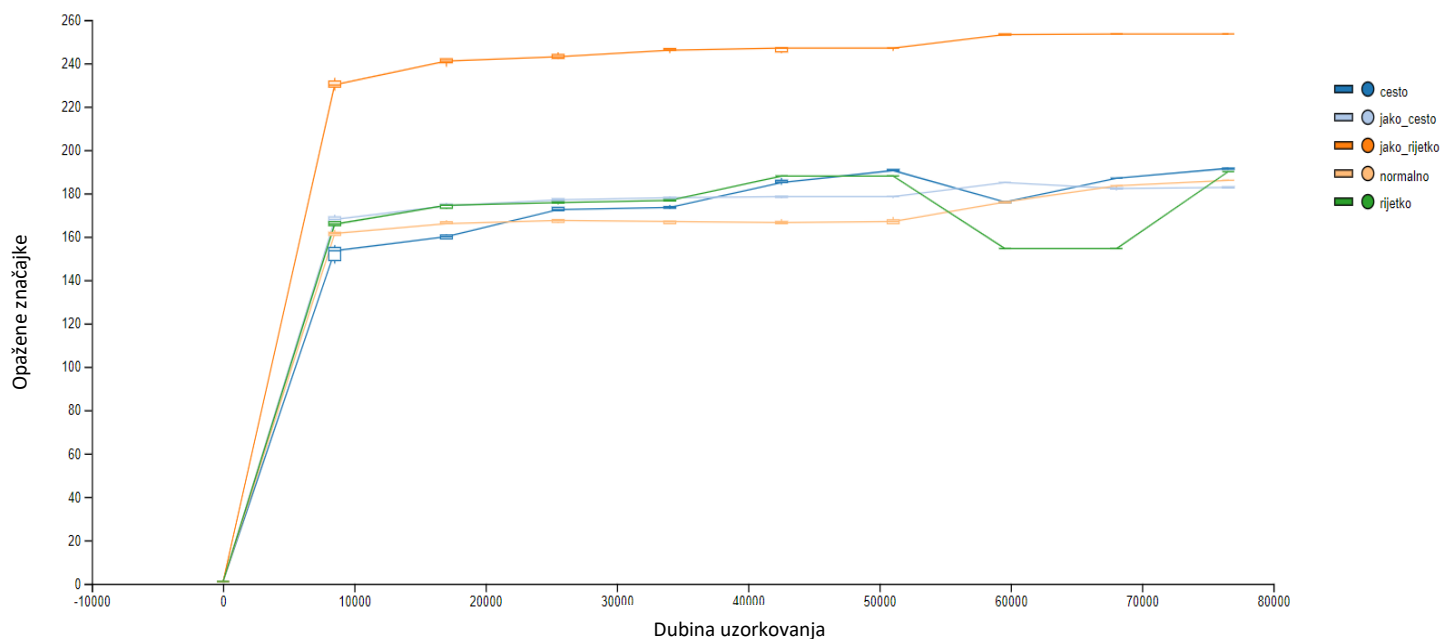
4.4. ALFA RAZNOLIKOST

Alfa raznolikost uzorka je mjera raznolikosti bakterijske populacije unutar jednog uzorka. Izračunom različitih mjera alfa raznolikosti dobiva se uvid u strukturu uzorka s obzirom na bogatstvo, odnosno brojnost vrsta unutar uzorka (broj različitih taksonomskih grupa) i ujednačenost, odnosno ravnomjernost vrsta (distribucija brojnosti tih taksonomskih grupa) ili kombinaciju ova dva svojstva (Willis, 2019).

Za izračun mjera raznolikosti korištena je dubina uzorkovanja od 20 849. Kako bi se provjerilo je li ta dubini uzorkovanja dovoljna kako bi obuhvatila sve prisutne jedinstvene sekvence (značajke) generiran je graf alfa razrijeđenosti. On se generira na način da se za svaku dubinu uzorkovanja, prikazanu na x-osi, uzme poduzorak sekvenci i tako n puta te se izračuna srednja vrijednost broja zapaženih sekvenci i standardna devijacija. One su na grafu prikazane *box-*

plotom. Kada krivulja dosegne plato znači da dodatnim povećanjem dubine uzorkovanja ne bi bile zapažene dodatne sekvence, odnosno da je bogatstvo uzorka u potpunosti uzeto u obzir u analizi.

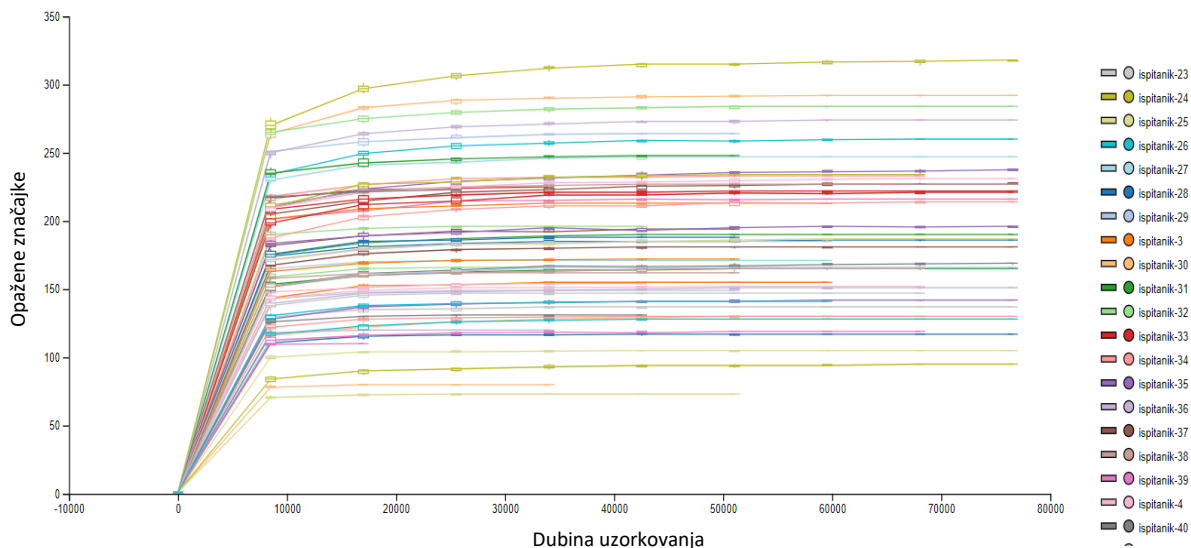
Slika 8 prikazuje graf razrijeđena za opažene značajke na 10 dubina uzorkovanja, a prikazan je za 5 grupa ispitanika ovisno o broju stolica.



Slika 8. Graf alfa razrjeđenja po grupama

Na grafu je vidljivo kako krivulja za grupu „rijetko“ oscilira, a za grupu „normalno“ raste i na najvišoj dubini uzorkovanja. Razlog tomu može biti mali broj uzoraka koji ulaze u sastav grupe pa je samim time prikupljeno malo sekvenci. Druga mogućnost je da u pojedinim uzorcima postoji velik broj rijetkih baterija koje se uočavaju samo na visokim dubinama uzorkovanja. Također je vidljivo kako najveći zapažen broj značajki, odnosno najvišu raznolikost, ima grupa „jako_rijetko“ u koju ulaze ispitanici s najnižim brojem stolica tjedno. To se opažanje podudara s ranije opisanim literaturnim navodima.

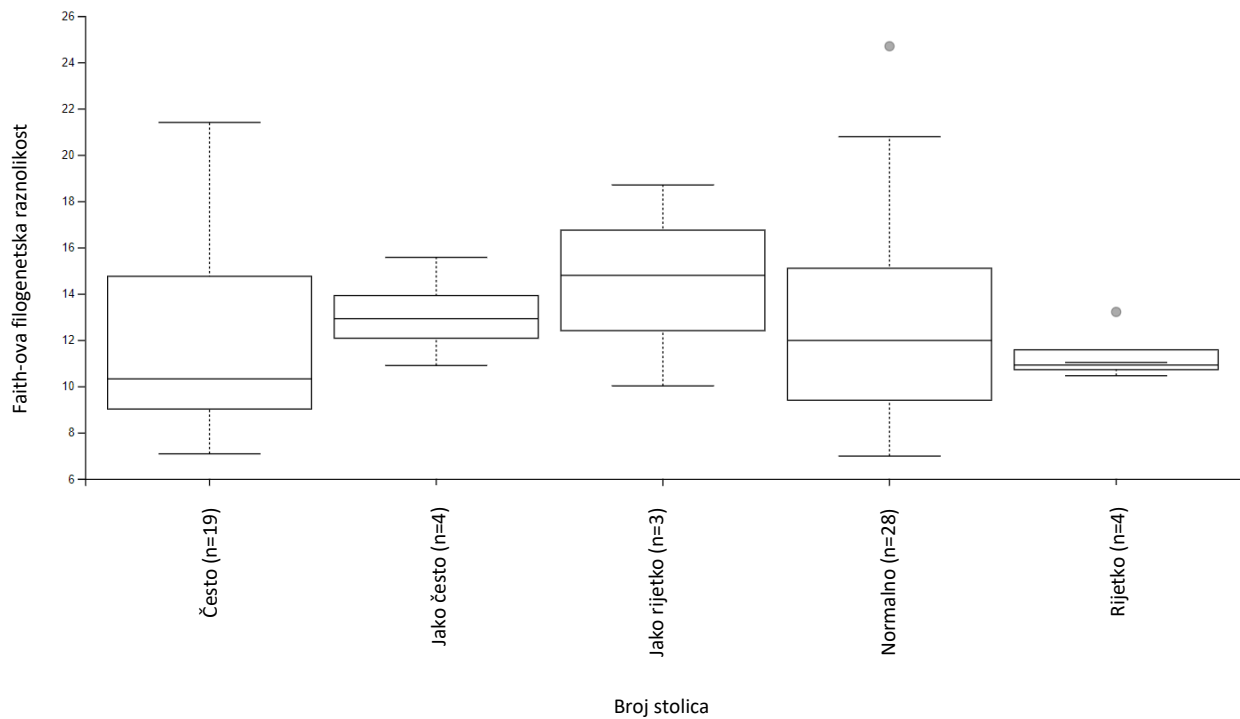
Ukoliko se pogleda graf razrjeđenja za ispitanike prikazan na slici 9, zamjećuje se plato što ukazuje na to da je dubina uzorkovanja bila dovoljna kako bi u analizi raznolikosti bili obuhvaćeni svi prisutni organizmi.



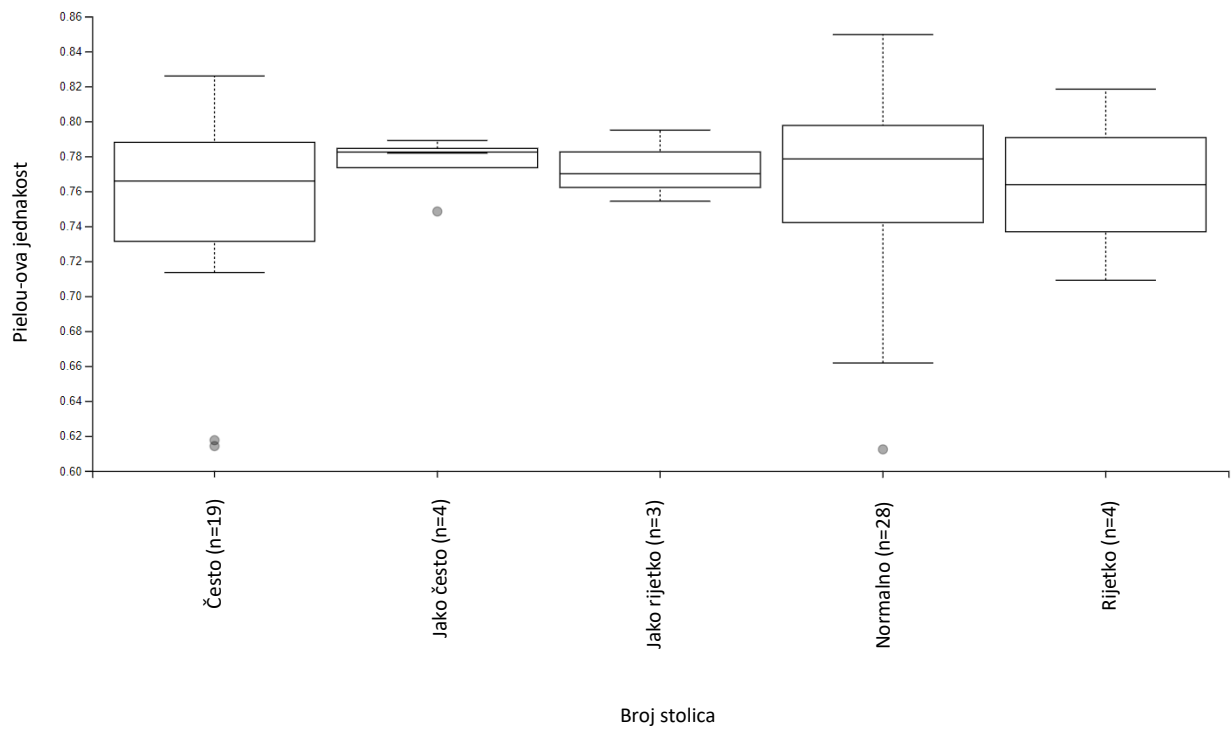
Slika 9. Graf alfa razrjeđenja za ispitanike

Dalje je promatrana značajnost alfa raznolikosti korištenjem metrika Faith-ove filogenetske raznolikosti i Pielou-ove jednakosti. Slika 10 prikazuje *box-plot* Faith-ove filogenetske raznolikosti, a slika 11 *box-plot* Pielou-ove jednakosti po grupama prema prosječnom tjednom broju stolica. Niti jedna od ovih mjera nije pokazala statističku značajnost, p-vrijednost za Faith-ovu raznolikost bila je 0,79, a za Pielou-ovu jednakost 0,95.

Također, nije uočena značajna alfa raznolikost kada se ona promatrala s obzirom na konzistenciju stolice niti kod podjele ispitanika u 2 grupe prema prosječnom dnevnom broju stolica.



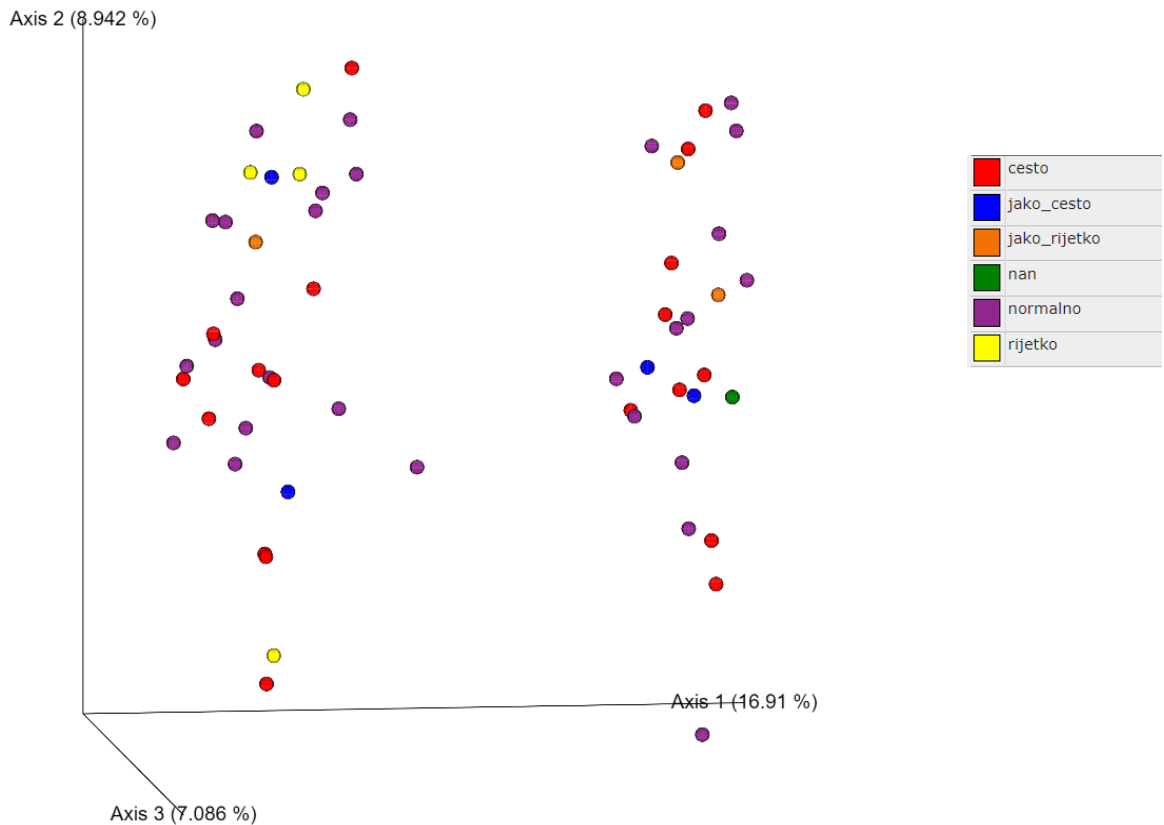
Slika 10. Faith-ova filogenetska raznolikost u odnosu na broj stolica tjedno



Slika 11. Pielou-ova jednakost u odnosu na broj stolica tjedno

4.5. BETA RAZNOLIKOST

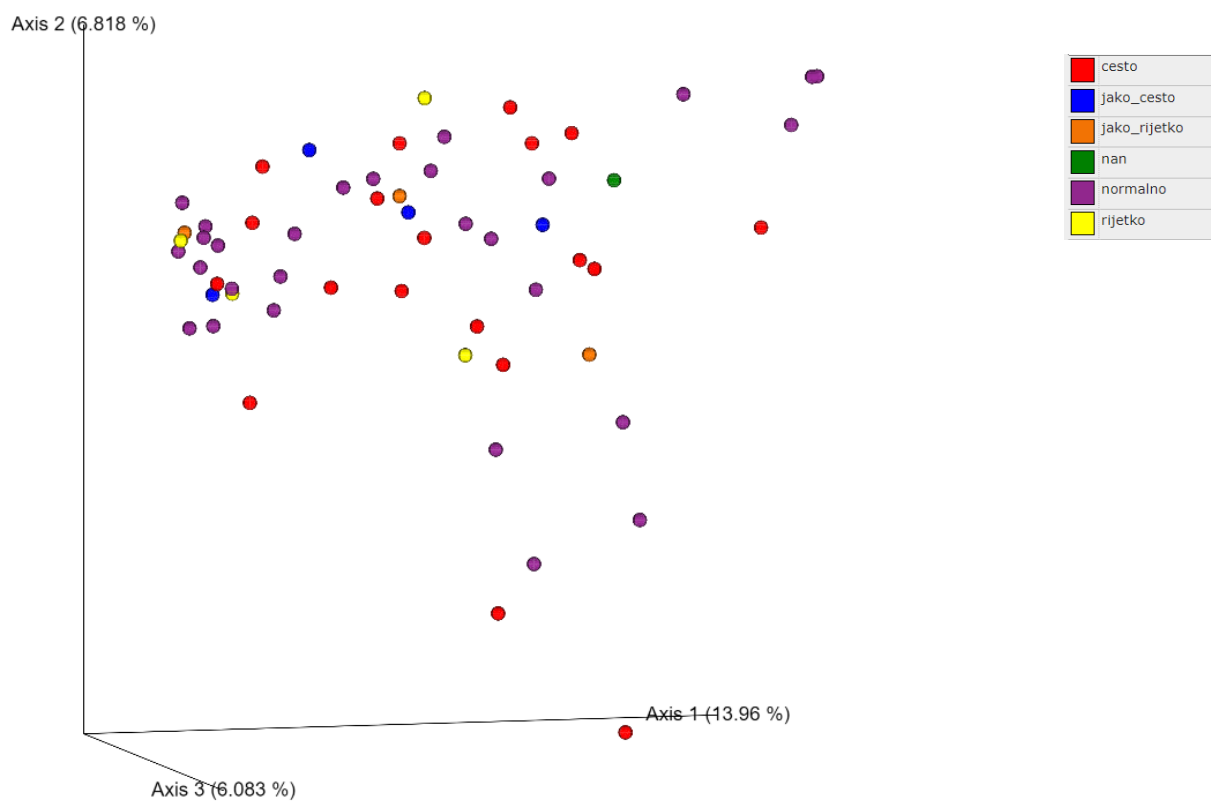
Beta raznolikost je raznolikost između različitih uzoraka. Prikazana je korištenjem PCoA metode i metrike netežinske UniFrac udaljenosti (slika 12).



Slika 12. Netežinska UniFrac udaljenost s obzirom na učestalost stolice

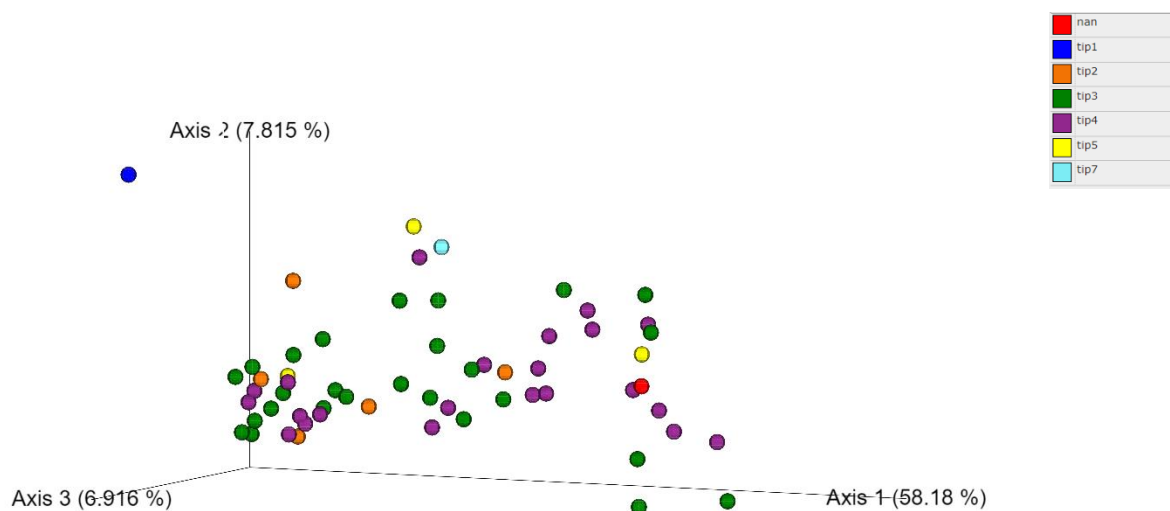
Iz prikaza je vidljivo kako je došlo do grupiranja uzoraka u 2 grupe, međutim ne na temelju učestalosti stolice niti na temelju konzistencije stolice.

Slika 13 prikazuje beta raznolikost PCoA metodom i korištenjem Bray-Curtisove udaljenosti s obzirom na učestalost stolice te je iz prikaza vidljivo kako korištenjem ove metrike nije došlo do grupiranja uzoraka.



Slika 13. Bray-Curtis-ova udaljenost s obzirom na učestalost stolice

Jednako tako, niti korištenjem metrike težinske UniFrac udaljenosti nije došlo do grupiranja uzoraka, a slika 14 prikazuje težinsku UniFrac udaljenost s obzirom na konzistenciju stolice. Iz prikaza je vidljivo kako je najveće odstupanje od svih uzoraka ima uzorak koji pripada grupi tip 1 stolce, međutim zbog neravnomjerne raspodjele uzoraka po grupama na temelju ovoga opažanja nije moguće iznositi zaključke.



Slika 14. Težinska UniFrac udaljenost s obzirom na konzistenciju stolice

PERMANOVA test je proveden kako bi se provjerila statistička značajnost beta raznolikosti. Sve testirane mjere beta raznolikosti nisu pokazale statističku značajnost. Tako je p-vrijednost netežinske UniFrac udaljenosti za podjelu ispitanika u 5 grupa s obzirom na učestalost stolice iznosila 0,577, kada su ispitanici podijeljeni u 2 grupe s obzirom na učestalost stolice p-vrijednost iznosila je 0,164, a p-vrijednost netežinske UniFrac udaljenosti s obzirom na konzistenciju stolice iznosila je 0,764.

Za težinsku UniFrac udaljenost p-vrijednost za podjelu ispitanika u 5 grupa s obzirom na učestalost stolice iznosila je 0,33, a za podjelu ispitanika u 2 grupa s obzirom na učestalost stolice iznosila je 0,087, dok je kod podjele s obzirom na konzistenciju stolice p-vrijednost iznosila 0,17.

P-vrijednost za Bray-Curtisovu udaljenost kod podjele ispitanika u 5 grupa s obzirom na učestalost stolice iznosila je 0,659.

5. ZAKLJUČCI

Na temelju provedenog istraživanja i dobivenih rezultata, može se zaključiti:

1. Metoda analize amplikona jednoga kraja (engl. *single-end*) uspješnija je od metode uparenih krajeva (engl. *paired-end*) kod analize sekvenci niže kvalitete.
2. Nije uočena značajna razlika u sastavu crijevne mikrobiote, u zastupljenosti koljena *Firmicutes* i *Bacteroidetes* niti u njihovom međusobnom omjeru među grupama s obzirom na učestalost i konzistenciju stolice.
3. Nije uočena značajna razlika u bogatstvu i raznolikosti sastava mikorbiote između grupa s obzirom na učestalost i konzistenciju stolice.
4. Analiza beta raznolikosti nije pokazala grupiranje uzoraka na temelju učestalosti ili konzistencije stolice.

Kako bi se sastav crijevne mikrobiote povezao s učestalosti ili konzistencijom stolice potrebno je provesti dodatna istraživanja u kojima će udio ispitanika u pojedinim grupama biti ravnomjerno raspoređen. Također, potrebno je provesti sekvenciranje u uvjetima koji bi rezultirali generiranim podacima bolje kvalitete, koja bi u potpunosti iskoristila mogućnosti amplikon sekvenciranja Illumina MiSeq platforme.

6. LITERATURA

- Allaband C, McDonald D, Vázquez-Baeza Y, Minich JJ, Tripathi A, Brenner DA, i sur. (2019) Microbiome 101: Studying, Analyzing, and Interpreting Gut Microbiome Data for Clinicians. *Clinical Gastroenterol Hepatol* **17**, 218–230. <https://doi.org/10.1016/j.cgh.2018.09.017>
- Alonso VR, Guarner F (2013) Intestinal microbiota composition in adults. U: Guarino A, Quigley EMM, Walker WA (ured.) Probiotic Bacteria and Their Effect on Human Health and Well-Being. Karger AG, Basel, str. 17-24. <https://doi.org/10.1159/000346875>
- Athanasopoulou K, Adamopoulos PG, Scorilas A (2023) Unveiling the Human Gastrointestinal Tract Microbiome: The Past, Present, and Future of Metagenomics. *Biomedicines* **11**, 827. <https://doi.org/10.3390/biomedicines11030827>
- Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, i sur. (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 90. <https://doi.org/10.1186/s40168-018-0470-z>
- Callahan BJ, McMurdie J, Holmes S (2023) Introduction to dada2. Bioconductor.org. <https://www.bioconductor.org/packages/devel/bioc/vignettes/dada2/inst/doc/dada2-intro.html>. Pristupljeno 9. lipnja 2023.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581–583. <https://doi.org/10.1038/nmeth.3869>
- Cao Y, Fanning S, Proos S, Jordan K, Srikumar S (2017) A review on the applications of next generation sequencing technologies as applied to food-related microbiome studies. *Front Microbiol* **8**, 1829. <https://doi.org/10.3389/fmicb.2017.01829>
- Chen Y, Zhou J, Wang L (2021) Role and Mechanism of Gut Microbiota in Human Disease. *Front Cell Infect Microbiol* **11**. <https://doi.org/10.3389/fcimb.2021.625913>
- Čipčić Paljetak H, Barešić A, Panek M, Perić M, Matijašić M, Lojkić I, i sur. (2022) Gut

- microbiota in mucosa and feces of newly diagnosed, treatment-naïve adult inflammatory bowel disease and irritable bowel syndrome patients. *Gut Microbes* **14**. <https://doi.org/10.1080/19490976.2022.2083419>
- de la Cuesta-Zuluaga J, Escobar JS (2016) Considerations For Optimizing Microbiome Analysis Using a Marker Gene. *Front Nutr* **3**, 26. <https://doi.org/10.3389/fnut.2016.00026>
- Donaldson GP, Lee SM, Mazmanian SK (2016) Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* **14**, 20–32. <https://doi.org/10.1038/nrmicro3552>
- Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476–3482. <https://doi.org/10.1093/bioinformatics/btv401>
- Estaki M, Jiang L, Bokulich NA, McDonald D, González A, Kosciolk T, i sur. (2020) QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. *Curr Protoc Bioinformatics* **70**. <https://doi.org/10.1002/cpbi.100>
- Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, i sur. (2016) Population-level analysis of gut microbiome variation. *Science* **352**, 560-564. <https://doi.org/10.1126/science.aad3503>
- Frame LA, Costa E, Jackson SA (2020) Current explorations of nutrition and the gut microbiome: A comprehensive evaluation of the review literature. *Nutr Rev* **78**, 798–812. <https://doi.org/10.1093/nutrit/nuz106>
- Galloway-Peña J, Hanson B (2020) Tools for Analysis of the Microbiome. *Dig Dis Sci* **65**, 674–685. <https://doi.org/10.1007/s10620-020-06091-y>
- Gomaa EZ (2020) Human gut microbiota/microbiome in health and diseases: a review. *Anton Leeuw Int J G* **113**, 2019–2040. <https://doi.org/10.1007/s10482-020-01474-7>
- Hadizadeh F, Walter S, Belheouane M, Bonfiglio F, Heinsen F-A, Andreasson A, i sur. (2017) Stool frequency is associated with gut microbiota composition. *Gut* **66**, 559–560. <https://doi.org/10.1136/gutjnl-2016>

- Hall M, Beiko RG (2018) 16S rRNA Gene Analysis with QIIME2. U: Walker JM (ured.) Methods in Molecular Biology, Humana Press, New York, str. 113–129. https://doi.org/10.1007/978-1-4939-8728-3_8
- Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN (2015) Role of the normal gut microbiota. *World J Gastroenterol* **21**, 8836–8847. <https://doi.org/10.3748/wjg.v21.i29.8787>
- Jo JH, Kennedy EA, Kong HH (2016) Research Techniques Made Simple: Bacterial 16S Ribosomal RNA Gene Sequencing in Cutaneous Research. *J Invest Dermatol* **136**, 23–27. <https://doi.org/10.1016/j.jid.2016.01.005>
- Kwon HJ, Lim JH, Kang D, Lim S, Park SJ, Kim JH (2019) Is stool frequency associated with the richness and community composition of gut microbiota? *Intest Res* **17**, 419–426. <https://doi.org/10.5217/ir.2018.00149>
- Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, i sur. (2012) Microbial culturomics: Paradigm shift in the human gut microbiome study. *Clin Microbiol Infec* **18**, 1185–1193. <https://doi.org/10.1111/1469-0691.12023>
- Lee YY, Erdogan A, Rao SSC (2014) How to assess regional and whole gut transit time with wireless motility capsule. *J Neurogastroenterol* **20**, 265–270. <https://doi.org/10.5056/jnm.2014.20.2.265>
- Manor O, Dai CL, Kornilov SA, Smith B, Price ND, Lovejoy JC, i sur. (2020) Health and disease markers correlate with gut microbiome composition across thousands of people. *Nat Commun* **11**, 5206. <https://doi.org/10.1038/s41467-020-18871-1>
- Nandhra GK, Mark EB, Di Tanna GL, Haase AM, Poulsen J, Christodoulides S, i sur. (2020) Normative values for region-specific colonic and gastrointestinal transit times in 111 healthy volunteers using the 3D-Transit electromagnet tracking system: Influence of age, gender, and body mass index. *J Neurogastroenterol* **32**, 13734. <https://doi.org/10.1111/nmo.13734>
- Pargin E, Roach MJ, Skye A, Papudeshi B, Inglis LK, Mallawaarachchi V, i sur. (2023) The human gut virome: composition, colonization, interactions, and impacts on human health.

Front Microbiol **14**. <https://doi.org/10.3389/fmicb.2023.963173>

Parthasarathy G, Chen J, Chen X, Chia N, O'Connor HM, Wolf PG, i sur. (2016) Relationship between Microbiota of the Colonic Mucosa vs Feces and Symptoms, Colonic Transit, and Methane Production in Female Patients with Chronic Constipation. *Gastroenterology* **150**, 367-379. <https://doi.org/10.1053/j.gastro.2015.10.005>

Pervez MT, Hasnain MJU, Abbas SH, Moustafa MF, Aslam N, Shah SSM (2022) A Comprehensive Review of Performance of Next-Generation Sequencing Platforms. *Biomed Res Int* **2022**, 3457806. <https://doi.org/10.1155/2022/3457806>

Plummer E, Twin J, Bulach DM, Garland SM, Tabrizi SN (2015) A Comparison of Three Bioinformatics Pipelines for the Analysis of Preterm Gut Microbiota using 16S rRNA Gene Sequencing Data. *J Proteomics Bioinform Citation* **8**, 283–291. <https://doi.org/10.4172/0974-276X.1000381>

Procházková N, Falony G, Dragsted LO, Licht TR, Raes J, Roager HM (2023) Advancing human gut microbiota research by considering gut transit time. *Gut* **72**, 180–191. <https://doi.org/10.1136/gutjnl-2022-328166>

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, i sur. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65. <https://doi.org/10.1038/nature08821>

Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggianno GAD, Gasbarrini A, i sur. (2019) What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. *Microorganisms* **7**, 14. <https://doi.org/10.3390/microorganisms7010014>

Rowland I, Gibson G, Heinken A, Scott K, Swann J, Thiele I, i sur. (2018) Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr* **57**, 1-24. <https://doi.org/10.1007/s00394-017-1445-8>

Schmidt TSB, Raes J, Bork P (2018) The Human Gut Microbiome: From Association to Modulation. *Cell* **172**, 1198–1215. <https://doi.org/10.1016/j.cell.2018.02.044>

- Shakya M, Lo CC, Chain PSG (2019) Advances and challenges in metatranscriptomic analysis. *Front Genet* **10**. <https://doi.org/10.3389/fgene.2019.00904>
- Štimac I, Martinković F (2021) Uvod u tehnologije sekvenciranja novih generacija. *Hrvatski veterinarski vjesnik* **29**, 37–47. <https://hrcak.srce.hr/263646>
- The Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214. <https://doi.org/10.1038/nature11234>
- Tigchelaar EF, Bonder MJ, Jankipersadsing SA, Fu J, Wijmenga C, Zhernakova A (2015) Gut microbiota composition associated with stool consistency. *Gut* **65**, 540–542. <https://doi.org/10.1136/gutjnl>
- Tropini C, Earle KA, Huang KC, Sonnenburg JL (2017) The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host Microbe* **21**, 433–442. <https://doi.org/10.1016/j.chom.2017.03.010>
- Vaga S, Lee S, Ji B, Andreasson A, Talley NJ, Agréus L, i sur. (2020) Compositional and functional differences of the mucosal microbiota along the intestine of healthy individuals. *Sci Rep* **10**. <https://doi.org/10.1038/s41598-020-71939-2>
- Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J (2016) Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57–62. <https://doi.org/10.1136/gutjnl-2015-309618>
- Vork L, Penders J, Jalanka J, Bojic S, van Kuijk SMJ, Salonen A, i sur. (2021) Does Day-to-Day Variability in Stool Consistency Link to the Fecal Microbiota Composition? *Front Cell Infect Microbiol* **11**. <https://doi.org/10.3389/fcimb.2021.639667>
- Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y (2020) Host variables confound gut microbiota studies of human disease. *Nature* **587**, 448–454. <https://doi.org/10.1038/s41586-020-2881-9>
- Willis AD (2019) Rarefaction, alpha diversity, and statistics. *Front Microbiol* **10**. <https://doi.org/10.3389/fmicb.2019.02407>
- Zoetendal EG, Von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans ADL, De Vos

WM (2002) Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol* **68**, 3401–3407. <https://doi.org/10.1128/AEM.68.7.3401-3407.2002>

PRILOZI

PRILOG 1: UPITNIK PILOT PROJEKTA „CRIJEVNA MIKROBIOTA RADNO AKTIVNE POPULACIJE RH“

Pilot projekt: Crijevna mikrobiota radno-aktivne populacije RH

Istraživanje provode Institut za Antropologiju i Prehrambeno-biotehnološki fakultet Sveučilišta u Zagrebu s ciljem određivanja sastava crijevne mikrobiote u zdravoj, radno-aktivnoj populaciji stanovništva RH koja bi se koristila kao kontrolna skupina u budućim istraživanjima. Istraživanje je odobrilo Etičko povjerenstvo Instituta za antropologiju. Istraživanje će biti provedeno u Republici Hrvatskoj. Pružene informacije će se koristiti pri analizi Vašeg mikrobioma te se mogu koristiti sada i u budućnosti u znanstvenim publikacijama i dodatnim istraživanjima. Ispitivanjem ovog upitnika pristajem pružiti minimalne potrebne informacije bez otkrivanja vlastitog identiteta i jamčim da neću krivotvoriti informacije.

Crijevna mikrobiota je zajednica mikroorganizama (bakterije, arheje, gljive i virusi) koja živi u našim crijevima. Ona se mijenja tijekom našeg života ovisno o prehrani, aktivnosti, zdravlju i mnogim drugim faktorima s kojima se svatko od nas u životu susreće i jedinstvena je u svakom pojedincu. Novija istraživanja sve veću važnost pridaju mikrobioti jer je "zdrava" mikrobiota povezana s pozitivnim zdravstvenim ishodima, dok se u slučaju "nezdrave", odnosno neuravnotežene mikrobiote, ona povezuje s bolestima u rasponu od sindroma iritabilnog crijeva preko alergija do karcinoma. Zdrava crijevna mikrobiota jedan je od faktora koji pojedincu osiguravaju zdrav i aktivan život.

* Required

1. Molimo unesite šifru s posudice za uzimanje uzorka stolice: *

2. Spol *

Mark only one oval.

- Muški
 Ženski

3. Dob *

Mark only one oval.

- 18 - 29
 30 - 39
 40 - 49
 50 - 59

4. Bračni status *

Mark only one oval.

- oženjen/udana
- živim u izvanbračnoj zajednici
- rastavljen/a
- slobodan/na
- ne želim odgovoriti

5. Obrazovanje *

Mark only one oval.

- Osnovna škola
- Srednja škola
- Viša škola
- Prediplomski studij
- Diplomski studij
- Doktorat
- ne želim odgovoriti

6. Visina u centimetrima: *

7. Masa u kilogramima: *

8. Na koji način ste rođeni? *

Mark only one oval.

- vaginalni porod
- carski rez
- ne znam

9. Jeste li uzimali antibiotike u zadnja 3 mjeseca? *

Mark only one oval.

DA

NE

10. Ako da, koji?

11. Jeste li tjelesno aktivni? *

Mark only one oval.

DA

NE

12. Koliko ste tjelesno aktivni na tjednoj bazi? *

Mark only one oval.

Manje od 1 sata

1 do 2 sata

2 do 3 sata

3 do 4 sata

Više od 4 sata

13. Kojim tjelesnim aktivnostima se bavite? *

Check all that apply.

Hodanje

Trčanje

Biciklizam

Plivanje

Društveni sportovi (nogomet, košarka, odbojka, hokej...)

Fitness, aerobik, pilates, joga i sl.

Other: _____

14. Pušite li? *

Mark only one oval.

- Da
- Ne
- Bivši sam pušač/ica

15. Moja prehrana se općenito može opisati kao prehrana: *

Mark only one oval.

- svejeda
- vegetarijanca
- vegana
- za posebne potrebe (dijabetičare, bubrežne bolesnike, srčane bolesnike...)
- Other: _____








16. Prosječna učestalost stolice? *

Mark only one oval.

- Više od 2 puta dnevno
- Jednom do 2 puta dnevno
- Jednom dnevno
- Svaki drugi dan
- Manje od 3 puta tjedno
- Patim od zatvora (manje od 2 stolice tjedno)

17. Izaberite uobičajeni tip stolice *

The Bristol Stool Form Scale

1		Odvojene tvrde kuglice (poput lješnjaka)
2		Poput kobasice, ali sa jasno vidljivim kuglicama
3		Poput kobasice sa raspucanom površinom
4		Poput kobasice, glatko
5		Mekane grudice s jasno raspoznatljivim rubovima
6		Komadići s nejasnim rubovima, konzistencije kaše
7		U potpunosti vodeno, bez ikakvih čvrstih dijelova

Mark only one oval.

- Tip 1
- Tip 2
- Tip 3
- Tip 4
- Tip 5
- Tip 6
- Tip 7

Životni stil i zdravlje

18. Za sebe smatrate da ste: *

Mark only one oval.

- Premršav/a
- Normalne tjelesne mase
- Imam nekoliko kilograma viška
- Pretio/la

19. Smatrate li da se debljate brzo? *

Mark only one oval.

- Ne
- Da
- Samo ako dulje vrijeme previše jedem

20. Želite li biti u boljoj formi? *

Mark only one oval.

- Da
- Ne

21. Što radite da biste poboljšali zdravstveno stanje ili smanjili tjelesnu masu? *

Check all that apply.

- Pokušavam se hraniti pravilnije
- Pazim na veličinu porcija
- Radim vježbe izdržljivosti (trčanje, hodanje, biciklizam, ples...)
- Radim vježbe za jačanje i tonus mišića (sklekovi, dizanje utega, čučnjevi, trbušnjaci..)
- Na restriktivskoj dijeti sam
- Ne radim ništa

22. Ako želite poboljšati formu, a ne uspijevate koji su razlozi koji Vas sprečavaju? *

Check all that apply.

- Stres
- Nedostatak vremena
- Nedostatak potpore okoline
- Financijski razlozi
- Nedostatak informacija
- Nedostatak upornosti

Other: _____

23. Koliko prosječno spavate dnevno? *

Mark only one oval.

- Manje od 5 sati
- Između 5 i 7 sati
- Više od 7 sati

24. Koliko prosječno jedete obroka dnevno? *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5 i više

25. Koliko često doručkujete? *

Mark only one oval.

- svaki dan
- ponekad
- rijetko
- nikad

26. Za ručak na poslu/školi: *

Check all that apply.

- donosite ručak od doma
- jedete hranu iz menze
- jedete brzu hranu (fast food)
- jedete grickalice
- jedete slatkiše
- preskačete ručak

Other: _____

27. Koristite li umjetne zaslađivače? *

Podrazumijevaju se saharin (E954), aspartam (E951), ciklambat (E952) i dr. Na deklaraciji prehrambenih proizvoda nalaze se istaknuti tekstualno ili kao E-brojevi

Mark only one oval.

- Da
- Ne
- Zasebno ne, ali pijem/jedem proizvode koji ih sadržavaju

28. Koliko ste često naduti? *

Mark only one oval.

- Nikad
- 1 mjesečno
- 2 - 3 puta mjesečno
- 1 tjedno
- 2 - 3 puta tjedno
- Svaki dan

29. Uzimate li neke od sljedećih lijekova? *

Check all that apply.

- Hormonalna terapija
- Analgetici
- Antidepresivi
- Dodaci prehrani
- Statini (kolesterol)
- Ništa od navedenog

Other: _____

30. Bolujete li od sljedećih bolesti/poremećaja: *

Check all that apply.

- Artritis
- Depresija
- Problemi sa spavanjem
- Astma
- Srčane bolesti
- Visok krvni tlak
- Dijabetes
- Sindrom iritabilnog crijeva
- Crohnova bolest
- Ulcerozni kolitis
- Celijakija
- Bolesti jetre
- Hepatitis
- Ništa od navedenog

31. Kakav Vam je COVID-19 status? *

Check all that apply.

- Prebolio/la
- Cijepljen/na
- Ništa od navedenog
- Ne želim odgovoriti

32. Alergičan/na sam na: *

Check all that apply.

- Pelud
- Orašaste plodove (badem, lješnak, orah, brazilski oraščić, indijski oraščić, pistacije...)
- Kikiriki
- Voće (jagoda, kruška, banana, breskva, limun, naranča, ananas, jabuka)
- Povrće (rajčica, špinat, kupus, paprika)
- Meso (svinjetina, govedina, piletina, puretina)
- Ribe, rakovi, školjke i ostali mekušci
- Žitarice s glutenom (pšenica, raž, ječam, pir..)
- Žitarice bez glutena (riža, kukuruz, proso, heljda...)
- Jaja i proizvode od jaja
- Mlijeko i mliječne proizvode
- Celer
- Lupina/vučji bob
- Soja i proizvodi od soje
- Sezam
- Sumporni dioksid i sulfiti
- Goruščica
- Nemam alergiju

Other: _____

33. Uzimate li probiotike? *

Mark only one oval.

- Da
- Ne
- Ne kao dodatke prehrani, ali pijem probiotičke jogurte

34. Uzimate li dodatke prehrani? *

Check all that apply.

- Ne
- Da, vitamine (C, E, multivitamine...)
- Da, minerale (magnezij, cink, kalcij...)
- Da, amino kiseline (BCAA, karnitin, glutamin)
- Da, proteinske dodatke (proteine sirutke, soje, konoplje...)
- Da, vlakna (prebiotike)
- Da, biljne suplemente (ginko, gospina trava, valerijana...)
- Omega 3 masne kiseline
- Other: _____

Prehrambene navike

35. Koliko često konzumirate kavu? *

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

36. Koliko često konzumirate čaj? *

Pod čaj ubrajamo crni, zeleni, bijeli, žuti i oolong tip čaja

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

37. Koliko često konzumirate jedno serviranje alkoholnog pića? *

Jedno serviranje alkoholnog pića odgovara 30 mL žestokog pića, 150 mL vina i 330 mL piva

Mark only one oval.

- Nikad
- Manje od 1 mjesečno
- 1 - 3 puta mjesečno
- 1 tjedno
- 2 - 3 puta tjedno
- 4 - 6 puta tjedno
- 1 dnevno
- Više puta dnevno

38. Koliko često konzumirate 100-postotne prirodne sokove? *

Prirodni sok dobiven je cijedenjem voća/povrća uz dodatak vode i bez dodanog šećera

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

39. Koliko često konzumirate gazirana bezalkoholna pića? *

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

40. Koliko često konzumirate mlijeko i mliječne proizvode? *

Pod mlijeko ubrajamo sva mlijeka životinjskog porijekla (kravlje, kozje, ovčje...) i njihove prerađevine (sir, jogurt, vrhnja...), a ne uključujemo mlijeka dobivena iz biljnih sirovina (sojino, bademovo, kokosovo...)

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

41. Koliko često konzumirate voće? *

Odnosi se isključivo na svježe, neprerađeno voće

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

42. Koliko često konzumirate povrće? *

Ne uključuje krumpir

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

43. Koliko često konzumirate meso? *

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

44. Koliko često konzumirate ribu? *

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

45. Koliko često konzumirate kruh i pekarske proizvode? *

Mark only one oval.

- Više puta dnevno
- Jednom dnevno
- Nekoliko puta tjedno
- Jednom tjedno
- Nekoliko puta mjesečno
- Jednom mjesečno
- Nekoliko puta godišnje
- Ne konzumiram

46. Što ste jučer jeli za doručak? *

Molimo navedite namirnice i procijenjenju količinu (npr. 2 kriške kruha s maslacem i marmeladom, zdjelicu mlijeka s čokoladnim rižinim pahuljicama..)

47. Što ste jučer jeli za ručak? *

Molimo navedite namirnice i procijenjenju količinu (npr. tanjur bistre juhe s tjesteninom, zagrebački odrezak s krumpirom, zdjelica zelene salate..)

48. Što ste jučer jeli za večeru? *

Molimo navedite namirnice i procijenjenju količinu (npr. 2 jaja na oko i 2 kajzerice...)

49. Što ste prekjučer jeli za doručak? *

Molimo navedite namirnice i procijenjenju količinu (npr. 2 kobasice i 2 jaja na oko..)

50. Što ste prekjučer jeli za ručak? *

Molimo navedite namirnice i procijenjenju količinu (npr. porciju variva od graška s noklicama i kobasicom)

51. Što ste prekjučer jeli za večeru? *

Molimo navedite namirnice i procijenjenju količinu (npr. tanjur tjestenine s umakom od tune i rajčice)

52. Koristite li maslinovo ulje kao glavni izvor masnoća u kuhanju? *

Mark only one oval.

DA

NE

53. Procijenite koliko kuhinjskih žlica maslinovog ulja konzumirate dnevno (uključuje ulje za prženje, u salatama, kad jedete vani...) *

54. Koliko porcija povrća jedete dnevno (porcija = 200 g) ? *
-
55. Koliko porcija voća jedete dnevno (porcija = 200 g), uključujući prirodne sokove ? *
-
56. Koliko porcija crvenog mesa ili mesnih prerađevina (šunka, kobasice...) jedete dnevno (porcija = 100-150 g) ?
-
57. Koliko porcija maslaca, margarina ili vrhnja jedete dnevno (porcija = 12 g) ?
-
58. Koliko zaslađenih ili gaziranih napitaka pijete dnevno?
-
59. Koliko čaša vina pijete tjedno?
-
60. Koliko porcija mahunarki jedete tjedno (porcija = 150g) ? *
-
61. Koliko porcija ribe ili školjkaša jedete tjedno (riba porcija = 100-150g, školjkaši porcija = 200g) ? *
-
62. Koliko kupovnih slatkiša ili kolača jedete tjedno? *
-
63. Koliko porcija orašastih plodova (uključujući kikiriki) jedete tjedno (porcija = 30g) ? *

64. Volite li više jesti piletinu, puretinu i zečevinu od junetine, svinjetine i kobasica? *

Mark only one oval.


Da

Ne

65. Koliko puta tjedno jedete povrće, tjesteninu ili rižu u umaku od rajčice? *

IZJAVA O IZVORNOSTI

Ja, Ivana Eranović, izjavljujem da je ovaj diplomski rad izvorni rezultat mojeg rada te da se u njegovoj izradi nisam koristila drugim izvorima, osim onih koji su u njemu navedeni.



Vlastoručni potpis