Utjecaj metode izolacije DNA na određen sastav intestinalne mikrobiote

Ćutić, Maja

Master's thesis / Diplomski rad

2017

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://urn.nsk.hr/urn:nbn:hr:159:806794

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-04-19**

prehrambeno biotehnološki fakultet

Repository / Repozitorij:

Repository of the Faculty of Food Technology and Biotechnology

zir.nsk.hr

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U ZAGREBU
PREHRAMBENO-BIOTEHNOLOŠKI FAKULTET

DIPLOMSKI RAD

Zagreb, travanj 2017.

Maja Ćutić
706/N

UTJECAJ METODE IZOLACIJE DNA NA ODREĐEN SASTAV INTESTINALNE MIKROBIOTE

Rad je izrađen u Kabinetu za bioinformatiku na Zavodu za biokemijsko inženjerstvo Prehrambeno-biotehnološkog fakulteta Sveučilišta u Zagrebu pod mentorstvom dr.sc. Jurice Žučka, doc. Prehrambeno-biotehnološkog fakulteta Sveučilišta u Zagrebu.

Veliko hvala mome mentoru, dr.sc. Jurici Žučku, doc. prije svega na strpljenju i nesebičnoj pomoći pri izradi ovog rada, na pozitivnom stavu i uvijek vrijednim savjetima.

Hvala kolegicama Marini, Steli i Teni na održavanju dobre radne atmosfere, motivaciji i podršci.

Hvala Vjeranu na podršci, ljubavi i konstantnom ohrabrenju.

Najveće hvala mojim roditeljima i bratu Tomislavu, bez čije potpore i razumijevanja moje obrazovanje i napredak ne bi bili mogući.

TEMELJNA DOKUMENTACIJSKA KARTICA

Diplomski rad

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

Znanstveno područje: Biotehničke znanosti
Znanstveno polje: Nutricionizam

UTJECAJ METODE IZOLACIJE DNA NA ODREĐEN SASTAV INTESTINALNE MIKROBIOTE

Maja Ćutić, 706/N

Sažetak: Crijevna mikrobiota predstavlja relativno novo, zanimljivo područje istraživanja povezano s mnogim stanjima ljudskog organizma, od održavanja homeostaze preko povremenih gastrointestinalnih poremećaja, do povezanosti s bolestima. Povezanost s pretilošću omogućava bolji uvid u prevenciju, razvoj i liječenje ove bolesti čija incidencija kontinuirano raste. Kako bi se bolje razumjela ta povezanost, koriste se suvremene metode analize koje uključuju sekvencioniranje bakterijskog 16S rRNA gena iz uzorka fecesa. Cilj ovog rada bio je usporediti tri metode za izolaciju DNA iz uzorka fecesa pretile osobe kako bi se vidjele razlike u dobivenim podacima i utjecaj ovog pretkoraka na daljnje analize. Također, promatrao se utjecaj dodavanja DNA Shield™ stabilizatora. Obradom podataka pomoću QIIME bioinformatičkog alata, dobiveni su podaci o DNA prinosu i bioraznolikosti uzoraka u ovisnosti o metodi izolacije. Za daljna istraživanja poželjno je koristiti jednu metodu kako bi se smanjila pogreška odnosno dobili sastavi mikrobiote koji su međusobno usporedivi.

Ključne riječi: crijevna mikrobiota, pretilost, metode izolacije DNA, 16S rRNA sekvencioniranje, QIIME program

Rad sadrži: 57 stranica, 24 slike, 4 tablica, 53 literaturna navoda, 00 priloga

Jezik izvornika: hrvatski

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

Mentor: doc.dr.sc. Jurica Žučko

Stručno povjerenstvo za ocjenu i obranu:
1.izv.prof.dr.sc. Antonio Starčević
2. doc.dr.sc. Jurica Žučko
3. doc.dr.sc. Ksenija Uroić
4. izv.prof.dr.sc. Jasna Novak (zamjena)

Datum obrane: 09. svibnja 2017.

BASIC DOCUMENTATION CARD

Graduate Thesis

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

Scientific area: Biotechnical Sciences
Scientific field: Nutrition

THE EFFECT OF DNA EXTRACTION METHOD ON THE GUT MICROBIOTA COMPOSTION

Maja Ćutić 706/N

Abstract: Human gut microbiota is rapidly becoming one of the most interesting fields of studies regarding health and disease development. Emerging evidences show relationship between gut microbiota and obesity, one of the fastest delevoping diseases worldwide. To better understand this connection, methods such as 16S rRNA sequencing are applied. The aim of this study was to compare three available methods of DNA extraction that were used to extract DNA from faeces of a obese person. Data collected from sequencing were processed using QIIME bioinformatical tool in order to observe differences in bacterial diversity and DNA yield depending on DNA extraction method. The influence of applying DNA Shield™ reagens was also observed. Io order to decrease biases in studies regarding analysis of the gut microbiota, it is recommended to use one specific method of DNA extraction.

Keywords: gut microbiota, obesity, DNA isolation methods, 16S rRNA sequencing, QIIME pipeline

Thesis contains: 57 pages, 24 figures, 4 tables, 53 references, 00 supplements

Original in: Croatian

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

Mentor: PhD, Jurica Žučko, Assistant professor

Reviewers:
1. PhD. Antonio Starčević, Associate professor
2. PhD, Jurica Žučko, Assistant professor
3. PhD. , Ksenija Uroić, Assistant professor
4. PhD. , Jasna Novak, Associate professor (substitute)

Thesis defended: May 9th, 2017

Sadržaj

1. UVOD 1
2. TEORIJSKI DIO 3
2.1. HUMANA CRIJEVNA MIKROBIOTA 3
2.2. POVEZANOST MIKROBIOTE I PRETILOSTI 5
2.3 UTJECAJ METODE IZOLACIJE DNA NA ODREĐIVANJE SASTAVA INTESTINALNE MIKROBIOTE (PRETILIH OSOBA) 9
2.4. ODREĐIVANJE SASTAVA MIKROBIOTE KORIŠTENJEM MARKER GENA 10
2.5. ILLUMINA MISEQ, QIIME PROGRAM I PAIRED-END SEKVENCIONIRANJE 11
3. EKSPERIMENTALNI DIO 17
3.1 PROTOKOL ZA IZOLACIJU DNA 17
3.1.1.Izolacija: Maxwell® 16 Tissue DNA Purification Kit 17
3.1.2. Izolacija: Quick-DNA™ Fecal/Soil Microbe Miniprep Kit 18
3.1.3. Izolacija Stool DNA Isolation Kit (Cat. 27600) Norgen Biotek Corp 19
3.2. PROTOKOL ZA SEKVENCIONIRANJE DNA 20
3.3. QIIME 22
3.3.1. Pre-procesiranje 23
3.3.2. Analize raznolikosti 25
3.3.3. Alfa raznolikost 27
3.3.4. Beta raznolikost 29
3.3.5. Skripta ″core diversity″ 30
3.3.6. Calypso web aplikacija 31
4. REZULTATI I RASPRAVA 32
4.1. ODREĐIVANJE TAKSONOMIJE 33
4.2. ALFA RAZNOLIKOST 39
4.3. BETA RAZNOLIKOST 46
4.4. STATISTIČKA ANALIZA 48
5. ZAKLJUČCI 53
6. LITERATURA 54

1. UVOD

Mikroorganizmi nastanjuju sva prirodna staništa pa tako i druge (više) organizme uključujući i čovjeka. Spoznajom o njihovoj prisutnosti u ljudskom organizmu te različitim utjecajima na nj, povoljnim i nepovoljnim, povećava se potreba za istraživanjem tih odnosa. Kako bi se razumio način istraživanja i rezultati, važno je razlikovati pojmove 'mikrobiota' i 'mikrobiom' koji se često pogrešno koriste kao sinonimi. Mikrobiota ljudskog organizma označava približno 10^{14} mikrobnih stanica koje doprinose fiziološkim procesima domaćina (ljudskog organizma), utječu na pojavnost bolesti te na samo ponašanje domaćina. U mikrobiotu ubrajamo razne viruse, bakterije, eukariote i archeae. Humani mikrobiom se odnosi na cjelokupni genom prisutnih mikroorganizama (1). Zanimljiv je podatak kako mikrobiom sadrži 100 puta više gena od humanog genoma, a broj stanica mikrobiote je 10 puta veći od broja humanih stanica. Fokus brojnih istraživanja i ovog rada većinom je na crijevnoj mikrobioti koja sadrži najveći broj mikrobnih stanica. Njihova brojnost, međusobni omjeri i razlike među različitim populacijskim grupama doprinose razumijevanju tog kompleksnog simbiotskog odnosa crijevne mikrobiote i ljudskog organizma (2).

Kako bi se shvatio taj odnos, potrebno je istražiti i definirati sastav mikrobiote zdrave osobe da bi se lakše uočile promjene povezane s bolestima. Time se bave European Metagenomics of the Human Intestinal Tract (MetaHIT) i Human Microbiome Project (HMP). Molekularne metode koje se koriste u istraživanju humanog mikrobioma su kultivacijski ovisne i kultivacijski neovisne metode (3). Kultivacijski ovisne metode, korištene u kliničkim istraživanjima, omogućile su dobar, ali nedostatan uvid u kompleksnost i raznolikost mikrobioma. Upotrebom kultivacijski neovisnih metoda u istraživanju humanog mikrobioma koje se baziraju na 16S rRNA sekvencioniranju, omogućena je identifikacija mikroorganizama koji se ne mogu uzgojiti u laboratorijskim uvjetima. Većini metoda zajednički je korak ekstrakcija DNA iz uzorka (4).

Predmet istraživanja ovog rada je usporedba različitih metoda izolacije DNA iz uzorka fecesa koji se analiziraju u sklopu projekta Biodinamik (Monitoring biodynamics of mixed cultures using the newly developed method of taking fingerprints – BioDinaMik) (5). Projekt se, između ostalog, bavi istraživanjem povezanosti mikrobiote i pretilosti te međusobnim ujtecajem. Izolacija DNA iz uzorka je nužna za daljnju obradu odnosno sekvencioniranje. Kako bi podaci bili reprezentativni, ispitivana je razlika u rezultatima ovisno o metodi (*Tool Kit*) koja je korištena. Cilj je bio utvrditi razlike u DNA prinosu (*DNA yield*) i broju različitih identificiranih vrsta između tri korištene metode. Također se tražila razlika s obzirom na

čuvanje i spremanje uzorka: zamrzavanje na -80°C te aplikacija DNA Shield™ stabilizatora (reagens koji služi za stabilizaciju uzoraka pri sobnoj temperaturi pri čemu se ne mijenja sastav uzorka).

U ovom radu sekvencionirana je V3-V4 regija 16S rRNA gena korištenjem Illumina MiSeq platforme te su korištenjem QIIME programa identificirani prisutni mikroorganizmi u uzorcima i određene razlike sastava mikrobiote u ovisnosti o korištenoj metodi izolacije DNA i metodi čuvanja uzorka.

3. TEORIJSKI DIO

2.1. HUMANA CRIJEVNA MIKROBIOTA

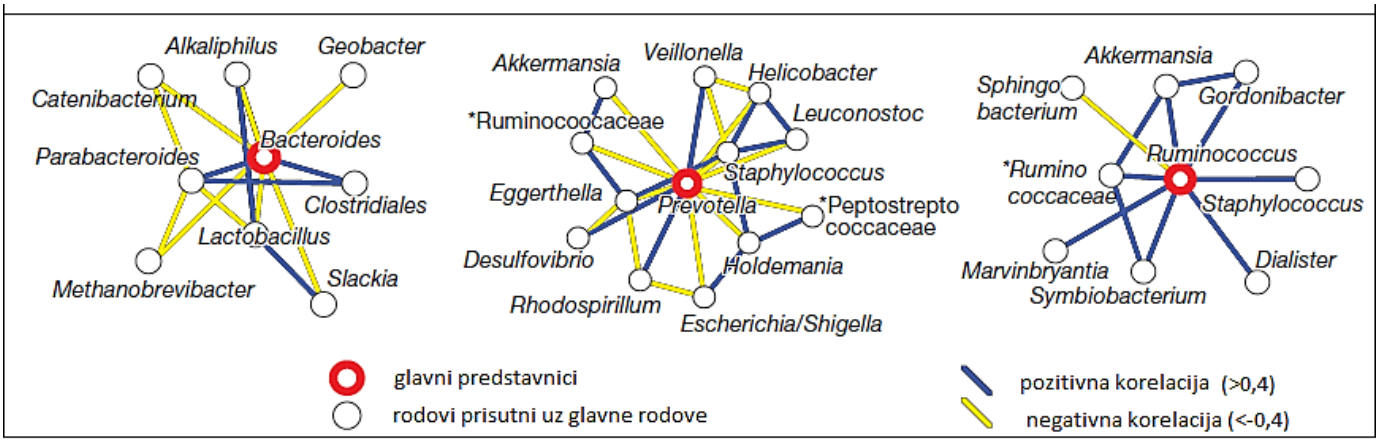
Crijevna mikrobiota predmet je sve većeg broja istraživanja iz više razloga. Od jednostavnih gastrointestinalnih smetnji koje su brzo uočljive preko povezanosti s pretilošću pa sve do, u novije vrijeme aktualnom, povezanosti s neuropsihijatrijskim poremećajima. Komunikacija između mozga i crijeva odvija se putem autonomnog i enteričnog živčanog sustava, neuroendokrinog te imunološkog sustava. Ta komunikacija se intenzivno istražuje s obzirom na nova saznanja o humanom mikrobiomu koja ukazuju na „bottom-up" princip odnosno direktan utjecaj mikrobiote na središnji živčani sustav – mozak (6). Ljudski organizam kao domaćin pruža anaerobne uvjete u zatvorenom i zaštićenom okolišu uz konstatan izvor hrane. Mikrobiota svojim rastom i iskorištenjem nutrijenata nedostupnih domaćinu pomaže u održavanju njegove homeostaze. Kompeticijom za hranu i prostor, stimulacijom sluznice imunološkog sustava te ometanjem virulencije, rezidencijalna mikrobiota regulira rast patogenih organizama (8)(9).

Iako se nameće pretpostavka da sve zdrave osobe imaju isti sastav crijevne mikrobiote, istraživanja ukazuju kako on varira među populacijama i u ovisnosti o vremenu. U svakom pojedincu može se identificirati više od 1 000 filotipova na razini vrste. Većina tih mikroorganizama su bakterije iz svega 8 koljena (phyla) od poznatih 55 (3). Kod odraslih osoba su to Bacteroidetes i Firmicutes dok Actinobacteria, Proteobacteria i Verrucomicrobia općenito manje doprinose sastavu iako su prisutna kod većine ljudi (10). Bez obzira što se glavna koljena uvijek identificiraju među populacijama, brojnost, omjeri i prisutnost vrsti uvelike variraju. U pokušaju pronalasku baznog sastava (*core set*) mikrobiote na razini vrste, identificirani su *Faecalibacterium prausnitzii, Roseburia intestinalis* i *Bacteroides uniformis*, no kod nekih osoba ne predstavljaju čak ni 0,5% od ukupne prisutne mikrobiote (11). Imajući to na umu, bazni sastav mikrobiote na razini vrste se čini nemoguće definirati.

U istraživanju iz 2011., Arumugam i sur. predlažu teoriju enterotipova prema kojoj se na temelju sastava mikrobiote ljudi mogu podijeliti u tri glavne grupe – enterotipa. Enterotipovi se međusobno razlikuju po sastavu i dominantnoj ulozi glavnih bakterijskih rodova - Bacteroides, Prevotella i Ruminococcus, a neovisni su o dobi, spolu, geografskom podrijetlu, tjelesnoj masi odnosno stupnju uhranjenosti(BMI) (Slika 1). Iako se razlikuju u sastavu na razini vrste, funkcionalna analiza mikriobioma pokazuje da velika zajednica ne znači nužno veliku molekularnu funkcionalnost. Npr. broj jedinki (količina, masa) *Escherichia* je vrlo mali, ali doprinosi u proizvodnji preko 90% proteina povezanih s pilima (proteinskim tvorbama koje

omogućuju vezanje bakterija na podlogu, organ ili tkivo domaćina). Time se pokazuje kako velika brojnost neke vrste ili roda ne može otkriti cjelokupnu funkcionalnu kompleksnost crijevne mikrobiote (12).

Sve veći broj metagenomskih istraživanja dovodi u pitanje ovu teoriju koja se često uzima kao definitivna. Podaci više pokazuju kontinuirane razlike unutar dominantnih taksonomskih kategorija nego strogo koncipirane enterotipove. Važnost u postavljanju teorija po kojoj se proučava sastav mikrobiote je njihov utjecaj na daljnji tijek istraživanja bolesti vezanih uz mikrobiotu. Kada bi se svi pacijenti mogli svrstati u jedan od 3 enterotipa, dijagnoza i liječenje bolesti bi brže napredovale u smjeru individualnog pristupa na temelju mikrobioma. S druge strane, pronalazak biomarkera za bolesti i dalje ipak zahtjeva uporabu naprednije statistike, metodologije i u konačnici veći broj uzoraka (13).



Slika 1. Enterotipovi, glavni predstavnici u odnosu s ostalim (12)

Crijevna mikrobiota razvija se zajedno s ljudskim organizmom pri čemu se povoljan sastav mikrobiote češće nalazi kod vaginalnog i terminskog poroda, dojenja i izlaganju mikroorganizmima za razliku od poroda carskim rezom, prijevremenog poroda, dohrane i terapije antibioticima. Prve bakterije koje koloniziraju crijevo nakon poroda su fakultativni anaerobi koji stvaraju uvjete za anaerobne mikroorganizme (primarno *Bifidobacterium* i *Bacteroides spp*). Vaginalnim porodom osigurava se veća zastupljenost *Bifidobacterium spp* nego carskim rezom. Također, isključivo dojena novorođenčad češće ima stabilniju bakterijsku zajednicu od novorođenčadi hranjene formulama. Uvođenjem krute hrane u prehranu pa sve do sedme godine života, dolazi do promjene sastava mikrobiote koja oko osamnaeste godine poprima punu zrelost i povoljan omjer koljena (phyla) (53).

2.2. POVEZANOST MIKROBIOTE I PRETILOSTI

Prema podacima WHO (World Health Organization) za 2014.-u godinu, 39 % muškaraca te 40 % žena imalo je prekomjernu tjelesnu masu. Iste godine, 11% muškaraca i 15% žena bilo je pretilo. Izračun koji se koristi kao pokazatelj stupnja uhranjenosti je *Body Mass Index* – BMI (Indeks Tjelesne Mase). U tablici 1 je prikazan opis stupnja uhranjenosti s odgovarajućim rasponima (7).

Tablica 1. Stupanj uhranjenosti prikazan kao raspon BMI(kg/m^2)

	BMI(Body Mass Index)	Opis
BMI= kg/m^2	<18,5 kg/m^2	pothranjenost
	18,6-24,9 kg/m^2	Normalna uhranjenost
	25,9-29,9 kg/m^2	Pretilost I.stupnja (Prekomjerna tjelesna masa)
	30,0-34,9 kg/m^2	Pretilost II.stupnja
	>35,0 kg/m^2	Pretilost III. stupnja

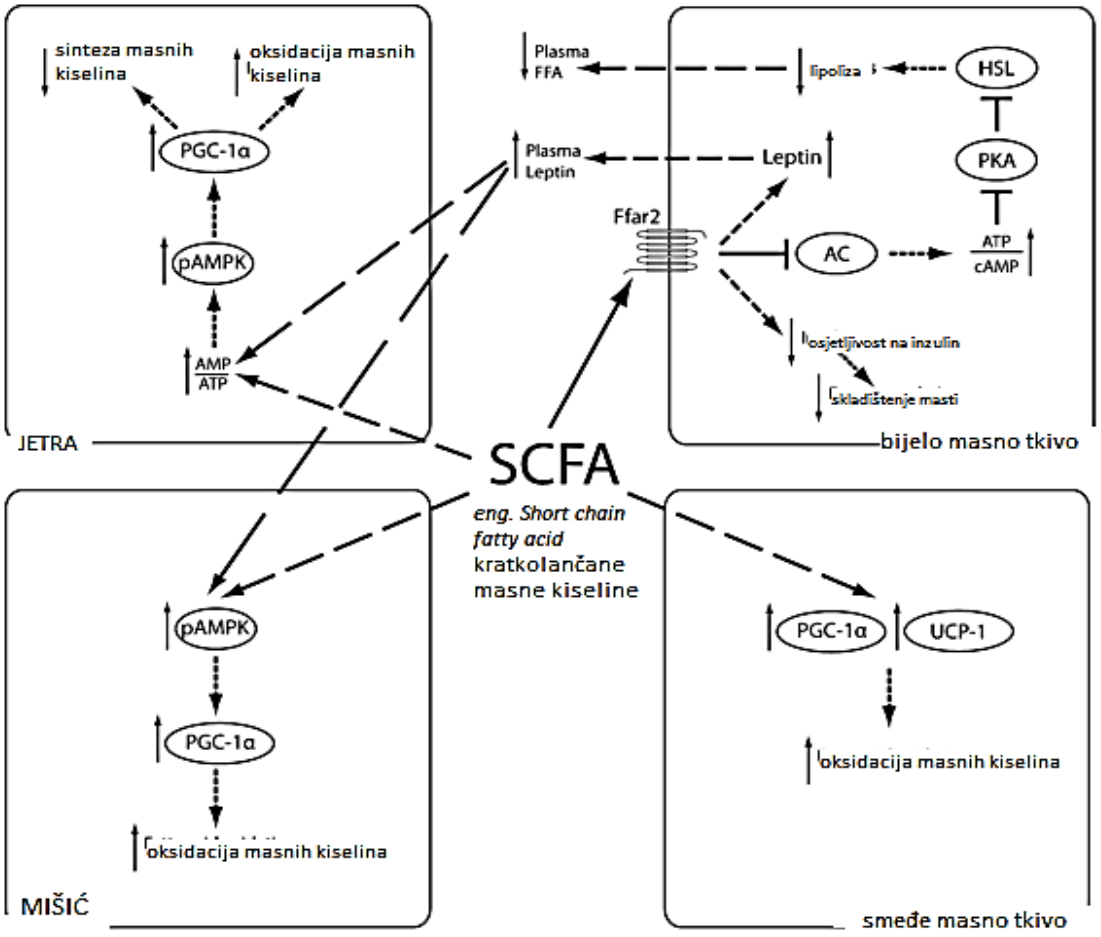
2013. godine *American Medical Association* definiralo je pretilost kao bolest pri čemu je glavni pokazatelj za dijagnozu BMI. Iako BMI kao pokazatelj stupnja uhranjenosti ne daje informacije o sastavu tijela, kod većine populacije postavljena granica od 30 kg/m^2 predstavlja početak problema koje vežemo uz pretilost. Oni mogu biti mehanički i metabolički te kombinacija (14). Sa pretilošću se sve više veže i pojam metaboličkog sindroma. Metabolički sindrom može se definirati kao skup međusobno povezanih fizioloških, biokemijskih, kliničkih i metaboličkih faktora koji međusobnim djelovanjem na organizam povećavaju rizik od obolijevanja od kardiovaskularnih bolesti i diabetesa tipa II. Dijagnosticira se kad se ispune sljedeći kriteriji: otpornost na inzulin, nakupljanje masnog tkiva u području abdomena, dislipidemija i hipertenzija. Jedna od glavnih karakteristika metaboličkog sindroma je stanje konstantne akutne upale organizma koju potiče akumulirano masno tkivo i hormoni kojeg ono luči (24).

S obzirom na prevalenciju debljine, njenog utjecaja na zajednicu i povezanosti s razvojem cijelog niza metaboličkih poremećaja, istražuju se uzroci obolijevanja i načini intervencija. Iako je glavna pretpostavka za razvoj pretilosti neravnoteža u unosu i potrošnji

energije, usporedba crijevne mikrobiote zdrave, normalno uhranjene i pretile osobe daje novi uvid na međusobni utjecaj mikrobiote i pretilosti.

Poznato je da mikrobiota pomaže u iskorištenju energije iz unešene hrane. U istraživanjima mikrobiote koriste se GF ("germ free") životinje uzgojene tako da ne posjeduju vlastitu crijevnu mikrobiotu. Uspoređujući GF i konvencionalno uzgojene miševe, GF su vidljivo mršaviji od konvencionalnih neovisno o načinu prehrane. U razdoblju od 14 dana nakon kolonizacije GF miševa mikrobiotom konvencionalno uzgojenog miša, pojavljuje se rezistentnost na inzulin i 60 % povećanje na masi iako je prehrana bila bazirana na složenim ugljikohidratima i smanjenom udjelu masti (15). Mikrobiota doprinosi iskorištenju energije hidrolizom i fermentacijom polisaharida koje domaćin ne može probaviti. Najviše bakterija nalazi se u debelom crijevu- kolonu, oko 10^{12} stanica po ml sadržaja. Fermentacijom polisaharida nastaju monosaharidi i kratkolančane masne kiseline (*eng. SCFA- short-chain fatty acid*) koji mogu biti iskorišteni kao izvor energije (16). Osim što neposredno služe kao izvor energije, kratkolančane masne kiseline posredno reguliraju unos hrane i metabolizam. Npr. fermentacijom inulina, neprobavljive oligofruktoze, nastaju veće količine kratkolančanih masnih kiselina. Njihova povećana proizvodnja utječe na sitost domaćina i smanjen unos hrane (16). Osnovne kratkolančane masne kiseline koje nastaju u ljudskom organizmu su acetat(C2), propionat(C3) i butirat(C4) u omjerima 60:25:15. Butirat je glavni izvor energije za kolonocite, acetat služi kao prekursor kolesterola u jetri dok je propionat supstrat u reakcijama glukoneogeneze (17). Istraživanja impliciraju kako propionat, uz ulogu u glukoneogenezi, ima važnu ulogu u osjećaju sitosti, inhibiciji sinteze kolesterola i masnih kiselina te regulaciji sekrecije i ekspresije adipokina. Osjećaj sitosti i smanjen osjećaj gladi rezultat su usporenog prolaska hrane kroz želudac (dulje vrijeme pražnjenja želuca). Mehaničke podražaje uvjetovane prolaskom hrane registriraju receptori aferentnih vlakana X. kranijalnog živca, vagusa, te prenose mozgu informaciju koja se interpretira kao sitost (18). Pozitivan utjecaj na regulaciju glukoze u krvi, razine inzulina te smanjenje razina LDL kolesterola zabilježeni su suplementacijom propionata u razdoblju od tjedan dana kod zdravih osoba (17). Kratkolančane masne kiseline općenito održavaju ravnotežu između sinteze i oksidacije masnih kiselina te hidrolize triglicerida (lipolize) u organizmu. Aktivira se oksidacija masnih kiselina, dok je sinteza novih i lipoliza inhibirana. Rezultat ovih reakcija je smanjena ukupna koncentracija slobodnih masnih kiselina u krvnoj plazmi što posljedično dovodi do smanjenja tjelesne mase. Iako taj mehanizam nije do kraja istražen, predloženi model uključuje direktan i indirektan utjecaj kratkolančanih masnih kiselina na jetru, mišiće, bijelo i smeđe masno tkivo. U mišićima

i jetri, kratkolančane masne kiseline fosforilaju i ativiraju pAMPK[1] direktno preko povećanja omjera AMP/ATP i indirektno preko Ffar2[2]- leptin puta u bijelom masnom tkivu. Djelovanjem na Ffar2 , smanjuje se osjetljivost na inzlin u bijelom masnom tkivu što dovodi do manje skladištenja masti. Vezanje kratkolančanih masnih kiselina na Ffr2 uzrokuje inhibiciju AC[3] koja katalizira pretvorbu ATP-a u cAMP[4]. Nakupljanje ATP-a će dovesti do inhibicije PKA[5] čija je zadaća aktivirati (fosforilacijom) HSL[6]. HSL regulira lipolizu u masnom tkivu i ispuštanje slobodnih masnih kiselina u plazmu. Njenom inhibicijom, koncentracija slobodnih masnih kiselina u plazmi pada (Slika 2)(19).

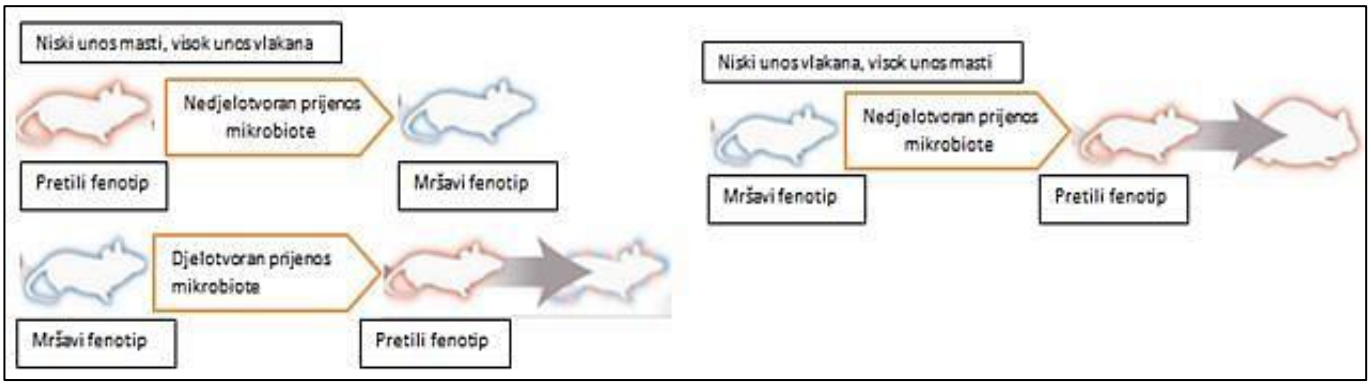


Slika 2. Utjecaj kratkolančanih masnih kiselina na regulaciju masti u organizmu (prilagođeno prema Den Besten, G. i sur.(2013))

[1] pAMPK- aktivirana protein kinaza
[2] Ffar2 – eng. Free fatty acid receptor 2- receptor slobodnih masnh kiselina na membrani adipoznog tkiva
[3] AC- adenilat ciklaza, enzim koji katalizira pretvorbu ATP u cAMP, transmembranski protein, aktivnost regulirana G proteinima
[4] cAMP – ciklički adenozin monofosfat, sekundarni stanični glasnik
[5] PKA- protein kinaza A
[6] HSL- hormon senzitivna lipaza

Na razinu glukoze u krvi utječu unos hrane, glukoneogeneza te iskorištenje organa. Uz djelovanje na glukoneogenezu putem pAMPK, kratkolančane masne kiseline mogu indirektno djelovati i na lučenje PYY (peptid YY) i GLP-1(glukagon peptid 1) hormona koji, među ostalim, djeluju na usporavanje pražnjenja želuca i smanjena pokretljivost (motilitet) crijeva. Zbog tog djelovanja ubrajaju se u supresore apetita. Koliki je točan utjecaj kratkolančanih masnih kiselina na cjelokupnu regulaciju metabolizma glukoze i lipida ipak nije poznat (19).

Jednako kao i definiranje prosječnog sastava crijevne mikrobiote kod zdravih, normalno uhranjenih odraslih osoba, prosječni sastav mikrobiote u pretilih osoba je također vrlo teško točno odrediti. Sastav mikrobiote pojedinca može se razlikovati u usporedbi s populacijom, no primjećuje se postojanost sastava samog pojedinca gledano u duljem vremenskom razdoblju (20). Na razini koljena (phyla), omjer koljena Firmicutes i Bacteroidetes pokazuje povećani broj koljena Firmicutes kod pretilih osoba. S vremenom, reguliranom prehranom, dolazi do promjene u tom omjeru u korist koljena Bacteroidetes. Ta promjena je u korelaciji s gubitkom tjelesne mase što dokazuje razlike između mikrobiote pretilih i normalno uhranjenih osoba (21). Transplatacija mikrobiote pretile osobe razvija adipoznost kod GF (″germ free″) miševa. Naknadni transfer mikrobiote normalno uhranjene osobe u pretilog miša ima reverzibilni efekt odnosno vraća se povoljan omjer koljena Bacteroidetes i Firmicutes. Ako se mišu s mikrobiotom normalno uhranjene osobe prenese mikrobiota pretile osobe, ne dolazi do promjene u fenotipu. Uvjet ovakvog ″selektivnog prijenosa″ je regulirana prehrana miša koja uključuje smanjen unos masti i povećani unos vlakana (Slika 3) (23).



Slika 3. Prijenos mikrobiote pretilog odnosno mršavog fenotipa miša ovisi o prehrani.(prilagođeno prema Walker i Parkhill, (2013))

Istraživanje provedeno na blizancima bavilo se proučavanjem sastava mikrobiote, ali i funkcionalne raznolikosti. Budući da je sastav na razini vrste gotovo nemoguće definirati, funkcionalna raznolikost je primjenjenija metoda jer uspoređuje broj različitih metaboličkih puteva i relativnu pojavnost svakog puta. Iako je pronađen mali broj metaboličkih puteva,

funkcionalna raznolikost je ipak ujednačena. Također, uočena je korelacija između veće funkcionalne raznolikosti mikrobioma i koljena Bacteroidetes, za razliku od mikrobioma gdje su dominantna koljena Firmicutes i Actinobacteria. Pretilost se može, prema tome, povezati s promjenama u sastavu mikrobiote na razini koljena, smanjenom funkcionalnom raznolikosti te izmjenjenom ekspresijom bakterijskih gena i metaboličkih puteva. Također, učinkovitije je definirati bazni sastav mikrobioma na razini funkcionalnosti budući da su odstupanja povezana s raznim fiziološkim stanjina i bolestima (22).Većina istraživanja se ipak proučava taksonomsku zastupljenost za razliku od funkcionalne zbog jednostavnijeg postupka istraživanja.

2.3 UTJECAJ METODE IZOLACIJE DNA NA ODREĐIVANJE SASTAVA INTESTINALNE MIKROBIOTE (PRETILIH OSOBA)

Ekstrakcija DNA iz uzoraka fecesa zahtjevnija je od izolacije iz ostalih uzoraka zbog kompleksnosti mikrobne flore, konzistencije sadržaja te raznih endogenih i prehrambenih faktora. S obzirom na te faktore, čistoća izolirane DNA je vrlo bitna za osjetljivost i mogućnost daljnjih analiza (33). U ovom radu istražuju se razlike između 3 metode (Tool Kit) za ekstrakciju DNA iz uzorka fecesa koje se koriste u Laboratoriju za tehnologiju antibiotika, enzima, probiotika i starter kultura, Zavoda za biokemijsko inženjerstvo Prehrambeno biotehnološkog fakulteta u Zagrebu. Za ekstrakciju su korišteni: (i) Maxwell® 16 Tissue DNA Purification Kit, (ii) Quick-DNA™ Fecal/Soil Microbe Miniprep Kit Zymo Researchi (iii) Stool DNA Isolation Kit (Cat. 27600) Norgen Biotek Corp. Također se tražila razlika s obzirom na čuvanje i spremanje uzorka: zamrzavanje na -80°C te aplikacija DNA Shield™ stabilizatora (reagens koji služi za stabilizaciju uzoraka pri sobnoj temperaturi pri čemu se ne mijenja sastav uzorka). Različite metode izolacije su testirane kako bi se odredili prinosi i bakterijska raznolikost dobivena sekvencioniranjem 16S rRNA pomoću Illumina MiSeq uređaja za sekvencioniranje.

Uzorci fecesa su uzeti od dvije osobe čiji je BMI>30 kg/m^2 . Uspoređivanje utjecaja ovih metoda na dobivenu bakterijsku bioraznolikost važan je rezultat za daljnja istraživanja u sklopu projekta BioDinaMik. Bitno je znati kolika je razlika u dobivenim podacima s obzirom na korišteni kit, budući da će se u daljnjim koracima oduzimati konkretne intervencije na prehranu ispitanika što će utjecati na sastav mikrobiote.

Istraživanja pokazuju kako svaka metoda ima svoje prednosti i mane, ovisno o tipu istraživanja za koja se koriste te da se bakterijska raznolikost kao i prinos razlikuju ovisno o

proizvođaču. I tog je razloga važno uračunato ovaj korak u planiranje istraživanja koje uključuje izolaciju i sekvencioniranje DNA mikrobiote (24).

2.4. ODREĐIVANJE SASTAVA MIKROBIOTE KORIŠTENJEM MARKER GENA

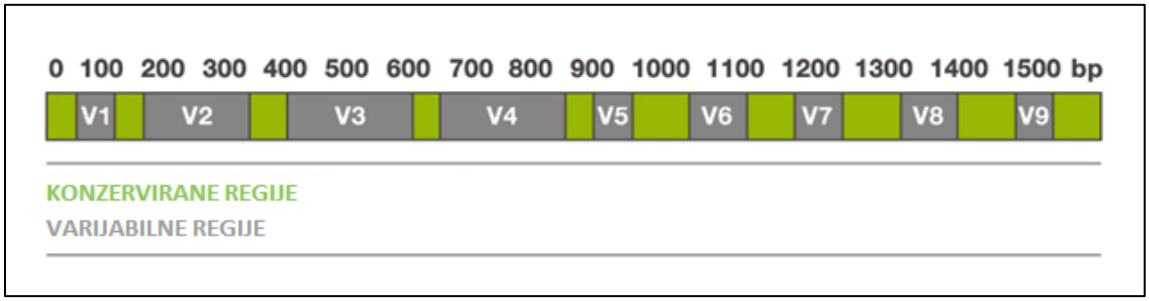
S obzirom da tradicionalne metode kultivacije i izolacije nisu prikladne za analizu pojedinačnih uzoraka mikrobiote gastrointestinalnog trakta, u posljednje vrijeme koriste se nove metode sekvencioniranja. Ove neinvazivne metode određivanja omogućuju dulje praćenje promjene sastava sadržaja uslijed nastupanja bolesti ili poremećaja te ulaska u stanje remisije.

Kako bi se identificirao mikrobiom i njegova uloga u ljudskom zdravlju i bolestima, od prevencije i pojavnosti pa do liječenja, pokrenut je Projekt humanog mikrobioma *(HMP)* te kasnije *European Metagenomics of the Human Intestinal Tract* (*MetaHIT*). Ovi projekti baziraju se na upotrebi metagenomike - ispitivanju genoma izoliranog iz mikrobne populacije uzorkovane iz prirodnog okruženja. Tako se izbjegava potreba za kultiviranjem pojedinačnih mikroorganizama i olakšava sama analiza. Obrađuje se mikrobna populacija s nekoliko mjesta na ljudskom tijelu: gastrointestinalni trakt, usna i nosna šupljina te urogenitalni trakt budući da se na tim mjestima nalazi najviše mikroorganizama. U sklopu istraživanja BioDinaMik, u ovom se radu posebno istražuje mikrobiom gastrointestinalnog trakta s obzirom na njegovu povezanost s metaboličkim sindromom i pretilošću, bolestima čija incidencija kontinuirani raste (32).

Eksperimantalne metode koje su dio nove generacije sekvencioniranja uključuju (i) „shotgun“ sekvencioniranje; metodu sekvencioniranja DNA koja omogućava uzorkovanje svih gena u svim organizmima u datom kompleksnom mikrobiološkom uzorku (ii) mikrobiološku metatranskriptomiku; analizu svih molekula RNA kodiranih grupom mikroorganizma u kompleksnom uzorku i (iii) sekvencioniranje 16S rRNA; metodu za identifikaciju i usporedbu bakterija prisutnih u uzorku. U ovom radu korištena je metoda sekvencioniranja 16S rRNA pomoću Illumina MiSeq uređaja za sekvencioniranje.

S obzirom na brzi rast broja genetičkih istraživanja, kao bitan alat za identifikaciju bakterijskih vrsta, nametnuli su se molekularni markeri. To su fragmenti DNA sekvence povezani s genomom koji se koriste za identifikaciju određene DNA sekvence (34). Ti fragmenti smješteni su i u kodirajućim i u ne kodirajućim regijama DNA te imaju nekoliko karakteristika - (i) većina su „housekeeping“ geni prisutni u svim bakterijskim vrstama, (ii) imaju visok stupanj polimorfizma, što im omogućava razlikovanje među bakterijskim vrstama

i (iii) markeri su konzervirani u pojedinim regijama što olakšava dizajniranje potrebnih početnica za PCR metodu. Više je vrsta markera koji se koriste u analizama, a neki od najkorištenijih su geni 16S rRNA, 23S rRNA, *rpoB, gyrB, dnaK* i drugi. Najpouzdanije rezulate klasifikacije bakterija dobiva se kombinacijom više biomarkera (35). Najčešće je korišteni gen 16S rRNA kao marker iako nije najpouzdaniji u klasifikaciji svih bakterijskih vrsta (vrlo bliskih vrsta sa velikom sekvencijskom sličnošću). Ipak, njegove prednosti su: prisutnost u gotovo svim bakterijama (moguće određivanje od koljena do vrste) i nepromjenjivost njegove funkcije tokom vremena. Iz tih se razloga koristi kod identifikacije rodova i vrsta bakterija koje se ne mogu identificirati prema biokemijskom profilu, sojeva koji se ne mogu identificirati uobičajenim postupcima ili za identifikaciju vrsti koje se rijetko povezuje sa zaraznim bolestima u ljudi (34). Sekvenca 16S rRNA gena duga je oko 1.5 kb i sadrži konzervirane, ali i varijabilne regije (Slika 4). Nasuprot genima koji kodiraju za enzime ima veći stupanj konzerviranosti jer rRNA ne podliježe čestim mutacijama, pa joj se ni struktura ne mijenja (36).



Slika 4. Prikaz 16S rRNA gena sa konzerviranim i varijabilnim regijama (37)

Bakterijski gen za 16S ribosomalnu RNA sadrži devet hipervarijabilnih regija koje pokazuju veliku različitost između bakterija, a nazvane su V1, V2, V3, V4, V5, V6, V7, V8 i V9, a istraživanja potvrđuju prednosti svake pojedine regije i kombinaciju određenih regija kod identifikacija vrsta. Samo jedna regija ne može se koristiti za razlikovanje svih bakterijskih vrsta jer hipervarijabilne regije pokazuju različiti stupanj sekvencijske raznolikosti (34). U svrhu ovog istraživanja obuhvaćene su regije V3 i V4 sa specifičnim početnicama.

2.5. ILLUMINA MISEQ, QIIME PROGRAM I PAIRED-END SEKVENCIONIRANJE

Kako bismo mogli ponovljivo sekvencionirati i analizirati 16S rRNA gen Illumina MiSeq uređajem (slika 5), potrebno se držati protokola. On uključuje: (i) ekstrakciju DNA, (ii) odabir početnica obzirom na varijabilne regije koje smo odlučili umnažati, (iii) pripremu knjižnice odnosno umnožavanje varijabilnih regija pomoću PCR-a, (iii) sekvencioniranje pomoću MiSeq uređaja i (iv) interpretaciju dobivenih rezultata pomoću npr. QIIME programa.

Illumina tehnologija sekvencioniranja koristi se za brza i precizna sekvencioniranja, te omogućava široku primjenu u genomici, transkriptomici i epigenomici (38)(39).

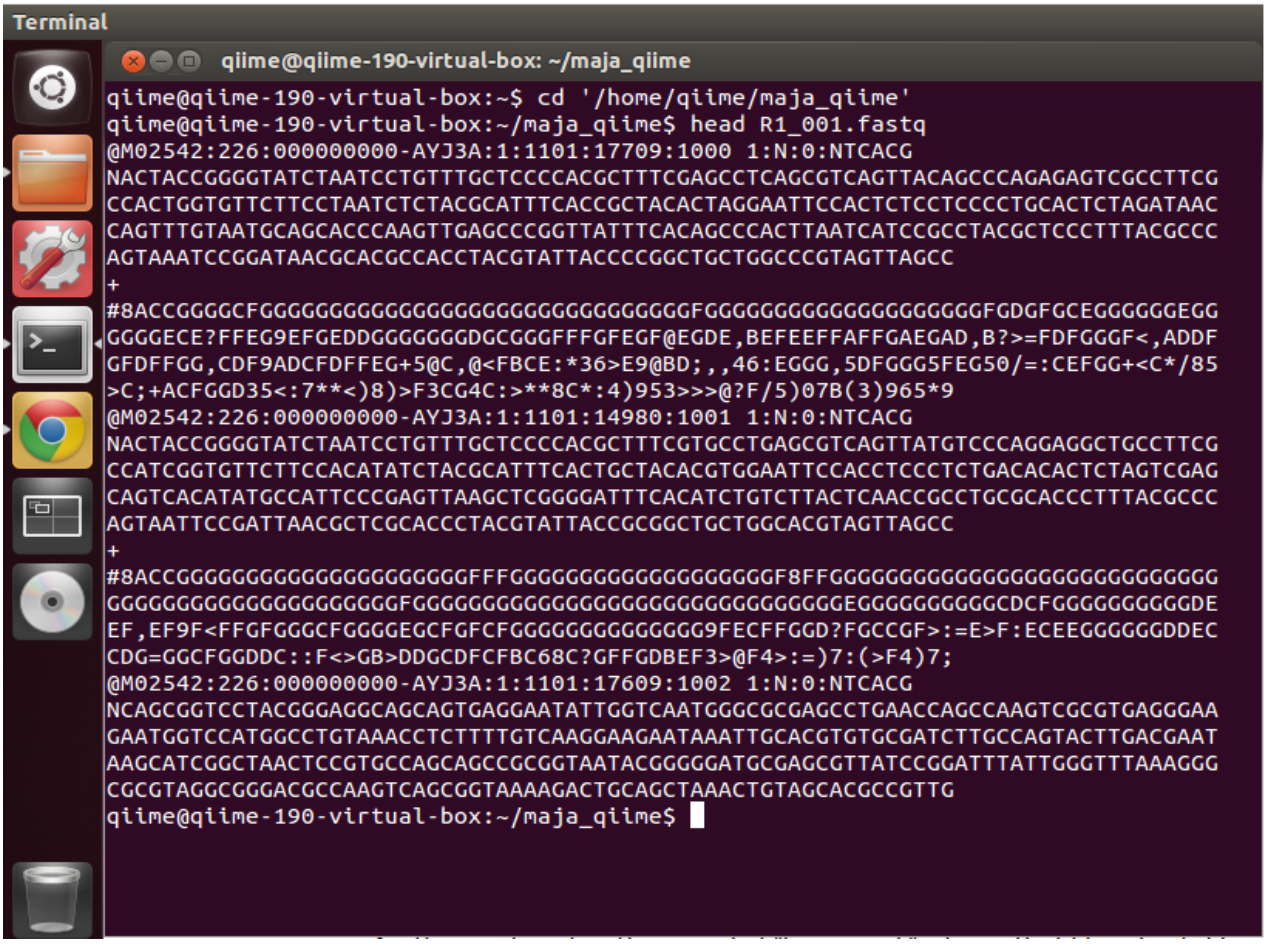
Slika 5. Illumina MiSeq uređaj(40)

Osnovni princip rada Illumina sekvencioniranja je detekcija fluoroscencije. Najprije se generiraju klonalni klasteri knjižnice molekula direktno na površini ploče. Nukleotidi koji se koriste tokom sekvencioniranja nalaze se na krajevima lanca i označeni su različitim fluoroforima. Krajevi lanaca su reverzibilni i tokom sekvencioniranja su svi nukleotidi protisnuti kroz ćeliju i pojedinačno se vežu na klastere. Nakon vezanja, višak nevezanih nukleotida se ispire, a snima se fluorescencija svakog klastera. Dodatkom enzima cijepaju se veze između nukleotida i klastera i kreće se sa novom rundom vezanja. Svaki nukleotid kodiran je emisijom različite boje te se podaci konvertiraju pomoću softvera u nukleotidnu sekvencu. Illumina MiSeq omogućuje generiranje klastera, amplifikaciju, sekvencioniranje i analizu podataka u jednom instrumentu (41).

Dobiveni „sirovi“ rezultati obrađeni su u QIIME programu (Quantitative Insight Into Microbial Ecology), javno dostupnom bioinformatičkom alatu za obavljanje mikrobioloških analiza iz sirovih podataka DNA sekvencioniranja. Dizajniran je tako da vodi korisnika od sirovih podataka dobivenih Illuminom do grafičkih prikaza i statističkih podataka potrebnih za publiciranje. Nudi brojne analize, od „demultiplexinga“ i filtriranja, preko OTU (operational taxonomic units) odabira, određivanja taksonomije, filogenetske rekonstrukcije i analize raznolikosti do vizualizacije dobivenih rezultata. QIIME alat je odabran zbog kvalitete i pouzdanosti u raznim načinima određivanja raznolikosti, statističkim metodama i vizualizacijskim alatima što potvrđuju i brojna istraživanja u kojima se koristi. Također je besplatno dostupan kao i opsežna korisnička dokumentacija za instalaciju i korištenje samog programa. Za analizu su potrebni podaci sekvencioniranja dobiveni Illuminom, Sangerom ili

nekom drugom tehnologijom sekvencioniranja, koje program analizira i daje korisne informacije o zajednicama prisutnima u svakom uzorku pojedinačno. Koristi se tako da se naredbe pišu u terminalu (terminalnom okviru), a rezultati se prikazuju grafički i tekstualno (42).

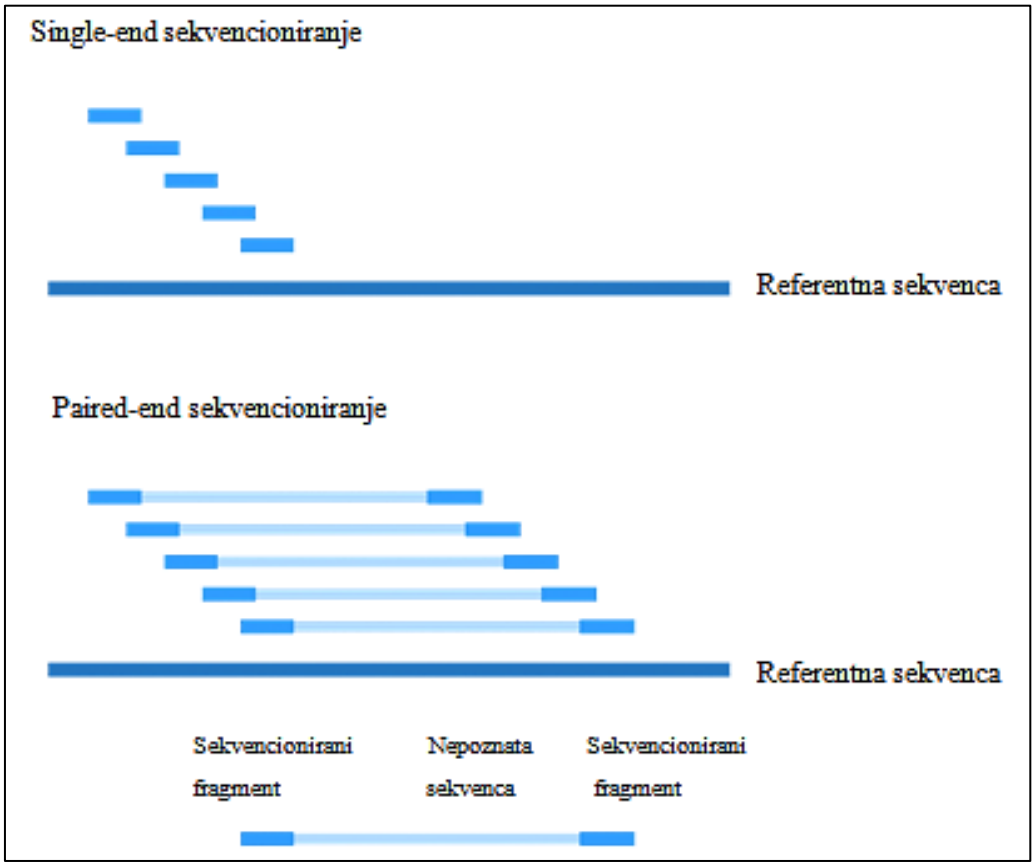
Formati podataka koji se koriste u ovom radu su: (i) FASTQ, (ii) FASTA i (iii) QUAL. (i) FASTQ (Slika 6.) je česti format za razmjenu podataka o sekvencioniranju koji kombinira podatke o očitanim sekvencama i ocjenu kvalitete svakog nukleotida u nizu. Negativna strana FASTQ formata je što se ne može otvarati u svim programima tj. nečitljiv je bez posebno instaliranog programa što može stvarati probleme. Sekvenca u (ii) FASTA formatu započinje s opisom sekvence u jednom redu nakon čega slijedi sama sekvenca. Nema podataka o točnosti sekvencioniranja što je i negativna strana ovog formata, jer se pretpostavlja da je svaki očitani nukleotid točan. Ovaj format je i dalje najčešći za pohranjivanje sekvenci jer je čitljiv u svim programima. Zadnji od navedenih formata je (iii) QUAL format i nadovezuje se na FASTA-u, tj. nadopunjuje ju. Pohranjuje ocjenu kvalitete za svaku bazu u sekvenci pohranjenu u FASTA formatu (43).



Slika 6. Primjer dokumenta u FASTQ formatu

Nova generacija sekvencioniranja, uključujući Illumina sustave, sposobna je provoditi tz. paired-end sekvenioniranje. Za razliku od single-end sekvencioniranja, koje sekvencionira samo jedan kraj DNA fragmenta u 5´-3´ smjeru i znatno je jednostavnije , pair-end sekvencioniranje sekvencionira oba kraja DNA fragmenta pri čemu se generiraju visoko kvalitetni slijedovi dužine do 600 pb (Slika7.). Metodom paird-end sekvencioniranja omogućava se dobivanje poravnanja u DNA regijama koje sadrže ponavljajuće nizove i popunjavanje praznina u konsenzus sekvencama. Ova metoda također može uočiti pregradnje u DNA, poput insercije, delecije ili inverzije (44).

Mikrobiota, odnosno njen sastav se prikazuje uz pomoć bioinformatičkih alata i programa kao QIIME koji obrađuju sirove podatke dobivene sekvencioniranjem. Kako bi se ti podaci analizirali i kategorizirali prema sličnim sekvencama, QIIME koristi OTU (*operational taxonomic unit*)- operativne taksonomske jedinice.

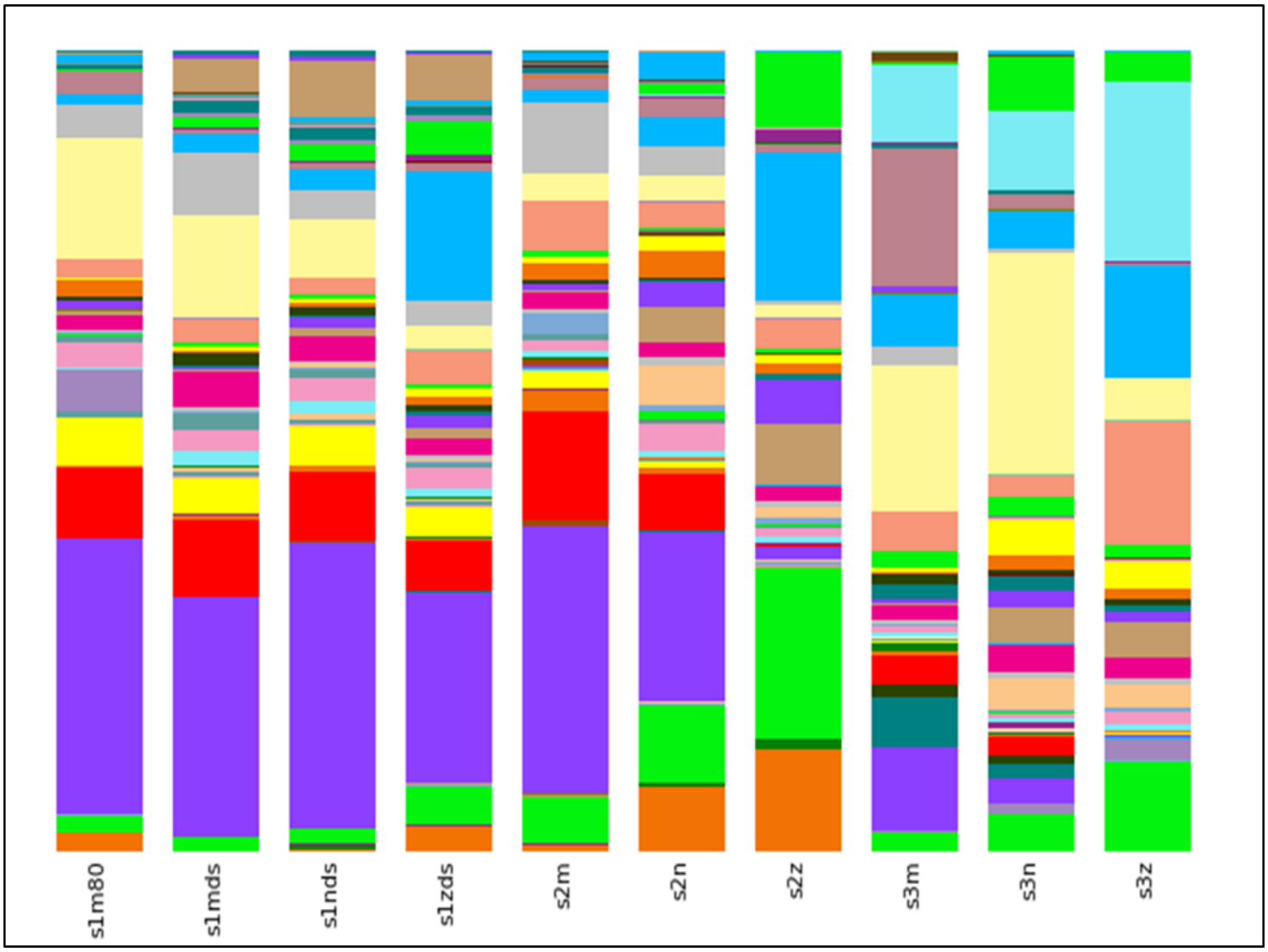


Slika 7. Razlika između single-end i paired-end sekvencioniranja (45)

Oni predstavljaju klastere sekvenci te određeni stupanj taksonomske srodnosti. Kada se dobivene sekvence grupiraju na 97% sličnosti, svaki klaster predstavlja vrstu. Kod grupacije na 98% ili 99% sličnosti, svaki klaster predstavlja vrstu. Tri su protokola (načina) za grupiranje

sekvenci u QIIME-u: (i) ″de novo″, (ii) ″closed-reference″ i (iii) ″open-reference″. (i) ″de novo″ pristup se koristi kad ne postoji vanjska referentna kolekcija sekvenci (npr. pri korištenju nestandardnog marker gena). Ovaj način grupira sve sekvence u uzorku, stoga je važno ne provoditi ga kod velikih setova podataka kao npr. HiSeq 2000 run zbog duljine samog procesa. (ii) ″closed-reference″ način grupira sekvence prema referentnoj kolekciji sekvenca. Svaka sekvenca koja nije uključena u kolekciji se isključuje iz daljne analize. Budući da su svi OTU klasteri već definirani u referentnoj kolekciji, bolja je izrada taksonomske klasifikacije i filogenetskog stabla. Također, mogu se grupirati sekvence velikih setova podataka. Mana ovog načina je nemogućnost prepoznavanja sekvenci koje nisu uključene u referentnu kolekciju. (iii) ″open-reference″ način grupiranja sekvenci koristi kombinaciju prethodna dva načina. Svaka sekvenca koja nije grupirana s obzirom na referentnu kolekciju se grupira ″de novo″. Prednost ovog načina je grupiranje svih sekvenci i brzina izvođenja procesa. Kao i kod ″closed-refernce″ načina grupiranja, potrebna je referentna kolekcija sekvenci (26).

Sastav mikrobiote prikazuje se putem analiza raznolikosti: (i) alfa raznolikost i (ii) beta raznolikost. (i) alfa raznolikost prikazuje koje su sve vrste prisutne u danom uzorku- bogatstvo vrsta ili nedostatak istih (Slika 8- prikaz raznolikosti unutar uzorka do razine roda). (ii) beta raznolikost predstvalja komparaciju mikrobne zajednice, temeljenu na njezinom sastavu, između svih uzoraka u analizi. Temeljni rezultat određivanja beta raznolikosti je kvadratni, šuplji matriks u kojem je izračunata „udaljenost“ ili različitost između svakog uzorka (27).



Slika 8. Alfa raznolikost unutar uzoraka - svaki rod označen je drugom bojom

3. EKSPERIMENTALNI DIO

3.1 PROTOKOL ZA IZOLACIJU DNA

Za ekstrakciju DNA iz uzoraka fecesa, u ovom radu su korišteni: (i) Maxwell® 16 Tissue DNA Purification Kit, (ii) Zymo Quick-DNA™ Fecal/Soil Microbe Miniprep Kit i (iii) Stool DNA Isolation Kit (Cat. 27600) Norgen Biotek Corp. Feces se klasificira kao tkivo zbog svog kompleksnog sastava. Izolacija je napravljena u Laboratoriju za tehnologiju antibiotika, enzima, probiotika i starter kultura, Zavoda za biokemijsko inženjerstvo Prehrambeno biotehnološkog fakulteta u Zagrebu. Korištenje svakog „Kit-a“ zahtjeva praćenje određenog protokola.

S obzirom da neki proizvođači u protokolima daju podatke za centrifugiranje u *RPM*(rotacija u minuti da bi se postigla određena G sila) dok drugi u *RCF*(relativna centrifugalna sila; mjerna jedinica g), važno je znati formulu koja povezuje te vrijednosti s obzirom na uređaj koji je dostupan za centrifugu. *r* predstavlja radijus rotora uređaja u cm

$$RPM = \sqrt{\frac{RCF}{(1.118 \times 10^{-5})\,(r)}}$$

Za sve korištene metode izolacije DNA slijeđene su upute proizvođače prema niže opisanom postupku.

3.1.1.Izolacija: Maxwell® 16 Tissue DNA Purification Kit

Izolacija Maxwell Kitom predstavlja jednostavno, učinkovito i automatizirano pročišćavanje genomske DNA iz uzoraka tkiva. Tkivo može biti svježe prikupljeno ili odmrznuto nakon skladištenja pri niskim temperaturama (-80°C). Pročišćena DNA je spremna za daljnje analize bez potrebe za predobradom, centrifugiranjem, precipitacijom ili rehidracijom DNA peleta. Dobivena DNA iznimne je čistoće i ne dolazi do kontaminacije među uzorcima. Koristi se za uzorke tkiva mase do 50 mg. Svježe ili odmrznuto tkivo treba staviti u prethodno dispenziranu jažicu. U nekim slučajevima ukupna RNA se može pročistiti sa genomskom DNA iz uzoraka stanica, te ju je potrebno ukloniti. Uklanjanje se vrši dodatkom RNaze, u ovom slučaju dodatkom 5µl RNaze A po mililitru pufera za eluaciju.

Protokol za pokretanje Maxwell® 16 uređaja (Slika 9) uključuje brisanje uređaja etanolom i zatvaranje vrata uređaja. Nakon toga se uređaj uključuje i pritisne se Run/Stop te se odabire program ovisno o tipu kita za izolaciju. Slijede se upute na displeju. Otvaraju se vrata kada postolje za kasetu izađe van, oni se namjeste u utore tako da etiketa gleda prema unutra, a žlijeb na sredini kasete prema nama. Odlijepi se pokrov sa kasete i uzorci se stavljaju u prvu jažicu, a klipovi u zadnju. Poslože se tubice za eluaciju i u svaku se doda 300µl pufera za eluaciju. Postupak izolacije započinje zatvaranjem vrata uređaja i pritiskom tipke Run/Stop. Sami postupak izolacije u Maxwell® 16 uređaju traje 30 – 45 minuta (46).

Slika 9. Maxwell® 16 uređaj za izolaciju DNA (46)

3.1.2. Izolacija: Quick-DNA™ Fecal/Soil Microbe Miniprep Kit

Izolacija Quick-DNA™ Fecal/Soil Microbe Miniprep Kitom je dizajnirana za jednostavnu i brzu izolaciju DNA mikrobnih i stanica domaćina iz različitih uzoraka uključujući ljude, ptice, miševe, štakore i sl. Uzorci fecesa se liziraju brzo i efikasno „bead beating“ metodom. Zatim se DNA izolira *Fast-Spin* tehnologijom nakon čega je potrebna filtracija kako bi se uklonili ostaci koji mogu ometati PCR.

1. Dodati do maks. 150 mg fekalnog sadržaja u ZR BashingBead™ epruvetu za liziranje (0.1 i 0.5 mm) te 750 µl otopine za liziranje.
2. Sadržaj se stavi u epruvetu od 2 ml gdje se sadržaj podvrgava vorteksiranju (″bead beatingu″) 5-20 minuta na maksimalnoj brzini, ovisno o proizvođaču vorteksa.
3. Sadržaj se stavi na centrifugiranje u minicentrifugu na ≥ 10 000 g na 1 minutu.

4. Odvojiti 400 µl supernatanta preko Zymo-Spin™ IV Spin Filtera u novu kolekcijsku epruvetu od 2 ml i centrifugira na 8 000 g jednu minutu.
5. Filtratu iz koraka 4. dodati 1 200 µl lizirajućeg pufera.
6. Prenijeti 800 µl sadržaja iz koraka 5. u Zymo-Spin™ IIC Column 4 i centrifugirati na 10 000 g 1 minutu.
7. Odliti tekući sadržaj te ponoviti korak 6. i prenijeti u novu kolekcijsku epruvetu.
8. Dodati 200 µ DNA Pre-Wash pufera i centrifugirati na 10 000 g 1 minutu.
9. Dodati 500 µl g-DNA Wash pufera u Zymo-Spin™ IIC Column 4 epruvetu i centrifugirati na 10 000 g 1 minutu.
10. Prenijeti sadržaj u čistu 1,5 ml mikrocentrifugalnu epruvetu i dodati 100 µl DNA eluacijskog pufera. Centrifugirati na 10 000 g 30 sekundi.
11. Staviti Zymo-Spin™ IV-HRC Spin filter u čistu epruvetu i centrifugirati 3 minute na 8 000 g.
12. Prenijeti eluiranu DNA kroz filter u priređenu 1,5 ml epruvetu i centrifugirati na 8 000 g 1 minutu. Filtrirana DNA je spremna za daljnju obradu (47).

3.1.3. Izolacija Stool DNA Isolation Kit (Cat. 27600) Norgen Biotek Corp

Izolacija Stool DNA Isolation Kitom je brza metoda za izolaciju DNA iz svježeg ili smrznutog uzorka fecesa. Omogućava istovremenu izolaciju DNA iz svih mikroorganizama u uzorku kao i stanica domaćina.

1. Dodati do 200 mg uzorka fecesa u epruvetu s kuglicama(bead tube) i 1ml lizirajućeg pufera L. Kratko vorteksirati.
2. Dodati 100 µl lizirajućeg aditiva A i kratko vorteksirati.
3. Centrifugirati na 14 000 RPM 2 minute. Prenijeti 600 µl čistog supernatanta u čistu mikrocentrifugalnu epruvetu.
4. Dodati 100 µl pufera I, protresti rukom i staviti na inkubaciju na led (10 minuta)
5. Centrifugirati 2 minute na 14 000 RPM kako bi se izdvojili ostaci stanica.
6. Prenijeti 500 µl čistog supernatanta pipetom (paziti da se ne dodiruju stijenke) u čistu epruvetu od 2 ml. Dodati jednaki volumen 70%-nog etanola. Vorteksirati.
7. Dodati 600 µl lizata s etanolom na kolonu i centrifugirati 1 minutu na 14 000 RPM. Odliti rezidualnu tekućinu i ponoviti postupak s ostatkom mješavine lizata i etanola.
8. Dodati 500 µl pufera C na kolonu. Centrifugirati na 10 000 RPM 1 minutu. Odliti rezidualnu tekućinu i složiti ponovno kolonu s kolekcijskom epruvetom.

9. Dodati 500 µl otopine za ispiranje A u kolonu i centrifugirati 1 minutu na 10 000 RPM. Ponoviti korake 8 i 9.
10. Centrifugirati kolone na 10 000 RPM 2 minute. Maknuti kolekcijsku epruvetu.
11. Staviti kolonu u čistu eluacijsku epruvetu od 1,7 ml.
12. Dodati 100 µl eluacijskog pufera B na kolonu i inkubirati na sobnoj temp. 1 minutu.
13. Centrifugirati 1 minutu na 10 000 RPM

3.2. PROTOKOL ZA SEKVENCIONIRANJE DNA

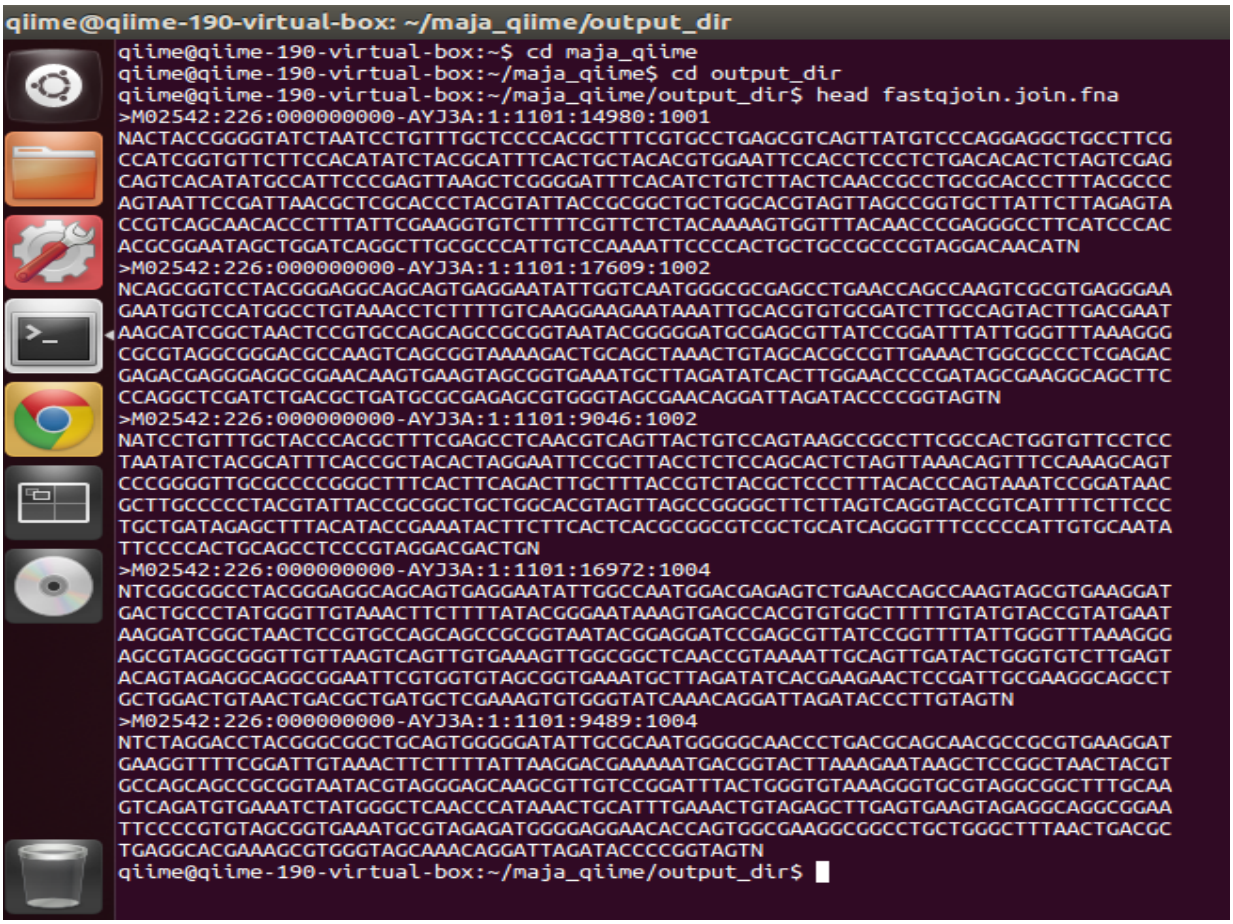
Izolirana DNA iz uzorka fecesa poslana je na sekvencioniranje u Molecular Research Lab (MR DNA). MR DNA (Shallowater, TX, USA) nudi besplatnu sveobuhvatnu analizu podataka (49). Kao povratna informacija dobiju se sirovi podaci i protokol koji je korišten pri sekvencioniranju. Podaci su podvrgnuti prosječno strogoj kontroli kvalitete koja odgovara za većinu projekata.

Za početak, V3 i V4 regija 16S rRNA gena umnožena je u 30 ciklusa PCR-a pomoću početnica specifičnih za te regije. Koristio se HotStarTaq Plus Master Mix Kit (Qiagen, SAD) pod sljedećim uvjetima: 94°C kroz 3 minute, zatim 28 ciklusa na 94°C po 30 sekundi, 53°C kroz 40 sekundi i 72°C kroz 1 minutu, te krajnji korak elongacije na 72°C, 5 minuta. Nakon amplifikacije provedena je provjera uspješnosti amplifikacije i intenzitet bandova na 2%-om agaroznom gelu. Više je uzoraka skupljeno zajedno, u jednakim omjerima, na temelju njihove molekularne mase i koncentracije DNA. Tako skupljeni uzorci pročišćeni su sa Ampure XP kuglicama. Pročišćeni PCR produkti korišteni su za pripremu DNA knjižnice prateći protokol Illumina TruSeq DNA knjižnica. Sekvencioniranje se provodilo u MR DNA (Shallowater, TX, USA) na već spomenutom MiSeq uređaju po uputama proizvođača. Kako bi amplifikacija bila maksimalna, MR DNA koristi kraće početnice, tj 50 pb dugi linker i barkod, i 20 pb je dužina same početnice. Budući da je korištena metoda pair-end sekvencioniranja, imamo 5´-3´ orijentaciju kao što je uobičajeno, ali i 3´-5´ orijentaciju u r1 i r2 datotekama. R1 i r2 datoteke predstavljaju sekvencu uzorka u 3´-5´ i u 5´-3´ smjeru.

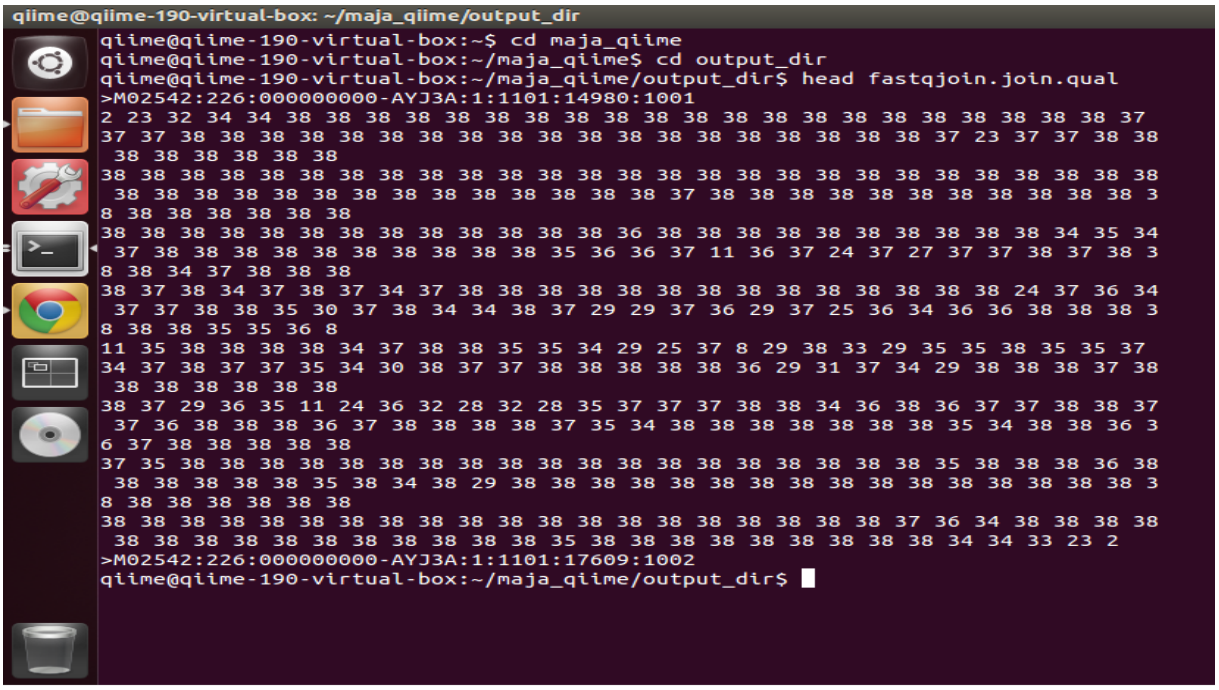
MR DNA koristi metodu mnogostrukog (eng. multiplex) sekvencioniranja koje uključuje istovremeno sekvencioniranje više uzoraka na jednoj ploči. Kako bi se identificirali uzorci na ploči koriste se specifične barkod sekvence (kratak slijed nukleotida) koje se dodaju svakom pojedinom uzorku i služe za kasnije razlikovanje uzoraka tokom analize. Istovremeno

analiziranje više uzoraka štedi se vrijeme, a i troškovi. Kasnije tijekom analize barkod sekvence se uklone kako bi dobili čistu sekvencu DNA koja se koristi u daljnjim analizama.

Sirovi podaci nakon analize dobiju se u FASTQ formatu, a MR DNA generira još nekoliko dokumenata u različitim formatima pomoću svojih analiza. FASTA podaci označeni su ekstenzijom .fna (slika 10), a QUAL podaci ekstenzijom .qual (slika 11).



Slika 10. Dokument u FASTA formatu



Slika 11. Dokument u QUAL formatu

Predstavljaju tekstualnu verziju sekvencioniranih podataka i ocjene pojedine kvalitete. Također, tu se nalazi i mapping dokument(-.txt). On sadrži svaki naziv svakog uzorka, sekvencu barkoda korištenog za svaki uzorak, sekvencu početnice korištene za umnožavanje uzorka, ime projekta u tekstualnom obliku i opis.(50). Dobiveni sirovi podaci u fastq formatu koristiti će se u daljnjoj analizi pomoću QIIME programa.

3.3. QIIME

Kao što je već ranije spomenuto, QIIME (Quantitative Insights Into Microbial Ecology) predstavlja bioinformatički alat za analizu mikrobioma iz sirovih podataka DNA sekvencioniranja (29). Za početak, potrebno je instalirati QIIME program na računalo prateći upute sa stranice www.qiime.org. S obzirom da je korišten Windows operacijski sustav, potrebno je, kao pretkorak, instalirati i program Virtual Box, https://www.virtualbox.org/wiki/Downloads. Nakon instalacije oba programa, kreće se s korištenjem samog QIIME programa koji se otvara unutar Virtual Box-a (51).

Obrada dobivenih sirovih podataka može se podijeliti na „upstream“ i „downstream“ korake. „Upstream“ koraci služe za generiranje OTU tablica, dok „downstream“ koraci služe za analizu i interpretaciju rezultata, te određivanje α i β raznolikosti i njihovu vizualizaciju.

3.3.1. Pre-procesiranje

Prvi korak pre-procesiranja je spajanje slijedova generiranih sekvencioniranjem. Spajanje slijedova je nužno jer pair-end metoda omogućava dobivanje dužih sekvenci – do 600 pb pri čemu se sekvenca očitava iz oba smjera što rezultira dvjema očitanim sljedovima koji se nalaze u dvije datoteke (r1 i r2). Ponuđene metode kojima se slijedovi mogu spojiti su: (i) fastq-join i (ii) SeqPrep. Korištena je prva metoda: fastq-join. Naredba koja se unosi u terminalni prozor QIIME progama je: `join_paired_ends.py -f forward_reads.fastq -r reverse_reads.fastq -m fastq-join -o output_dir.`

Argument `-f forward_reads.fastq` pokazuje na datoteku koja sadržava očitane slijedove DNA u prednjem smjeru u FASTQ formatu, a `-r reverse_reads.fastq` je argument za datoteku sa slijedovima u obrnutom smjeru. `-o` predstavlja direktorij za spremanje rezultata naredbe. Kao rezultat `join_paired_ends.py` naredbe dobili smo spojene slijedove u FASTQ formatu kao datoteku fastqjoin.join.fastq. Uz to, u direktoriju sa rezultatima, dobiju se dvije dodatne datoteke u FASTQ formatu: fastqjoin.un1.fastq i fastqjoin.un2.fastq. Predstavljau nesparene slijedove, a cilj je da ih bude što manje.

Kako bi mogli dalje raditi u QIIME-u, podaci se moraju iz FASTQ formata pretvoriti u FATSA i QUAL format. Cijela naredba za pretvorbu glasi: `convert_fastaqual_fastq.py -f fastqjoin.join.fastq -c fastq_to_fastaqual -o output_dir. convert_fastaqual_fastq.py` je naredba za pretvorbu, `-f` predstavlja FASTQ datoteku koju želimo pretvoriti u FASTA dokument, `--c` je dio naredbe koji govori koji format želimo pretvoriti u koji, u ovom slučaju to je FASTQ u FASTA i QUAL. `-o` označava direktorij za spremanje rezultata naredbe. Rezultat ove naredbe su dva dokumenta: fastqjoin.join.fna (FASTA format) i fastqjoin.join.qual (QUAL format).

Daljnji korak je provjera mapping dokumenta koji sadrži sve informacije o uzorku, početnicama i barkodu. Bitna je njegova kompatibilnost s QIIME-om kako bi se analiza mogla nastaviti bez problema. Naredba glasi: `validate_mapping_file.py -m F-mapping.txt -o validate_mapping_file_output.` `-m` predstavlja mapping dokument(ovdje nazvan mapping_file.txt), a `-o` označava direktorij za spremanje rezultata naredbe. Ako postoji neki problem u mapping dokumentu, on je najavljen u terminalnom prozoru, a može se vidjeti u HTML dokumentu koji je dio rezultata (Slika 12). Greška(error) označena je crvenom bojom i mora se ispraviti kako ne bi ometala daljnje korake. Upozorenje (″warning″) označeno je žutom bojom, može se i zanemariti, ali preporuka je ispraviti ih sve.

Krajnji rezultat ove naredbe je tekstualni dokument koji sadrži sve informacije potrebne za daljnji rad i kompatibilan je sa QIIME programom (mapping_file_corrected.txt).

No errors or warnings detected.

Mapping file error and warning details.

Notes for interpreting this report:

- Errors will be listed in red, warnings in yellow.
- Mouse over an error or warning in a cell for more details.
- Errors in the header row may mask other errors, so these should be corrected first.
- Modifications to your mapping file to fix certain issues may result in different errors. You should run `validate_mapping_file.py` until no errors (nor warnings, ideally) are found.

Some general rules about formatting mapping files (see here for additional details):

- Header characters should only contain alphanumeric and _ characters only.
- Valid characters for SampleID fields are alphanumeric and . only.
- Other fields allow alphanumeric and +-%. / : , ; _ characters.

General issues with your mapping file (i.e., those that do not pertain to a particular cell) will be listed here, if any:

SampleID	BarcodeSequence	LinkerPrimerSequence	BarcodeName	ProjectName	ExtractionKit	Origin	storage	Description
s2m	CATTGCGT	CCTACGGGNGGCWGCAG	illwhoi341bar14	120616JZ341F	maxwell	LS	80	F3
s1m80	CCACCTTA	CCTACGGGNGGCWGCAG	illwhoi341bar15	120616JZ341F	maxwell	LS	80	F4
s3m	CCACGGTG	CCTACGGGNGGCWGCAG	illwhoi341bar16	120616JZ341F	maxwell	ID	80	F5
s1zds	ATCGATGG	CCTACGGGNGGCWGCAG	illwhoi341bar25	120616JZ341F	zymo	LS	shield	F14
s1mds	ATCGGCGG	CCTACGGGNGGCWGCAG	illwhoi341bar26	120616JZ341F	maxwell	LS	shield	F15
s1nds	ATCTAGCG	CCTACGGGNGGCWGCAG	illwhoi341bar27	120616JZ341F	norgen	LS	shield	F16
s2z	ATCTAGGA	CCTACGGGNGGCWGCAG	illwhoi341bar28	120616JZ341F	zymo	LS	80	F17
s2n	ATCTCGAT	CCTACGGGNGGCWGCAG	illwhoi341bar29	120616JZ341F	norgen	LS	80	F18
s3z	ATGAACAG	CCTACGGGNGGCWGCAG	illwhoi341bar30	120616JZ341F	zymo	ID	80	F19
s3n	ATGAACTT	CCTACGGGNGGCWGCAG	illwhoi341bar31	120616JZ341F	norgen	ID	80	F20

Slika 12. Prikaz html dokumenta mapping file-a

Zadnji korak pre-procesiranja je demultipleksiranje i filtritanje kvalitete sekvenci. Napredak tehnologije sekvencioniranja (NGS-nove generacije sekvencioniranja) omogućio je multipleksiranje – sekvencioniranje više uzoraka odjednom. Naredba split_libraries.py se koristi za razdvajanje spojenih uzoraka (multiplexed). Pri tome sefiltriraju sekvence koje ne zadovoljavaju određenu kvalitetu i svakog niza u setu se preimenuje da ima jedinstveno ime s naznačenim nazivom uzorka. Time se formiraju podaci za „downstream“ analizu. Sekvence uzoraka koji se ne nalaze u mapping dokumentu ili sekvence bez odgovarajuće početnice se isključuju iz rezultata. Naredba unešena u terminalni prozor glasi `split_libraries.py -m F-mapping_corrected.txt -f fastqjoin.join.fna -q fastqjoin.join.qual -b 8 - o split_libraries_output.` `-m` predstavlja mapping dokument iz prethodnog koraka koji je kompatibilan sa QIIME-om. `-f` i `-q` su dokumenti koje smo dobili pretvorbom FASTQ dokumenta (koji je sadržavao spojene slijedove). `-b` predstavlja duljinu barkoda (standardna duljina je 12, u našem slučaju je 8) koja se vidi u mapping dokumentu, a `-o` je direktorij za pohranu rezultata. Minimalna duljina sekvence je 200 nukleotida, a maksimalna 1000. Minimalna prosječna dopuštena kvalitativna vrijednost definirana Phred bodovanjem je 25, a početnice i barkodovi se uklanjaju iz sekvenci.

Maksimalni broj nesigurnih baza je 6, isto kao i maksimalna dužina homopolimera. Rezultat ove naredbe su tri dokumenta: (i) seqs.fna dokument u FASTA formatu - sadrži sve sekvence koje zadovoljavaju postavljene parametre, te svaka sekvenca sada sadrži pripadajuću oznaku uzorka iz mapping dokumenta, (ii) histogram.txt - sadrži broj sekvenci određene duljine (iii) split_library_log.txt - sadrži sažetak split_libraries.py analize – informacije o sekvencama koje su prošle kontrolu kvalitete i kako su raspoređene diljem različitih uzoraka.

3.3.2. Analize raznolikosti

Analize raznolikosti započinju određivanjem taksonomske jedinice (*OTU- operational taxonomic unit),* operativne definicije koja se koristi za razvrstavanje grupa usko povezanih sekvenci. QIIME nudi tri strategije odnosno protokola - (i) ″de novo″, (ii) ″closed-reference″ i (iii) ″open-reference″. U ovom radu je korištena (i) de novo strategija gdje su slijedovi grupirani međusobno, bez ikakve vanjske referentne kolekcije sekvenci. Prednost ovog pristupa je grupiranje svih slijedova, a mana je vrijeme trajanja same analize- može potrajati dulje ako ima više slijedova. Odabir OTU-a sastoji se od više podkoraka. Započinje grupiranjem svih sekvenci svakog uzorka u OTU-e na temelju njihove odabrane sekvencijske sličnosti. U QIIME-u OTU-i predstavljaju klastere sekvenci te određeni stupanj taksonomske srodnosti. Recimo, kada su sekvence grupirane na temelju 97% sličnosti, svaki klaster predstavlja vrstu. Postoji više metoda grupiranja kojim se QIIME koristi. Ovdje je korištena uclust, koja stvara „sjemenke“ sekvenci koje onda stvaraju klastere bazirane na osnovi postotka identičnosti (52). Naredba u terminalnom prozoru: `pick_otus.py –i seqs.fna –m uclust –s 0.97 –o picked_otus`. `seqs.fna` poziva dokument `seqs.fna` generiran u prethodnom koraku koji predstavlja sve sekvence koje su zadovoljile postavljene paramete kvalitete, `uclust` je odabrana metoda grupiranja, a 0.97 je postavljeni prag sličnosti sekvenci. Kao rezultat se dobiju dva dokumenta: (i) seqs_otus.txt, datoteka koja sadrži tablicu sa brojem klastera i sekvencama pridruženima tom klasteru i (ii) seqs_otus.log dokument koji sadrži popisane sve postavljene parametre.

Sljedeći korak je odabir reprezentativne sekvence iz svakog OTU-a za daljnje analize (taksonomska identifikacija OTU-a i filogenetsko poravnavanje) pomoću naredbe `pick_rep_set.py`. Puna naredba koja se upisuje glasi: `pick_rep_set.py –i seqs_otus.txt –f seqs.fna –o rep_set.fna`. `–i` predstavlja OTU mapping

dokument, a `-f` FASTA dokment sa svim zadovoljavajućim sekvencama. Rezultat je FASTA dokument seqs_rep_set.fna koji sadrži jednu reprezentativnu sekvencu po OTU-u.

Zatim slijedi poravnavanje reprezentativnih sekvenci sa taksonomskom bazom podataka. QIIME nudi tri metode: (i) PyNAST, (ii) MUSCLE i (iii) INFERNAL. Ovdje se koristi (i) PyNAST metoda u kojoj NAST algoritam poravnava svaku sekvencu s najboljom sekvencom iz već poravnate baze podataka referentnih sekvenci. Naredba koja se upisuje: `align_seqs.py -i rep_set.fna -m pynast -o aligned.` Sadrži dokument koji je generiran u prethodnom koraku te odabranu metodu za poravnavanje (`-m pynast`) i izlazni direktorij. Određeni parametri su automatski postavljeni kao što su (i) minimalna duljina sekvence koja se poravnava – 75% od mediana dužine ulaznih sekvenci i (ii) minimalni postotak identičnosti koji sekvenca mora imati prema najboljem BLAST pogotku da bi bila uključena u poravnavanje – 0,75. Rezultati naredbe QIIME generira u tri dokumenta: (i) rep_set_aligned.fasta, FASTA dokument sa svim poravnanim sekvencama, (ii) rep_set_failures.fasta, FASTA dokument sa svim sekvencama koje nisu zadovoljile specificirane kriterije i (iii) rep_set_log.txt, log dokument koji sadrži informacije koje se odnose na rezultate dobivene odabranom metodom.

Ovako poravnate sekvence koriste se u daljnjim koracima, Naredba `assign_taxonomy.py` dodjeljuje taksonomiju svakoj reprezentativnoj sekvenci. Metode koje QIIME nudi su: (i) BLAST, (ii) RDP klasifikator, (iii) RTAX, (iv) mothur, (v) uclust i (vi) sortmerna. Rezultat koji se dobije su dokumenti formata log (rep_set_aligned_tax_assignments.log) i text (rep_set_aligned_tax_assignments.txt). Naredba koja se upisuje glasi: `assign_taxonomy.py -i rep_set_aligned.fasta -m rdp -o taxonomy.` `-m` pokazuje koja se metoda koristi; u ovom slučaju RDP klasifikator. Ostali parametri postavljeni su automatski i nisu mjenjani kao minimalna pouzdanosti za dodjeljivanje taksonomije koja iznosi 0,5. Konsenzus dodijeljena taksonomija implementirana u programu pruža najdetaljniji opis porijekla dijeljen među najmanje 90% sekvenci unutar pojedinog OTU-a.

`filter_alignment.py` je naredba koja se koristi prije generiranja filogenetskog stabla. Služi uklanjanju praznina u poravnanju. Cijela naredba glasi: `-filter_alignment.py -i rep_set_aligned.fasta -o filtered_alignment.` `-i` predstavlja početni FASTA dokument generiran u

prethodnim koracima. Rezultat izvođenja naredbe je dokument u FASTA formatu (rep_set_aligned_pfiltered.fasta) .

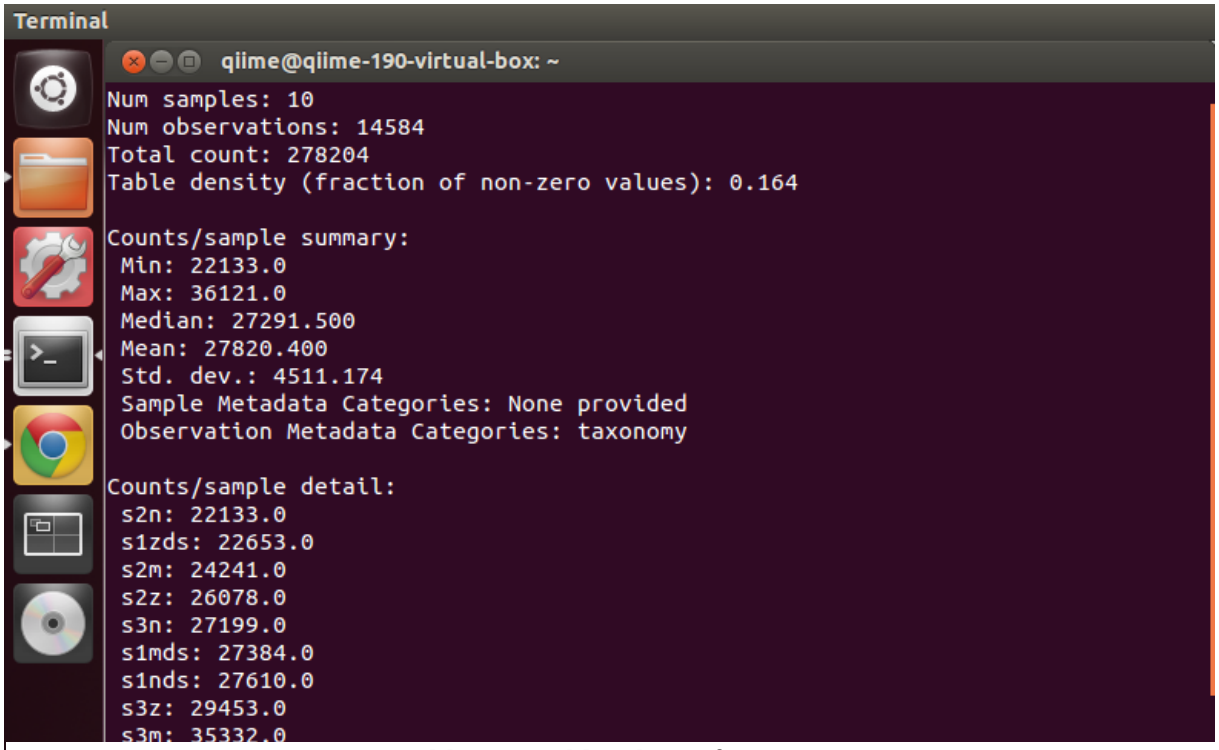
Mnoge „downstream“ analize zahtijevaju prisutnost filogenetskog stabla. Naredba `make_phylogeny.py` stvara filogenetsko drvo koristeći FastTree metodu i filtrirano poravnanje iz prethodnog koraka. `make_phylogeny.py –i rep_set_ aligned_pfiltered.fasta –t fasttree –o rep_phylo.tre` stvara dokument koji se može vizualizirati pomoću brojnih programa za vizualizaciju kao što su TopiaryTool ili FigTree.

Naredbom `make_otu_table.py –i seqs_otus.txt –t rep_set_aligned_tax_assignments.txt –o otu_table.txt` QIIME sastavlja tablicu OTU zastupljenosti u svakom uzorku s taksonomskom identifikacijom za svaki OTU. Krajnji rezultat je tablica u biom formatu (slika 13), u kojoj stupci odgovaraju uzorcima, a redovi OTU-ima i broju puta u kojem se uzorak pojavljuje u određenom OTU-u. Kako bi se vizualizirala taksonomska raznolikost, može se izabrati između nekoliko načina vizualizacije (npr. taxonomy heatmap) ovisno o vrsti rada u kojem se rezultati prikazuju.

3.3.3. Alfa raznolikost

Alfa raznolikost predstavlja raznolikost unutar uzorka te opisuje bogatstvo vrsta ili nedostatak istih u jednom uzorku. Određivanje alfa raznolikosti uključuje četiri koraka. Prvi uključuje stvaranje prorijeđenih OTU tablica koje predstavljaju bazu za određivanje alfa raznolikosti. Time se ne dobivaju krivulje raznolikosti po broju sekvenci u uzorku, nego se kreira serija poduzorkovanih OTU tablica slučajnim uzorkovanjem iz glavne OTU tablice. Svi uzorci koji sadrže manje sekvenci od traženog broja su izostavljeni iz analize. Cijela naredba prvog koraka glasi `multiple_rarefactions.py –i otu_table.txt –m 2000 –x 2200 –s 2000 –o rarefaction_tables`.

Ulazne informacije uključuju OTU tablicu i određene parametre. `-x` kao maksimalni broj sekvenci za prorijeđivanje se određuje iz tablice biom formata koja se dobije naredbom `biom summarize-table -i otu_table.biom` (slika 13). Ovdje je ta vrijednost iznosila 22000 jer je to bio najmanji broj sekvenci u svim uzorcima. `-m` predstavlja minimalni broj sekvenci za prorijeđivanje, a u ovom slučaju iznosio je 2000. `-s` je veličina svakog koraka između minimuma i maksimuma sekvenci. Ovdje je izabrana 2000 kao s vrijednost. Naredba generira brojne biom dokumente čiji broj ovisi o minimalnom i maksimalnom broju sekvenci po uzorku. Dokumenti su istog formata kao i ulazna OTU tablica.



Slika 13. Tablica biom formata

Sljedeći korak određivanja alfa raznolikosti uključuje mjerenje raznolikosti za svaku OTU tablicu dobivenu u prethodnom koraku. QIIME nudi više od 20 različitih metrika za mjerenje raznolikosti, no najčešće se koriste (i) Chao1, (ii) Observed OTUs i (iii) Phylogenetic Diversity. Naredba glasi `alpha_diveristy.py -i rarefaction_tables -m chao1, PD_whole_tree, observed_otus -t rep_phylo.tre -o adiv_chao1_PD_obsotus`. Naredba uključuje OTU tablice iz prethodnog koraka, odabrane metrike i filogenetsko stablo ukoliko se koriste filogenetske metrike kao što je Phylogenetic Diversity. Dokumenti koji se dobiju su tekstualnog oblika i uključuju tablice sa stupcima koji predstavljaju odabrane metrike i redovima koji odgovaraju uzorcima i izračunatim raznolikostima.

Tablice iz prethodnih koraka su temelj za sljedeći korak, a to je razvrstavanje rezultata alfa raznolikosti. Cilj je povezivanje brojnih dokumenata u jedan kako bi se mogle generirati krivulje raznolikosti. Naredba glasi: `collate_alpha.py -i adiv_chao1_PD_obsotus -o alpha_collated`. Kao krajnji rezultat ove naredbe dobije jedan dokument za svaku metriku koji smo koristili (chao1.txt, PD_whole_tree.txt i observed_otus.txt) u kojem se nalazi tablica. Dokumenti sadrže mjeru alfa raznolikosti za svaki uzorak i uređen je tako da su uzorci poredani od najmanjeg do najvećeg broja sekvenci.
Zadnji korak određivanja alfa raznolikosti je generiranje dijagrama za vizualizaciju pomoću mapping dokumenta i tablica iz prethodnog koraka. Svaka krivulja predstavlja jedan uzorak. Naredba glasi: `make_rarefaction_plots.py i alpha_collated -m Fassting_Map_corrected.txt -o rarefaction_plots`. Pomoću dijagrama za vizualizaciju lakše se uspoređuje sama raznolikost. Kao rezultat QIIME daje html dokument pomoću kojeg se vizualiziraju sve krivulje.

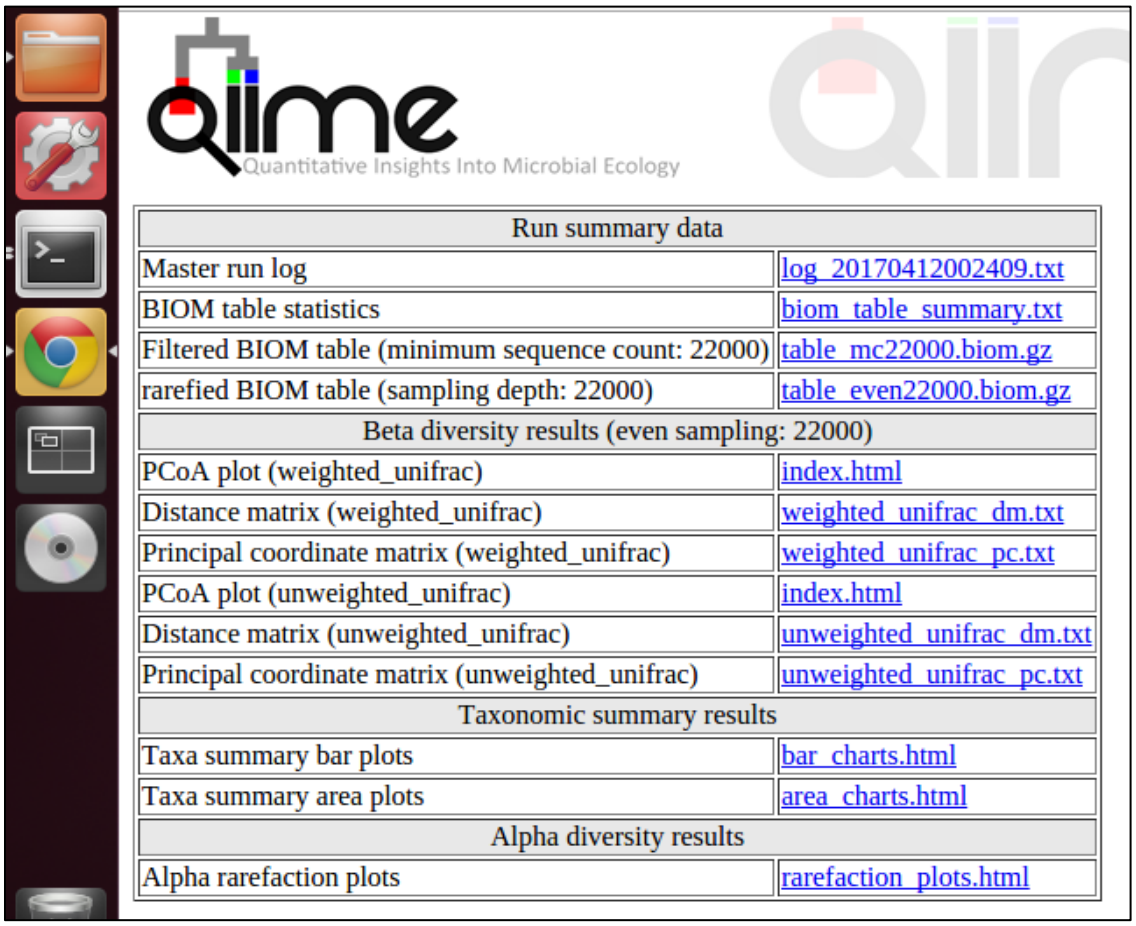
3.3.4. Beta raznolikost

Beta raznolikost predstavlja raznolikost između više uzoraka u analizi. Time se dobiva slika sličnosti mikrobnih zajednica, u ovom slučaju crijevne mikrobiote. Rezultat ovih usporedbi je kvadratni matriks gdje je izračunata različitost ili „udaljenost“ između svakog uzorka. Dobiveni podaci se mogu vizualizirati sa Principal Coordinates Analysis (PCoA) analizom. QIIME nudi više opcija za izračunavanje beta raznolikosti, tzv. metrika, a ovdje se koristi UniFrac (54). Kako bi se izvršilo određivanje raznolikosti, koristi se samo jedna naredba koji obavlja istovremeno više koraka: (i) single_rarefaction.py, korak koji modificira OTU tablice da bi se uklonila heterogenost u uzorcima (25). Ovaj korak služi za nasumičnu analizu istog broja sekvenci iz svakog uzorka, neovisno o tome koliko je sekvenci u pojedinom uzorku, da bi se postigla analiza zajednica na jednakom nivou, (ii) beta_diversity.py, koristi se za procjenjivanje beta raznolikosti pomoću weighted unifrac metode za kvantitativno mjerenje i unweighted unifrac metode za kvalitativno mjerenje raznolikosti, (iii) principal_coordinates.py, korak za usporedbu grupa uzoraka na temelju filogenetskog stabla ili metrika udaljenosti baziranih na broju OTU-ova i (iv) make_emperor.py, zadnji korak koji generira emperor dijagrame. Emperor je interaktivni alat za analiziranje, vizualizaciju i razumijevanje mikrobnih podataka pri čemu omogućava 3D vizualizaciju PCoA dijagama (55). Naredba glasi: `beta_diversity_through_plots.py -i otu_table.txt -m mapping_file_corrected.txt -t rep_phylo.tre -o bdiv`. Za njenu provedbu, potrebno je imati OTU tablicu, filogenetsko stablo i mapping dokument. Rezultat je

trodimenzionalni prikaz svakog uzorka kao točke u prostoru pri čemu je njihova udaljenost obrnuto proporcionalna njihovoj sličnosti.

3.3.5. Skripta ″core diversity″

Završni korak analize podataka u QIIME programu je izrada skripte core diversity kojoj je svrha spojiti nekoliko prethodnih QIIME analiza kako bi se generirao osnovni tijek rada. On sadrži BIOM tablicu, mapping dokument i filogenetsko stablo (Slika 14). Naredba koja se unosi glasi: `core_diversity_analyses.py -i otu_table.biom -m mapping_file_corrected.txt -e 22000 -t rep_phylo.tre -o core_output.` Skripta je rađena nakraju analize kako bi se prošlo svi koraci potrebni za razumijevanje rezultata.



Run summary data	
Master run log	log_20170412002409.txt
BIOM table statistics	biom_table_summary.txt
Filtered BIOM table (minimum sequence count: 22000)	table_mc22000.biom.gz
rarefied BIOM table (sampling depth: 22000)	table_even22000.biom.gz
Beta diversity results (even sampling: 22000)	
PCoA plot (weighted_unifrac)	index.html
Distance matrix (weighted_unifrac)	weighted_unifrac_dm.txt
Principal coordinate matrix (weighted_unifrac)	weighted_unifrac_pc.txt
PCoA plot (unweighted_unifrac)	index.html
Distance matrix (unweighted_unifrac)	unweighted_unifrac_dm.txt
Principal coordinate matrix (unweighted_unifrac)	unweighted_unifrac_pc.txt
Taxonomic summary results	
Taxa summary bar plots	bar_charts.html
Taxa summary area plots	area_charts.html
Alpha diversity results	
Alpha rarefaction plots	rarefaction_plots.html

Slika 14. Prikaz skripte core diversity u html formatu

3.3.6. Calypso web aplikacija

Za statističku obradu dobivenih podataka te različite vizualne pristupe i multivarijabilne analize korištena je online web aplikacija Calypso, specijalizirana za kompleksnu analizu obrađenih podataka sekvencioniranja. Kao ulazne podatke Calypso traži biom tablicu taksonomske identifikacije intestinalne mikrobiote te mapping dokument koji sadrži sve informacije o uzorcima. Mapping dokument je potrebno dodatno urediti prema detaljnim uputama koje Calypso nudi na svojoj web stranici (53).

4. REZULTATI I RASPRAVA

Tema ovog rada bila je usporedba različitih metoda („kit“-ova) za izolaciju DNA iz uzorka fecesa koji se analiziraju u sklopu projekta BioDinaMik (Monitoring biodynamics of mixed cultures using the newly developed method of taking fingerprints – BioDinaMik). Cilj je bio utvrditi razlike u DNA prinosu (″DNA yield″) i broju različitih identificiranih vrsta između tri korištene metode: (i) Maxwell® 16 Tissue DNA Purification Kit, (ii) Quick-DNA™ Fecal/Soil Microbe Miniprep Kit Zymo Research i (iii) Stool DNA Isolation Kit (Cat. 27600) Norgen Biotek Corp. Također se tražila razlika s obzirom na čuvanje i spremanje uzorka: zamrzavanje na -80°C te aplikacija DNA ShieldTM stabilizatora (reagens koji služi za stabilizaciju uzoraka pri sobnoj temperaturi pri čemu se ne mijenja sastav uzorka). S obzirom na mogućnost korištenja različitih dostupnih metoda za izolaciju DNA, istraživala se razlika u podacima koji se dobiju obradom podataka sekvencioniranja. Uzorci izolirani prema prije spomenutim protokolima, poslani su na sekvencioniranje te su dobiveni sirovi podaci u FASTQ formatu. Podaci su obrađeni u QIIME programu s obzirom da se navedni program koristi i u daljnjim istraživanjima u sklopu projekta. Sličnosti odnosno razlike u podacima dobivene korištenjem različitih metoda izolacije DNA od velike su važnosti zbog međusobne usporedivosti uzoraka. Analiza je vršena na uzorcima fecesa dobivenih od pretilih osoba, međutim rezultati usporedbe metoda izolacije DNA se mogu preslikati na cjelokupnu populaciju. Analizom podataka u QIIME programu, određeni su OTU-i (operational taxonomic unit) te alfa i beta raznolikost. Alfa raznolikost nam daje uvid u raznolikost bakterijskih vrsta unutar svakog uzorka, a beta raznolikost daje uvid u raznolikost između svih deset analiziranih uzoraka. Zbog mogućnosti vizualizacije rezultata u QIIME-u, alfa raznolikost je prikazana krivuljama, dok je beta raznolikost prikazana u trodimenzionalnom prikazu. Ovi podaci daju uvid u važnost odabira i korištenja određene metode za izolaciju DNA kao prvog koraka u sekvencioniranju jer nam daju podatak da li sve metode ravnomjerno izoliraju NDA svih bakterijskuh vrsta. Kako bi se dobiveni rezultati bolje razumijeli, u tablici 2 su navedene šifre uzoraka i njihov opis te koncentracija DNA (ng/µl) u svakom uzorku. Koncentracija je mjerena BioSpec Nano spektrofotometrom za mjerenje koncentracije nukleinskih kiselina. Volumen uzorka iznosio je 2µl, a kao slijepa proba služio je pufer za eluaciju iz Maxwell kita za izolaciju. Duljina puta iznosila je 0,7 mm (28). Koncentracija DNA u uzorku pokazuje razlike u prinosu između različitih metoda.

Tablica 2. Šifre uzorka, opis i koncentracija DNA u svakom uzorku

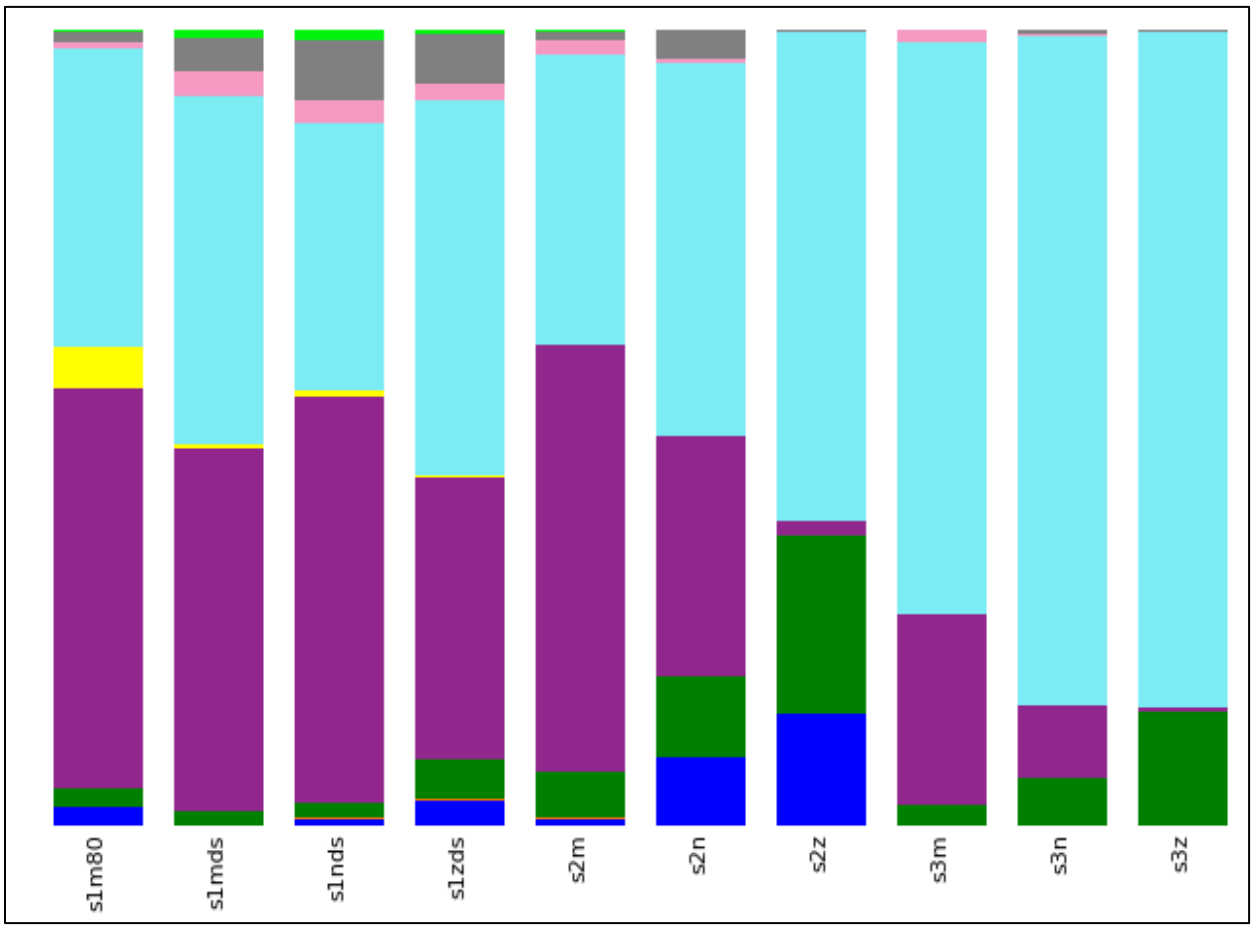
Uzorak-šifra	Opis	Koncentracija DNA(ng/µl)
s1m80	1.uzorak, Maxwell kit(smrzavanje na -80°C)	211,30
s1mds	1.uzorak, Maxwell kit(dodatak DNA Shield)	166,79
s1nds	1.uzorak, Norgen kit(dodatak DNA Shield)	31,5
s1zds	1.uzorak, Zymo kit(dodatak DNA Shield)	42,29
s2m	2.uzorak, Maxwell kit	116,73
s2n	2.uzorak, Norgen kit	92
s2z	2.uzorak, Zymo kit	70,74
s3m	3.uzorak, Maxwell kit	82,72
s3n	3.uzorak, Norgen kit	86,2
s3z	3.uzorak, Zymo kit	10,27
*Napomena: Uzorci 1 i 2 su od fecesa iste osobe, a uzorak 3 je od druge osobe, Samo uzorci označeni ds (s1xds) su odmah stavljeni u DNA Shield, ostali su, po standardnoj proceduri skladišteni na -80°C		

Iz tablice 2 se vidi kako se najveći prinos DNA ostvaruje korištenjem Maxwell metode u uzorku s1m80. Maxwell metoda također daje najveći prinos u uzorcima tretiranim DNA Shield stabilizatorom i u drugom uzorku, dok je u trećem uzorku najbolji prinos postignut korištenjem Norgen metode. Najmanji prinos u uzorcima 2 i 3 daje Zymo metoda, dok je u uzorku 1 najmanji prinos dobiven korištenjem Norgen metode. Manja koncetracija DNA u uzorku, odnosno prinos, rezultat je nemogućnosti metode da filtrira i pročisti kompleksni uzorak kao što je uzorak fecesa.

4.1. ODREĐIVANJE TAKSONOMIJE

Uz pomoć QIIME programa, određene su taksonomske jedinice odnosno prisutnost bakterija u svakom uzorku. Analiza je napravljena do razine roda, a može se vidjeti na slici 16. Oznake na slikama (grafovima) redom označavaju: slovom „p“ označeno je koljeno (phylum), slovom „c“ razred (classis); slovo „o“ označava red (ordo), slovo „f“ predstavlja porodicu (familia), a slovo „g“ rod (genus). Iako nas primarno zanimaju bakterije iz uzoraka, vidimo i prisutnost carstva Arcahea

Slika 15 prikazuje taksonomiju svih deset uzoraka do razine koljena.



Legend	Taxonomy	Total %	s1m80 %	s1mds %	s1nds %	s1zds %	s2m %	s2n %	s2z %	s3m %	s3n %	s3z %
	None;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	k__Archaea;p__Euryarchaeota	3.0%	2.3%	0.0%	0.9%	3.2%	0.8%	8.4%	13.9%	0.0%	0.0%	0.0%
	k__Bacteria;Other	0.1%	0.0%	0.1%	0.1%	0.1%	0.2%	0.1%	0.1%	0.1%	0.0%	0.0%
	k__Bacteria;p__Actinobacteria	7.2%	2.2%	1.7%	1.8%	5.2%	5.9%	10.2%	22.4%	2.4%	5.9%	14.3%
	k__Bacteria;p__Bacteroidetes	30.2%	50.3%	45.7%	51.2%	35.2%	53.6%	30.1%	1.8%	24.1%	9.2%	0.4%
	k__Bacteria;p__Cyanobacteria	0.7%	5.3%	0.5%	0.6%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	k__Bacteria;p__Firmicutes	54.8%	37.5%	43.6%	33.5%	47.2%	36.5%	47.0%	61.4%	71.9%	84.0%	84.9%
	k__Bacteria;p__Proteobacteria	1.3%	0.8%	3.1%	3.0%	2.1%	1.8%	0.4%	0.1%	1.3%	0.3%	0.1%
	k__Bacteria;p__Synergistetes	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	k__Bacteria;p__TM7	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	k__Bacteria;p__Tenericutes	2.5%	1.3%	4.4%	7.6%	6.2%	1.0%	3.6%	0.1%	0.1%	0.5%	0.2%
	k__Bacteria;p__Verrucomicrobia	0.3%	0.2%	0.9%	1.2%	0.5%	0.2%	0.1%	0.0%	0.0%	0.0%	0.0%

Slika 15. Grafički prikaz taksonomskih jedinica u uzorcima fecesa taksonomska razina koljena

Iz grafičkog prikaza može se vidjeti razlika u rezultatima ovisno o metodi izolacije DNA i broju uzorka. Najveći udio kroz sve uzorke imaju koljena Firmicutes (54,8%) i Bacteroidetes

(30,2%). Također, u svim uzorcima se mogu detekirati koljena Proteobacteria i Actinobacteria koja su prisutna kod većine ljudi.

U prvom uzorku, s1m80, vidi se veći broj udio koljena Bacteroidetes (50,3%) u odnosu na Firmicutes (37,5%) i Proteobacteria (0,8%) što nije u skladu s pretpostavkom da je kod pretilih osoba koljeno Firmicutes više zastupljeno u odnosu na Bacteroidetes. Također u prvom uzorku je detektirana značajno veća prisutnost Cyanobacteria od 5,3% za razliku od sljedeća tri uzorka gdje je pristunost redom s1mds-0,5%, s1nds-0,6% i s1zds-0,3%. S obzirom na te rezultate, možemo nagađati da da je vrijeme proteklo do smrzavanja uzorka na -80°C (1 h) utjecalo na umnažanje bakterija koljena Cyanobacteria u usporedbi s uzorcima gdje je DNA Shield^{TM} stabilizator dodan odmah po izuzimanju. Usporedba s1m80 i s1mds, dakle uzoraka koji su izolirani istom metodom- Maxwell kit, pokazuje razlike s obzirom na čuvanje uzoraka nakon prikupljanja. Dolazi do smanjenja koljenja Bacteroidetes i Cyanobacteria, dok se udio Firmicutes povećava s 37,5% na 43,6%. S obziorm da razlika nije značajno možemo spekulirati da do nje dolazi zbog faktora slučajnostui u cijelom postupku određivanja bioraznolikosti.

Gledamo li usporedbu između uzoraka izoliranih istom metodom, također se vide razlike u rezultatima. s1m80 i s2m uzorci izolirani su Maxwell metodom od uzoraka iste osobe čuvanih na -80, ali pokazuju veliku razliku u udjelu glavnih koljena. Koljeno Bacteroidetes – u s2m 53,6% dok je u s3m udio 24,1%. Udio Firmicutes u s2m je 36,5% dok je u s3m 71,9%. Treće najbrojnije koljeno je Actinobacteria čiji je postotak u s2m 5,9% dok je u s3m 2,4%.

Između s2n i s3n, uzoraka izoliranih Norgen Biotek Corp metodom također se vidi razlika u glavnim koljenima; porast koljena Firmicutes s 47% na 84%, pad koljena Bacteroidetes s 30,1% na 9,2% te Actinobacteria s 22,4% na 5,9%.

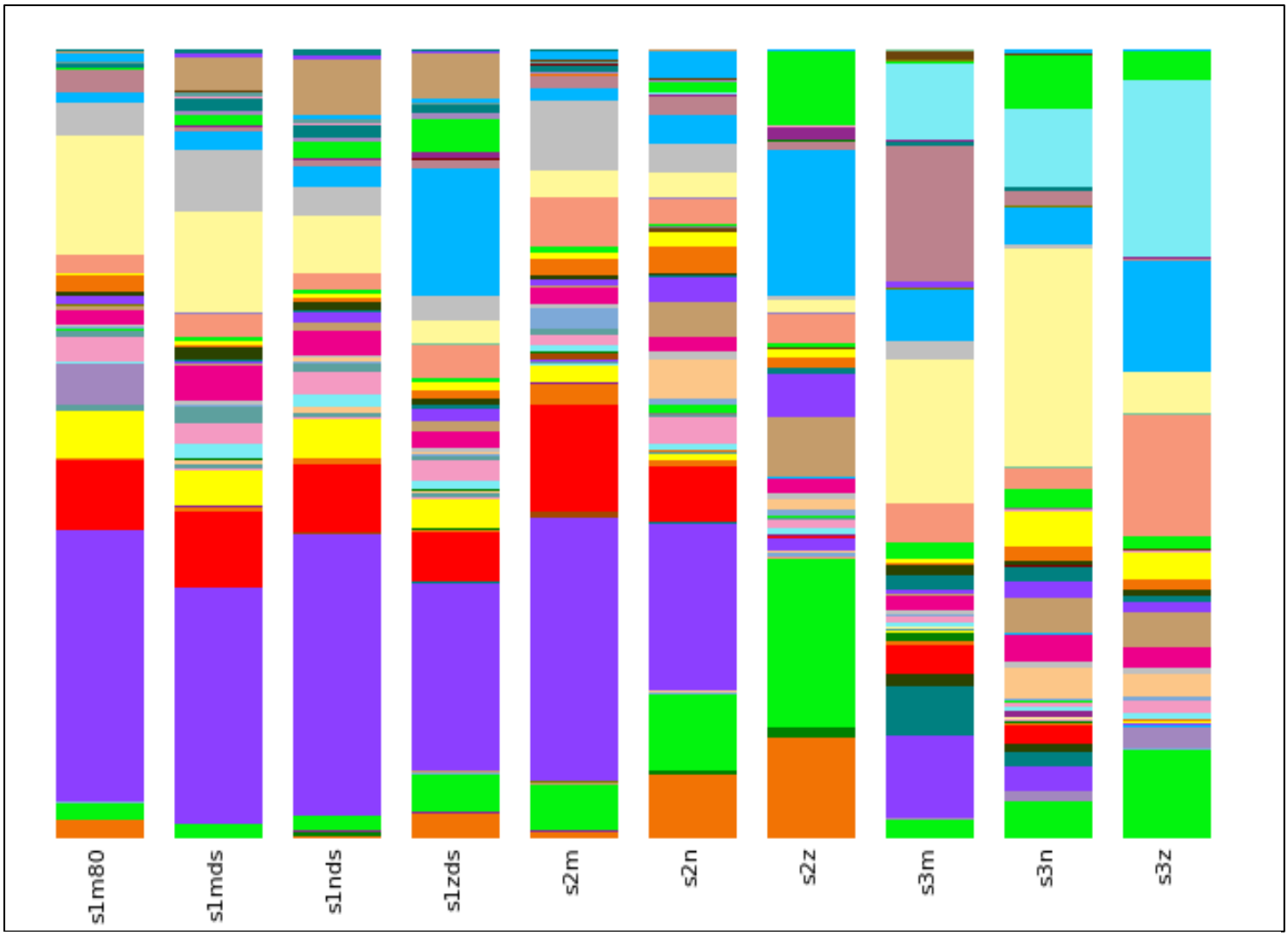
s2z i s3z izolirani su zadnjom metodom, Zymo Research Quick-DNA™ Fecal/Soil Microbe Miniprep Kit. Razlika u glavnim koljenima je ponovno prisutna, u s2z udio koljena Firmicutes je 61,4% dok u s3z iznosi 84,9%. Postotak koljena Bacteroidetes s 1,8% pada na 0,4% u s3z. U s2z je zabilježen najveći udio Actinobacteria- 22,4%, ali je i taj postotak manji u s3z- 14,3%.

Ukupno gledajući u uzorku 1, izuzev s1m80, u udio koljena Bacteroidetes vidi se razlika koja se kreće između 35%(s1zds) preko 45,7%(s1nds) do 51,2%(s1nds). Koljeno Firmicutes detektirano je s manjim razlikama, 33,5% u s1nds preko 43,6% u s1mds do 47,2% u s1zds.

Detekcija koljena Proteobacteria je najsličnija među svim metodama: s1mds- 3,1%; s1nds- 3,0% i s1zds- 2,1%.

Uzorak 2 pokazuje veliku razliku u detekciji koljena Bacteroidetes: 53,6% u uzorku s2m, 30,1% u s2n i samo 1,8% u s2z. Ako se pogleda ranija usporedba s2z i s3z, vidljivo je da Zymo metoda ne omogućava detekciju Bacteroidetes koljena. U s1m80 uzorku, najmanji udio je isto upravo u uzorku iz Zymo metode.

Padom udjela Bacteroidetes, vidljiv je porast koljena Eurychaeota iz carsta Archaea: od s2m- 0,8% preko s2n- 8,4% do s2z- 13,9%. Zanimljivo, ovo koljeno nije uopće detektirano u uzorku broj 3 što može značiti pogrešku u provođenju izolacije zadnjeg uzorka ili sekvencioniranju gena. Koljeno Firmicutes detektirano je u rasponu udjela od 25 %- s2m- 36,5%, s2n- 47,0% i s2z- 61,4%.. Između sva tri uzorka, Zymo metoda daje najviši udio Firmicutes koljena. Koljeno Actinobacteria je također više detektirano u ovom uzorku za razliku od prethodnog. Najviše ima u s2z- 22,4%, zatim u s2n- 10,2% i na kraju s2m- 5,9%. Također, slično je vidljivo u uzorku broj 3 gdje je također metodom Zymo zabilježeno najviše Actinobacteria u odnosu na ostale 2 metode. U uzorku 3 je najveći postotak Firmicutes koljena među svim uzorcima, a Zymo metoda daje naviši udio kao što je već i opisano. s3n i s3z imaju gotovo isti postotak 84,09% i 84,9% dok je u s3m 71,9% bacteria Fimicutes. Koljeno Bacteroidetes detektirano je u s3m- 24,1% te s3n- 9,2% dok se Zymo metodom(s3z) dobije svega 0,4%.



Slika 15. Grafički prikaz taksonomskih jedinica do razine roda

Tablica 3 prikazuje popis 10 najzastupljenijih rodova kroz svih deset uzoraka.

Bakterijski rod	% zastupljenosti u svih 10 uzoraka
Bacteroides	19,3%
Faecalibacterium	9%
Ruminococcus	7,1%
Bifidobacterium	6,5%
*Rikenellaceae g_*	6,0%
*Ruminococceae g_*	4,7%

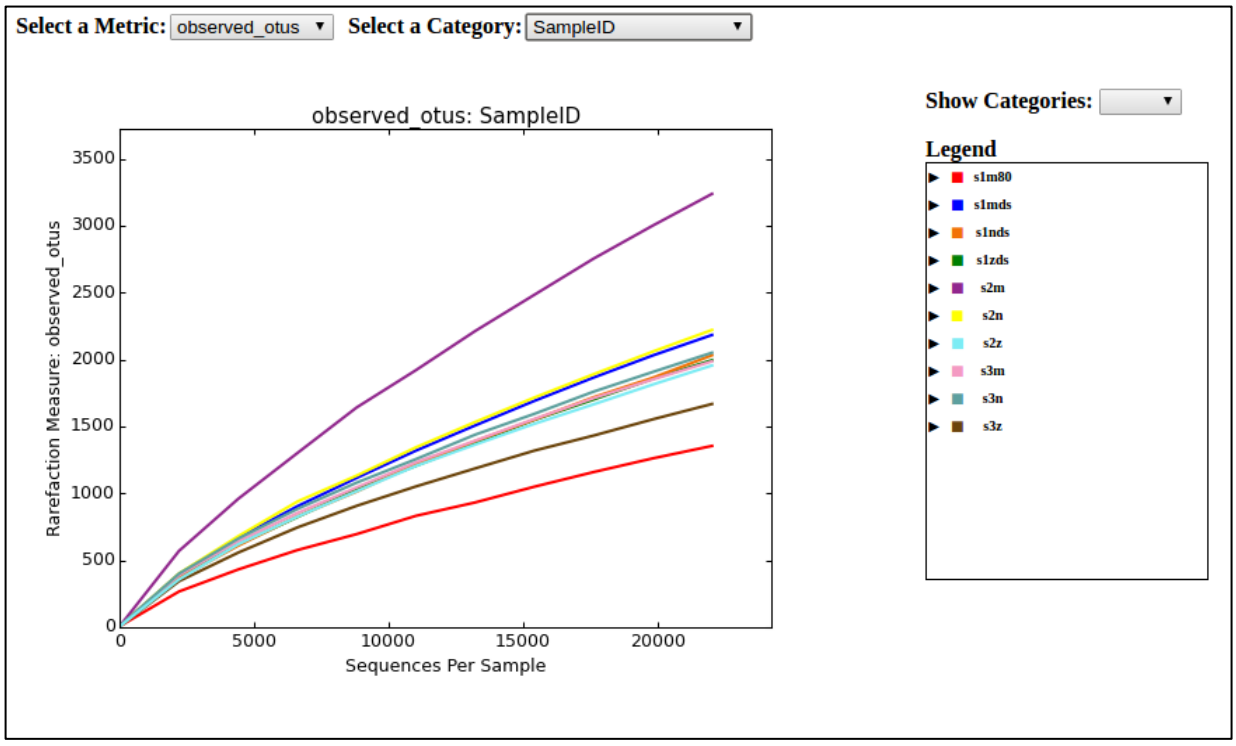
Catenibacterium	4,2%
Oscillospira	3,5%
Dialister	2,9%
Eubacterium	2,9%
Methanobrevibacter	2,7%

Iako je u uzorcima detektirana velika raznolikost rodova, u Tablici 3 je popis 10 najzastupljnijih. Najzastupljeniji rod je Bacteroides- 19,3%, a slijede ga Faecalibacterium- 9% te Ruminococcus-7,1%. Zadnji najzastupljeniji rod iz tablice je Methanobrevibacter s 2,7% na razini svih uzoraka.

4.2. ALFA RAZNOLIKOST

Za određivanje alfa raznolikosti u QIIME-u, postoji više ponuđenih metrika. Najčešće se koriste: (i) ″PD whole tree″, (ii) ″chao1″ i (iii) ″observed otus″. Serijom naredbi u QIIME-u generiraju se grafove krivulja razrjeđivanja koje služe za vizualizaciju samih rezultata. Te krivulje služe za utvrđivanje bioraznolikosti bakterijskih zajednica, odnosno da bi se utvrdilo da li je dubina sekvencioniranja bila dovoljna za pojedini uzorak. Krivulje se stvaraju tako da se svaka zajednica nasumično poduzorkuje bez zamjenjivanja u različitim intervalima, te se za svaki interval odredi prosječan broju OTU-a i prikaže u odnosu na veličinu poduzorka. Točka u kojoj se broj OTU-a ne povećava s daljnjim uzorkovanjem je točka u kojoj je uzeto dovoljno uzoraka za preciznu karakterizaciju zajednice.

Na slici 17 prikazana je alfa raznolikost koristeći *observed otus* metriku. Na x-osi prikazan je broj sekvenci po uzorku, a na y-osi broj opaženih (određenih) OTU-ova.

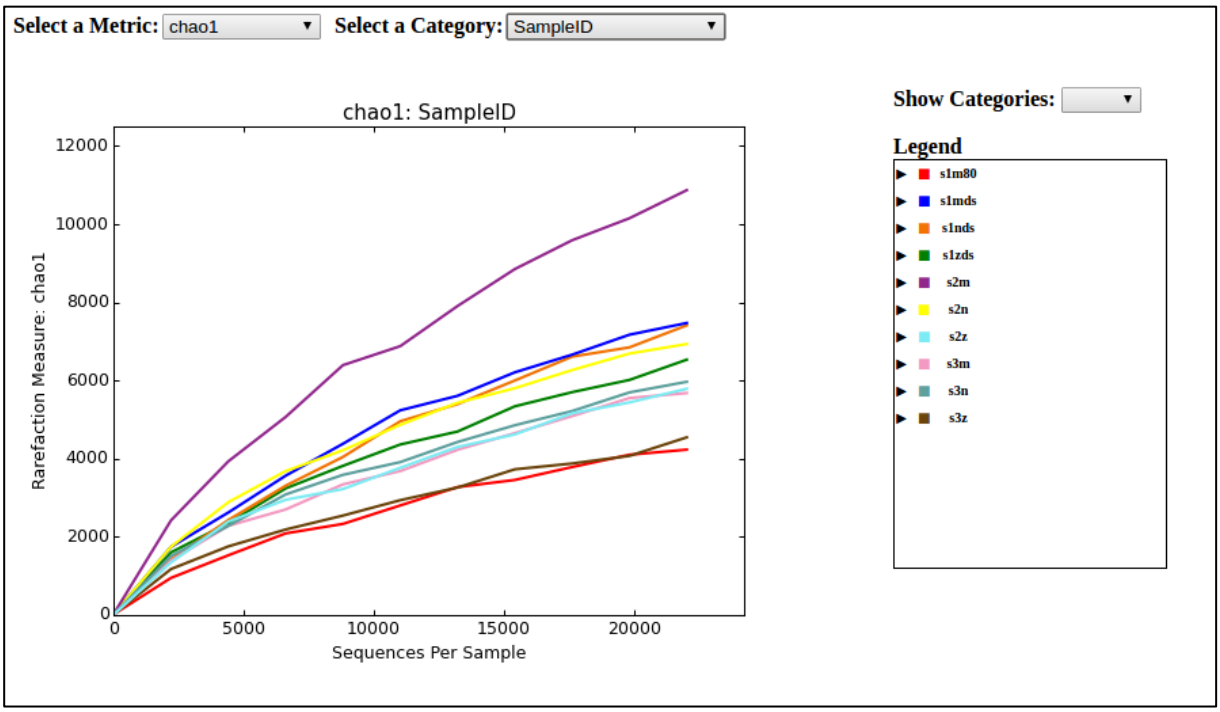


Slika 16. Krivulje razrjeđenja u *observed otus* metrici

Iz grafa možemo vidjeti da je najveća α-raznolikost prisutna u uzorku s2m koji predstavlja uzorak izolirane DNA Maxwell metodom bez dodataka DNA Shield stabilizatora. Nakon toga slijedi s2n uzorak te s1mds. Zatim slijede s3n te s2z, s3m, s1nds i s1nzds koji imaju prilično jednaku raznolikost dok je najmanja raznolikost primjećena u s1m80 i s3z uzorcima. Po opaženim rezultatima alfa raznolikosti uzorak s2m značajno odskače od ostalih uzoraka. Kako bi se sa sigurnošću moglo zaključiti da li je taj uzorak znatno bogatiji bakterijskim vrstama ili

je došlo do pogreške u samoj pripremi uzorka ili sekvencioniranju trebalo bi ponoviti postupak za taj uzorak. Vidimo da krivulje prestaju na 22 000 sekvenci po uzorku, iz razloga što je to bio maksimalni broj sekvenci u jednom od deset uzoraka, a kako bi svi uzorci bili jednako zastupljeni tada sve uzorke poduzorkujemo na taj broj sekvenci. Krivulje nisu postignule izgled hiperbole, odnosno nisu se ustalile što može indicirati na to da rezultati broja vrsta nisu maksimalni, odnosno da bi veći broj sekvenci po uzorku dao i pouzdanije podatke, odnosno veći broj vrsta u svakom uzorku.

Graf metrike *chao1* može se vidjeti na slici 18. Kao i na prethodnom grafu, uzorak s2m ima najveću raznolikost, a slijede ga s1mds, s2n, s3n te s2z, s3m, s1nds i s1nzds te s2s i s1m80 s najmanjom raznolikošću.

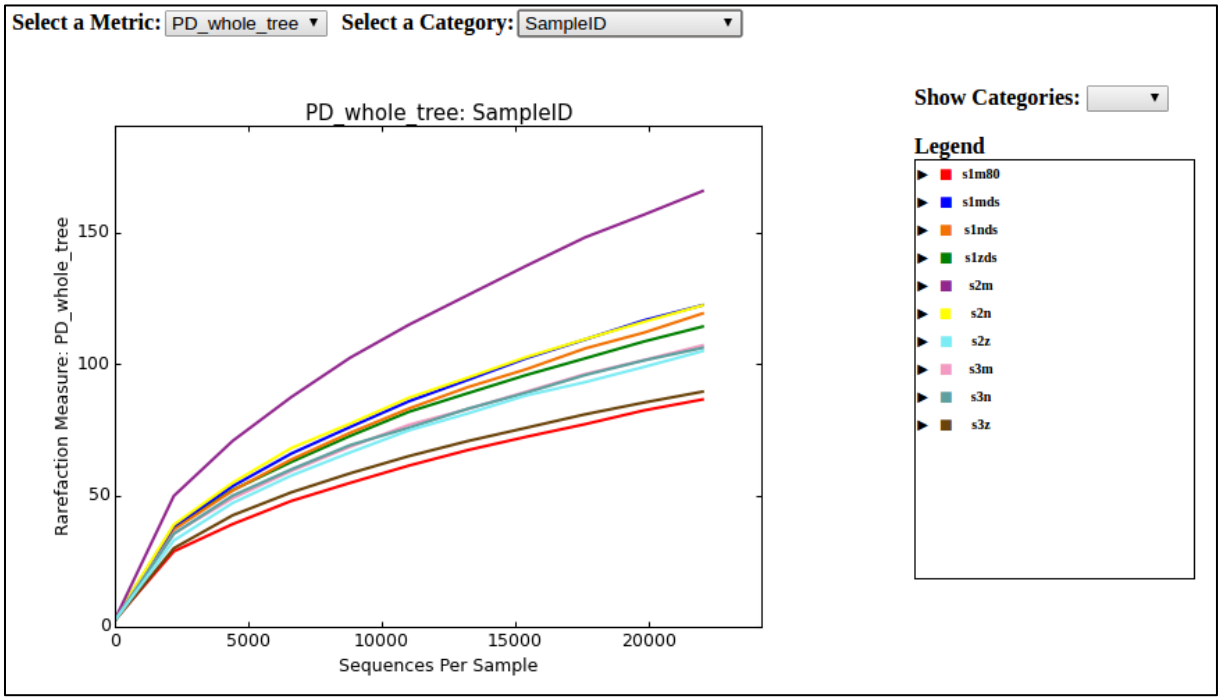


Slika 18. Krivulje razrjeđenja u *chao1* metrici

Zadnja korištena metrika za određivanje α-raznolikosti je *PD whole tree* čiji se graf može vidjeti na slici 19. Kao i na prethodna dva grafa, najveća raznolikost u odnosu na razrjeđenja, zabilježena je u uzorku s2m gdje je korištena metoda izolacije Maxwell, a slijede uzorci s1mds, s2n, s3n te s2z, s3m, s1nds i s1nzds. Najmanja raznolikost zabilježenja je kod uzoraka s2s i s1m80.

U tablici 3 prikazane su točke korištene za izradu krivulja razrjeđenja u sve tri metrike za određivanje α-raznolikosti, koje je generirao QIIME program. Prvi stupac predstavlja oznaku

uzoraka, a drugi stupac predstavlja broj sekvenci po uzorku počevši od 2000 do 22 000 prema uputama unesenima u terminal QIIME programa. Nakon toga slijede izračunate mjere razrjeđenja i eventualne pogreške za svaku metriku, kojih u našem slučaju nema. Iz tih podataka QIIME stvara već prikazane krivulje radi vizualizacije dobivenih rezultata.



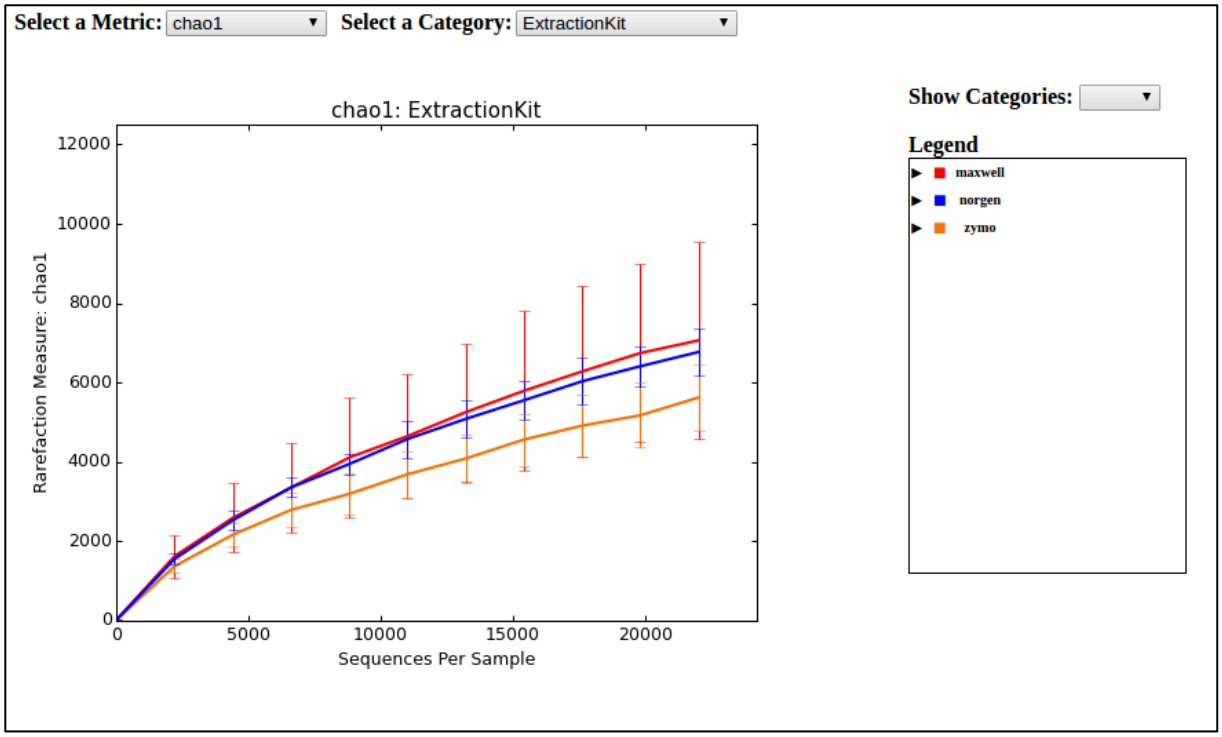
Slika 19. Krivulje razrjeđenja u PD whole tree metrici

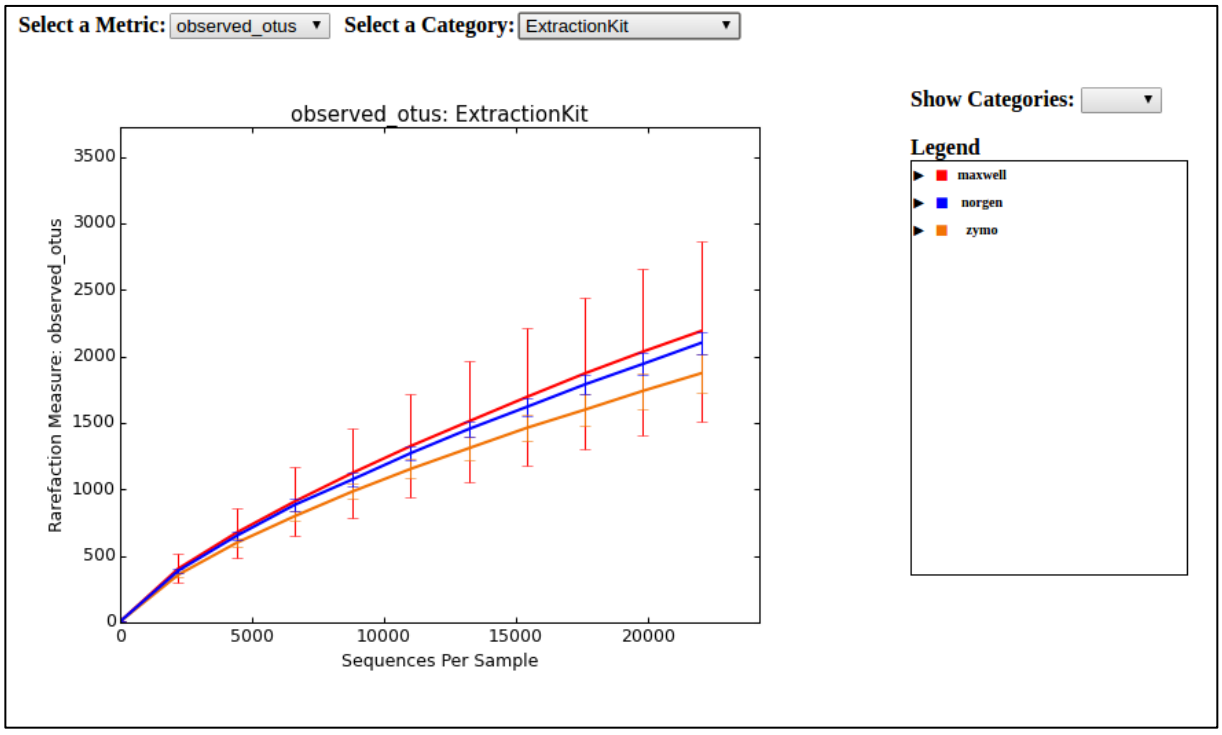
Tablica 4. Podaci potrebni za crtanje krivulja razrjeđenja u *PD whole tree*, *chao1* i *observed otus* metrikama

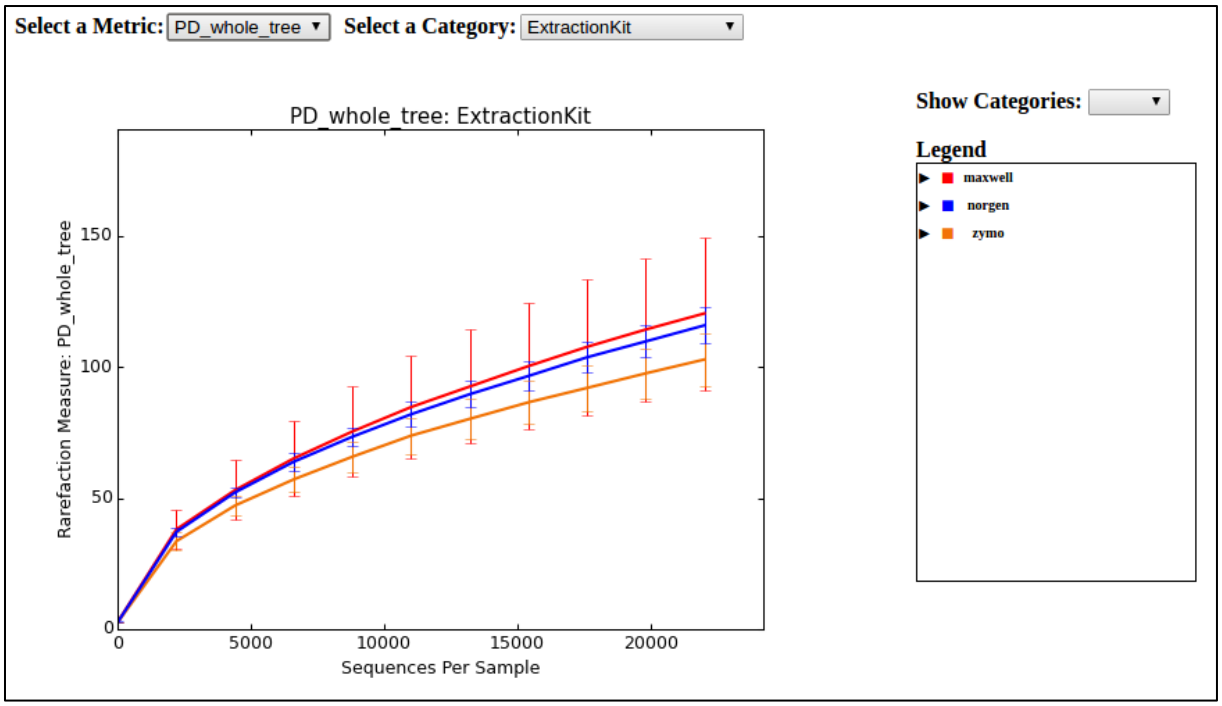
SampleID	Seqs/Sample	PD_whole_tree Ave.	PD_whole_tree Err.	chao1 Ave.	chao1 Err.	observed_otus Ave.	observed_otus Err.
s1m80	10.0	2.795	nan	20.600	nan	7.900	nan
s1m80	2209.0	28.716	nan	951.360	nan	267.000	nan
s1m80	4408.0	39.060	nan	1529.498	nan	431.700	nan
s1m80	6607.0	47.885	nan	2089.680	nan	576.700	nan
s1m80	8806.0	54.749	nan	2332.448	nan	695.900	nan
s1m80	11005.0	61.322	nan	2804.800	nan	832.400	nan
s1m80	13204.0	67.213	nan	3274.909	nan	932.300	nan
s1m80	15403.0	72.296	nan	3457.171	nan	1050.100	nan
s1m80	17602.0	77.091	nan	3787.299	nan	1159.500	nan
s1m80	19801.0	82.319	nan	4102.450	nan	1261.800	nan
s1m80	22000.0	86.498	nan	4232.875	nan	1354.900	nan
s1mds	10.0	2.675	nan	21.300	nan	8.100	nan
s1mds	2209.0	38.263	nan	1745.556	nan	396.400	nan
s1mds	4408.0	53.494	nan	2624.648	nan	668.300	nan
s1mds	6607.0	65.904	nan	3565.945	nan	904.100	nan
s1mds	8806.0	76.001	nan	4382.549	nan	1115.900	nan
s1mds	11005.0	85.786	nan	5239.002	nan	1319.700	nan
s1mds	13204.0	93.915	nan	5609.771	nan	1507.800	nan
s1mds	15403.0	102.193	nan	6211.464	nan	1691.500	nan
s1mds	17602.0	109.308	nan	6663.113	nan	1865.600	nan
s1mds	19801.0	116.564	nan	7174.047	nan	2030.800	nan
s1mds	22000.0	122.349	nan	7473.178	nan	2183.900	nan
s1nds	10.0	2.428	nan	18.500	nan	7.800	nan
s1nds	2209.0	37.039	nan	1423.552	nan	367.800	nan
s1nds	4408.0	51.799	nan	2432.963	nan	609.600	nan
s1nds	6607.0	63.666	nan	3309.955	nan	828.400	nan
s1nds	8806.0	73.808	nan	4045.628	nan	1014.200	nan
s1nds	11005.0	83.035	nan	4958.926	nan	1218.200	nan
s1nds	13204.0	91.147	nan	5399.425	nan	1390.300	nan
s1nds	15403.0	98.007	nan	6002.070	nan	1551.600	nan
s1nds	17602.0	105.899	nan	6612.567	nan	1718.400	nan
s1nds	19801.0	111.903	nan	6849.863	nan	1864.200	nan
s1nds	22000.0	119.206	nan	7412.946	nan	2032.400	nan
s1zds	10.0	2.865	nan	20.700	nan	8.300	nan
s1zds	2209.0	37.977	nan	1603.383	nan	377.700	nan
s1zds	4408.0	52.013	nan	2354.235	nan	617.800	nan
s1zds	6607.0	62.712	nan	3240.450	nan	823.900	nan
s1zds	8806.0	72.600	nan	3818.011	nan	1032.100	nan
s1zds	11005.0	81.788	nan	4364.391	nan	1206.500	nan
s1zds	13204.0	88.844	nan	4696.458	nan	1378.200	nan
s1zds	15403.0	95.803	nan	5341.281	nan	1549.800	nan
s1zds	17602.0	102.126	nan	5703.769	nan	1701.800	nan
s1zds	19801.0	108.536	nan	6016.616	nan	1856.500	nan
s1zds	22000.0	114.217	nan	6535.133	nan	1995.300	nan

s2m	10.0	2.915	nan	32.200	nan	8.400	nan
s2m	2209.0	49.796	nan	2423.516	nan	570.300	nan
s2m	4408.0	70.706	nan	3933.913	nan	959.500	nan
s2m	6607.0	87.413	nan	5069.453	nan	1302.000	nan
s2m	8806.0	102.365	nan	6394.849	nan	1642.200	nan
s2m	11005.0	114.901	nan	6880.625	nan	1920.800	nan
s2m	13204.0	126.156	nan	7907.343	nan	2213.100	nan
s2m	15403.0	137.354	nan	8853.687	nan	2484.100	nan
s2m	17602.0	148.167	nan	9591.222	nan	2754.300	nan
s2m	19801.0	156.782	nan	10154.163	nan	3001.200	nan
s2m	22000.0	165.784	nan	10872.463	nan	3238.300	nan
s2n	10.0	3.041	nan	29.500	nan	8.900	nan
s2n	2209.0	38.869	nan	1747.368	nan	401.600	nan
s2n	4408.0	54.786	nan	2888.490	nan	680.400	nan
s2n	6607.0	67.985	nan	3683.313	nan	939.500	nan
s2n	8806.0	77.289	nan	4215.820	nan	1133.100	nan
s2n	11005.0	87.133	nan	4867.770	nan	1344.200	nan
s2n	13204.0	94.764	nan	5431.328	nan	1533.600	nan
s2n	15403.0	102.623	nan	5804.565	nan	1716.400	nan
s2n	17602.0	109.421	nan	6269.071	nan	1890.700	nan
s2n	19801.0	116.011	nan	6691.196	nan	2060.100	nan
s2n	22000.0	122.280	nan	6935.861	nan	2221.300	nan
s2z	10.0	2.540	nan	14.270	nan	7.000	nan
s2z	2209.0	32.833	nan	1345.977	nan	357.000	nan
s2z	4408.0	47.084	nan	2391.298	nan	619.900	nan
s2z	6607.0	57.530	nan	2949.963	nan	826.200	nan
s2z	8806.0	66.335	nan	3224.213	nan	1016.100	nan
s2z	11005.0	74.714	nan	3768.288	nan	1205.200	nan
s2z	13204.0	81.148	nan	4300.925	nan	1364.600	nan
s2z	15403.0	88.037	nan	4622.994	nan	1520.000	nan
s2z	17602.0	93.079	nan	5162.637	nan	1665.500	nan
s2z	19801.0	98.858	nan	5440.879	nan	1813.700	nan
s2z	22000.0	104.884	nan	5789.863	nan	1956.000	nan
s3m	10.0	2.742	nan	18.300	nan	8.000	nan
s3m	2209.0	35.565	nan	1394.899	nan	388.500	nan
s3m	4408.0	48.903	nan	2284.549	nan	640.800	nan
s3m	6607.0	59.314	nan	2702.027	nan	853.600	nan
s3m	8806.0	68.500	nan	3344.521	nan	1044.900	nan
s3m	11005.0	76.828	nan	3684.163	nan	1232.100	nan
s3m	13204.0	82.843	nan	4229.603	nan	1394.100	nan
s3m	15403.0	89.464	nan	4655.719	nan	1556.000	nan
s3m	17602.0	96.141	nan	5090.963	nan	1711.800	nan
s3m	19801.0	101.466	nan	5552.254	nan	1851.300	nan
s3m	22000.0	107.067	nan	5681.177	nan	1986.900	nan
s3n	10.0	2.577	nan	18.450	nan	8.300	nan
s3n	2209.0	35.407	nan	1510.221	nan	400.200	nan
s3n	4408.0	49.891	nan	2281.608	nan	662.200	nan
s3n	6607.0	59.906	nan	3086.259	nan	883.800	nan
s3n	8806.0	69.111	nan	3588.238	nan	1080.500	nan
s3n	11005.0	75.728	nan	3913.899	nan	1256.000	nan
s3n	13204.0	82.919	nan	4428.556	nan	1438.800	nan
s3n	15403.0	89.071	nan	4857.514	nan	1594.500	nan
s3n	17602.0	95.775	nan	5220.708	nan	1760.300	nan
s3n	19801.0	101.341	nan	5697.040	nan	1907.500	nan
s3n	22000.0	106.116	nan	5969.053	nan	2051.400	nan
s3z	10.0	2.421	nan	21.300	nan	7.200	nan
s3z	2209.0	29.893	nan	1178.006	nan	344.100	nan
s3z	4408.0	42.411	nan	1757.051	nan	557.700	nan
s3z	6607.0	51.156	nan	2189.100	nan	745.900	nan
s3z	8806.0	58.390	nan	2545.666	nan	907.100	nan
s3z	11005.0	64.934	nan	2939.502	nan	1051.900	nan
s3z	13204.0	70.651	nan	3263.407	nan	1186.500	nan
s3z	15403.0	75.723	nan	3729.754	nan	1319.700	nan
s3z	17602.0	80.837	nan	3878.312	nan	1432.100	nan
s3z	19801.0	85.310	nan	4072.896	nan	1553.500	nan
s3z	22000.0	89.470	nan	4547.083	nan	1668.300	nan

Nakon grafova u kojima smo promatrali α-raznolikost u svakom uzorku, QIIME daje opciju generiranja takvih grafova s obzirom na metodu izolacije koja je unešena kao podatak u *mapping* dokumentu. Slika 20 prikazuje raznolikost prema metodi korištenja u svim uzorcima.







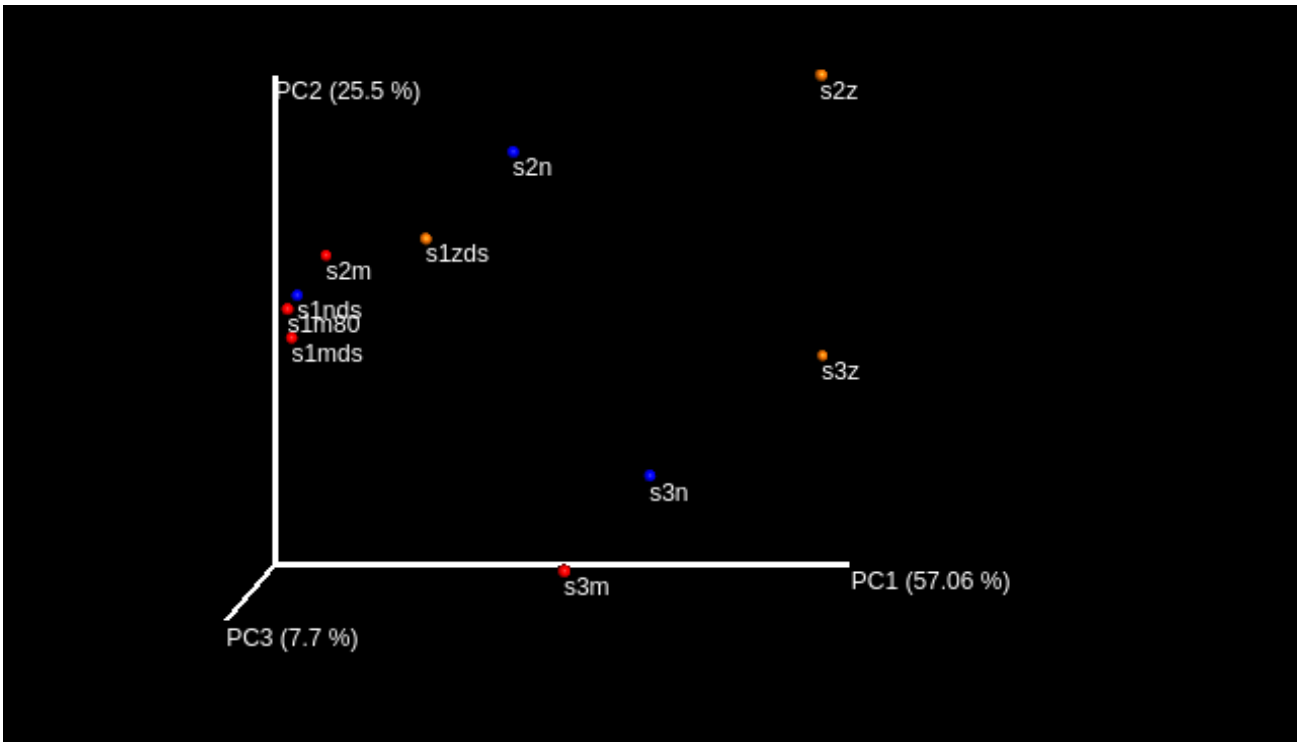
Slika 20. Krivulje razrjeđenja: a) *chao1*, b) *observed otus*, c) *PD whole tree*

Sve tri korištene metrike pokazuju kako se, u prosjeku, najveća alfa raznolikost dobiva korištenjem Maxwell metode izolacije DNA, ali uz najveći iznos varijacija od svih metoda . Gledajući rezultate pojedinačno po uzorcima, vidimo kako metoda izolacije po Maxwellu u određenim uzorcima daje veću raznolikost, dok u drugima puno manju. Druga metoda izolacije koja generira najveću raznolikost je Norgen, a zadnja po raznolikosti je Zymo.

4.3. BETA RAZNOLIKOST

Uz α-raznolikost određena je i β-raznolikost koja predstavlja usporedbu sličnosti uzoraka međusobno baziranu na određenim taksonomskim vrstama u svakom uzorku. Ona određuje mjeru udaljenosti ili različitosti između svih uzoraka međusobno. U ovom radu korištene su dvije metrike za određivanje β-raznolikosti dostupne u QIIME-u – (i) ″Weighted UniFrac″ (slika 21) i (ii) ″Un-Weighted UniFrac″ (slika 22). ″UniFrac″ izračunava udaljenost između parova uzoraka tako da smješta sve taksonomske jedinice prisutne u jednom ili oba uzorka na filogenetsko stablo. Grana stabla koja vodi do taksonomske jedinica zajedničke za oba uzorka se označava kao dijeljena dok ona koja vodi do taksonomske jedinice prisutne u samo jednom uzorku kao nedijeljena. Udaljenost između uzoraka se izračunava kao omjer dužine nedijeljenih grana i ukupne dužine grana (dijeljene + nedijeljene). Razlika između ″Weighted″ i ″Un-Weighted UniFrac″ metrike je u tome da ″Weighted UniFrac″ kod izračunavanja udaljenosti između uzoraka uzima u obzir i zastupljenost pojedinih taksonomskih jedinica, a ne samo njihovu prisutnost. ″Principal Coordinates Analysis″ (PCoA) je metoda korištena za istraživanje i sličnosti i različitosti podataka beta raznolikosti. Bazira se na matrici sličnosti (ili udaljenosti) i za svaku vrijednost dodjeljuje lokaciju u, našem slučaju, u trodimenzionalnom prostoru.

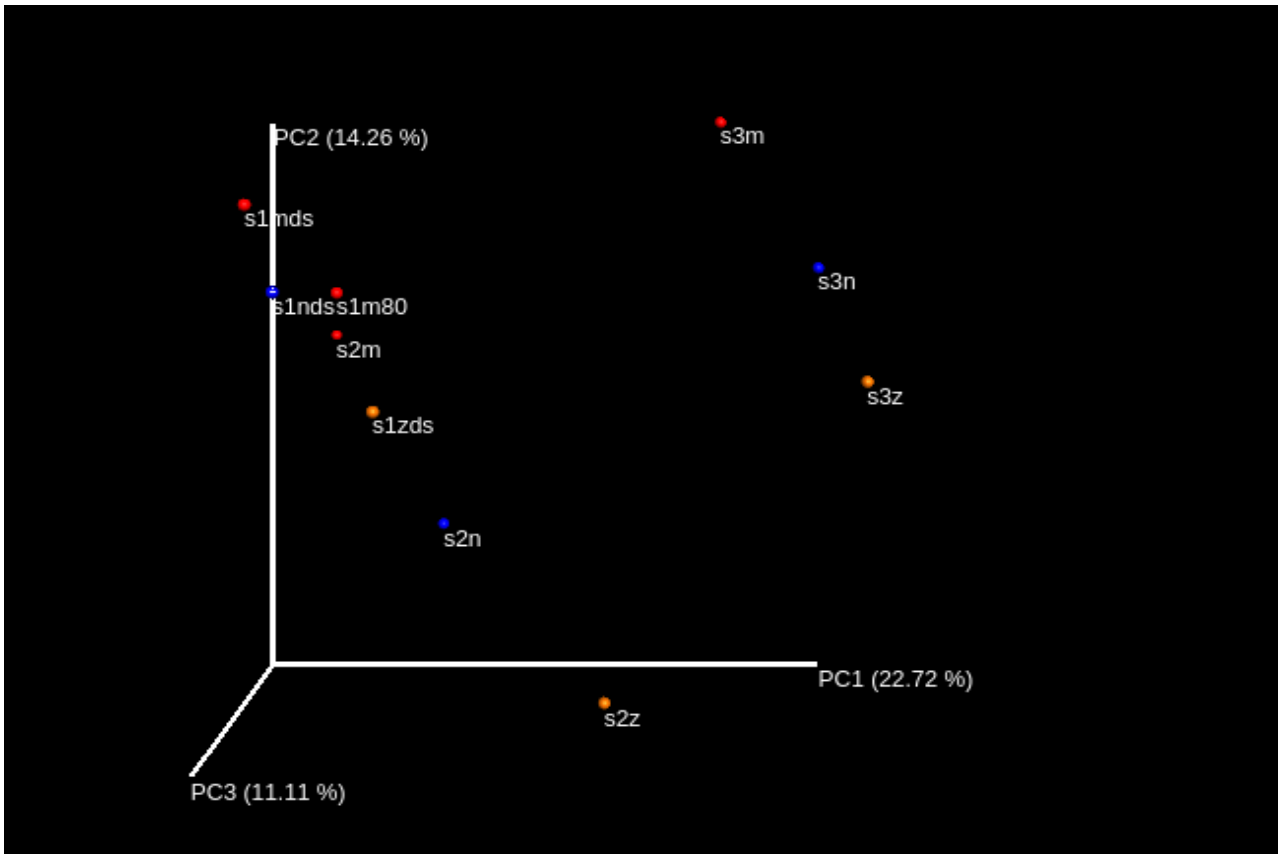
″Weighted UniFrac″ mjera prilagođena je za otkrivanje razlika zastupljenosti mikroorganizama, čak i kada skupine organizama prisutne u svakom uzorku ostaju iste.



Slika 21. PCoA prikaz β-raznolikosti Weighted metodom

PCoA analiza pokazuje da je β-raznolikost uzoraka s1mds, s1m80, s1nds vrlo slična kao i uzoraka s2m, s1zds i s2n. To se poklapa sa rezultatima na slici15 gdje se vidi da uzorci s1mds, s1m80, s1nds dijele veliki broj istih bakterijskih vrsta. Uzorci s2m, s1zds i s2n nisu tako usko grupirani zbog manje sličnosti u broju bakterijskih vrsta. Uzorak s2z ima najmanje zajedničkih bakterijskih vrsta u odnosu na ostale uzorke. Uzorci s3m i s3n s3z također imaju malo zajedničkih bakterijskih vrsta u odnosu na ostale uzorke, ali se ipak blago grupiraju što se također poklapa sa slikom15.

″Un-Weighted UniFrac″ mjera prilagođena je za detektiranje razlike u prisutnosti ili odsutnosti vrsta bakterija u različitim uzorcima. Rezultati se, očekivano, razlikuju od rezultata dobivenih Weighted mjerom, jer su to dvije metode (kvalitativna i kvantitativna) zasnovane na različitim temeljima te je za očekivati da se isti uzorci u potpunosti ne podudaraju. Ipak, na slici 23 uzorci su grupirano slično kao i ″Weighted″ methodom. Najveću međusobnu sličnost imaju uzorci s1mds, s1m80, s1nds i s2m te s1dzs. S obzirom na malu udaljenost od uzorka s1zds, možemo zaključiti da uzorak s2m ima veću sličnost u prisutnosti bakterijskih vrsta od primjerice uzorka s2z. Uzorci s3m, s3n i s3z imaju najveću β- raznolikost u odnosu na ostale uzorke, ali se međusobno pokazuju sličnosti i blago grupiranje.



Slika 22. PCoA prikaz β-raznolikosti Un-Weighted metodom

Uzorci koji su bili tretirani DNA Shield™ stabilizatorom – s1mds, s1nds i s1ds daju rezultate veće međusobne sličnosti u odnosu na uzorke koji nisu bili tretirani stabilizatorom: s2m, s2n,s2z te s3m, s3n i s3z, iako unutar pojedinog uzorka ne daju najveću izmjerenu raznolikost Takva grupacija uzoraka nameće zaključak kako DNA Shield™ stabilizator može stabilizirati uzorak prije same izolacije te time ujednačiti rezultate samog sekvencioranja. s3m, s3n i s3z uzorci također pokazuju određeno grupiranje s obzirom na međusobnu sličnost. Također, unutar samih uzoraka nema velike bakterijske raznolikosti. Uzorak s2m pokazuje najveću bakterijsku raznolikost unutar samog uzorka, ali uspoređujući s ostalim metodama izolacije u uzorcima s2n i s2z, nema međusobne sličnosti iako su sva tri uzorka uzeta iz zajedničkog početnog uzorka. Čuvanje i skladištenje fecesa na -80°C daje najmanju alfa raznolikost, ali veliku sličnost s uzorcima tretiranim DNA Shield™ stabilizatorom.

4.4. STATISTIČKA ANALIZA

Statistička obrada podataka pokazuje kako grupiranje uzoraka po metodi izolacije nije statistički značajno što se može pretpostaviti iz prethodno dobivenih rezultata. QIIME program

nudi opciju ″compare categories″(eng. usporedba kategorija) koja provjerava je li grupiranje uzoraka statistički značajno. Može se izabrati između dvije opcije: (i)″adonis″ i (ii) ″ANOSIM″, a u ovom radu korištena je ″ANOSIM″. Bazira se na ANOVA statističkim testovima. Naredba koja se unosi u terminalni prozor glasi:

```
compare_categories.py --method anosim -i unweighted_unifrac_dm.txt -m mapping_file_dorada_2sub.txt -c ExtractionKit -o anosim_output2sub
```

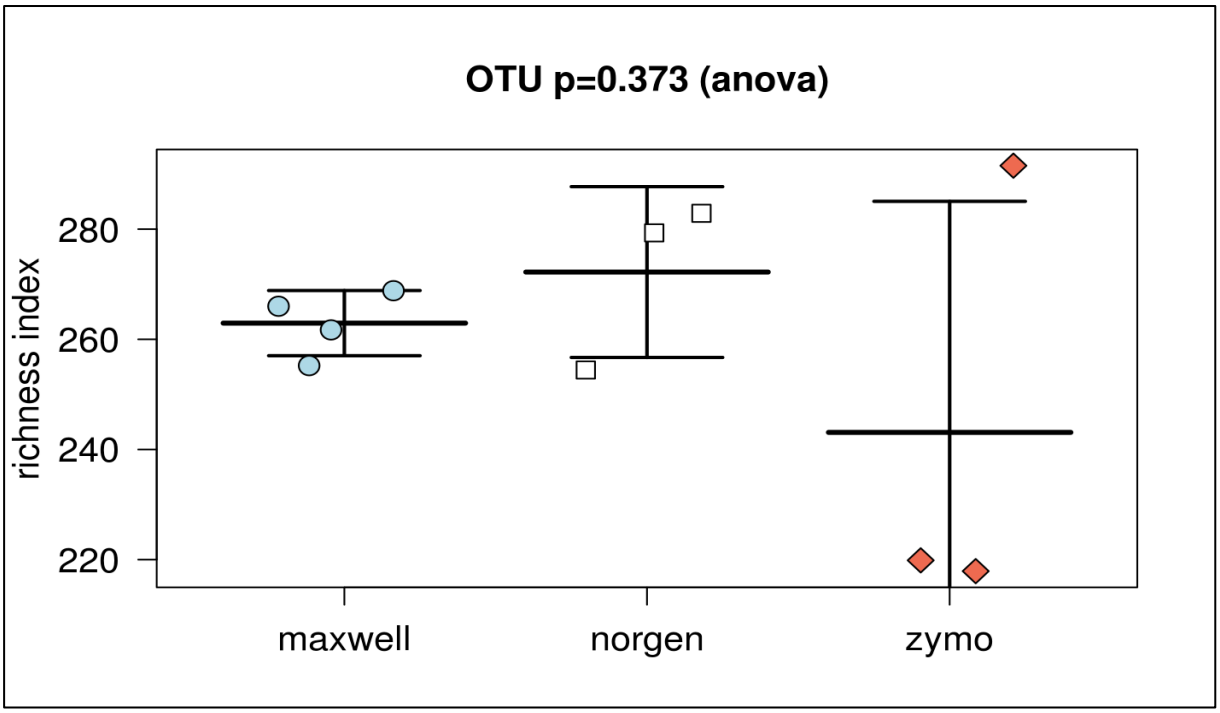
Naredba koristi ″Un-Weighted UniFrac″ podatke dobivene prethodnim beta analizama kako bi se izračunala statistička značajnost grupiranja. Rezultat je P vrijednost koja iznosi 0,374 (>0,05) što znači da grupiranje uzoraka prema metodi izolacije DNA(*ExtractionKit*) nije statistički značajno.

Kako bi se provjerila statistička značajnost grupiranja uzoraka prema osobi koja je donirala uzorke, ponovno je korištena ranije navedena naredba u programu QIIME – compare diversity.py, s razlikom u kategoriji usporedbe. U ovom slučaju kategorija je postavljena na ˙Subject˙ što označava osobe koje su donirale uzorke. Kao što je navedeno ranije u radu, uzorci 1 i 2 su od fecesa iste osobe, a uzorak 3 je od druge osobe. Iako su uzorci 1 i 2 od iste osobe, treba uzeti u obzir kako se mikrobiota mijenja ovisno o različitim vanjskim utjecajima, stoga je njen sastav sličan, a ne potpuno identičan iako su uzorci od iste osobe. Potpuna naredba unesena u terminal glasi:

```
compare_categories.py --method anosim -i unweighted_unifrac_dm.txt -m mapping_file_dorada_2sub.txt -c Subject -o anosim_output2sub
```

Rezultat je P vrijednost koja iznosi 0,014 (<0,05) što pokazuje kako je grupiranje uzoraka prema osobi koja je dala uzorke statistički značajno. Iznos R vrijednost od 0,86 (0-1) međutim potvrđuje da postoji nesličnost između grupa, što je i vidljivo iz rezultata beta raznolikosti. Ove analize potvrđuju prethodno dobivene rezultate.

U svrhu statističke analize, u ovom radu je korištena i web aplikacija Calypso. Aplikacija, uz analizu, nudi i vizualizaciju rezultata ″Stripchartom″. Koriste se različiti indeksi koji koriste različite metrike kako bi izračunali mikrobnu raznolikost pojedinog uzorka - alfa raznolikost. U ovom radu korišten je Richness indeks koji u obzir osim ukupne raznolikosti mikrobnog sastava, uzima u obzir i broj prisutnih taksonomskih grupa ili OTU-a, što znači da je kvantitativan.



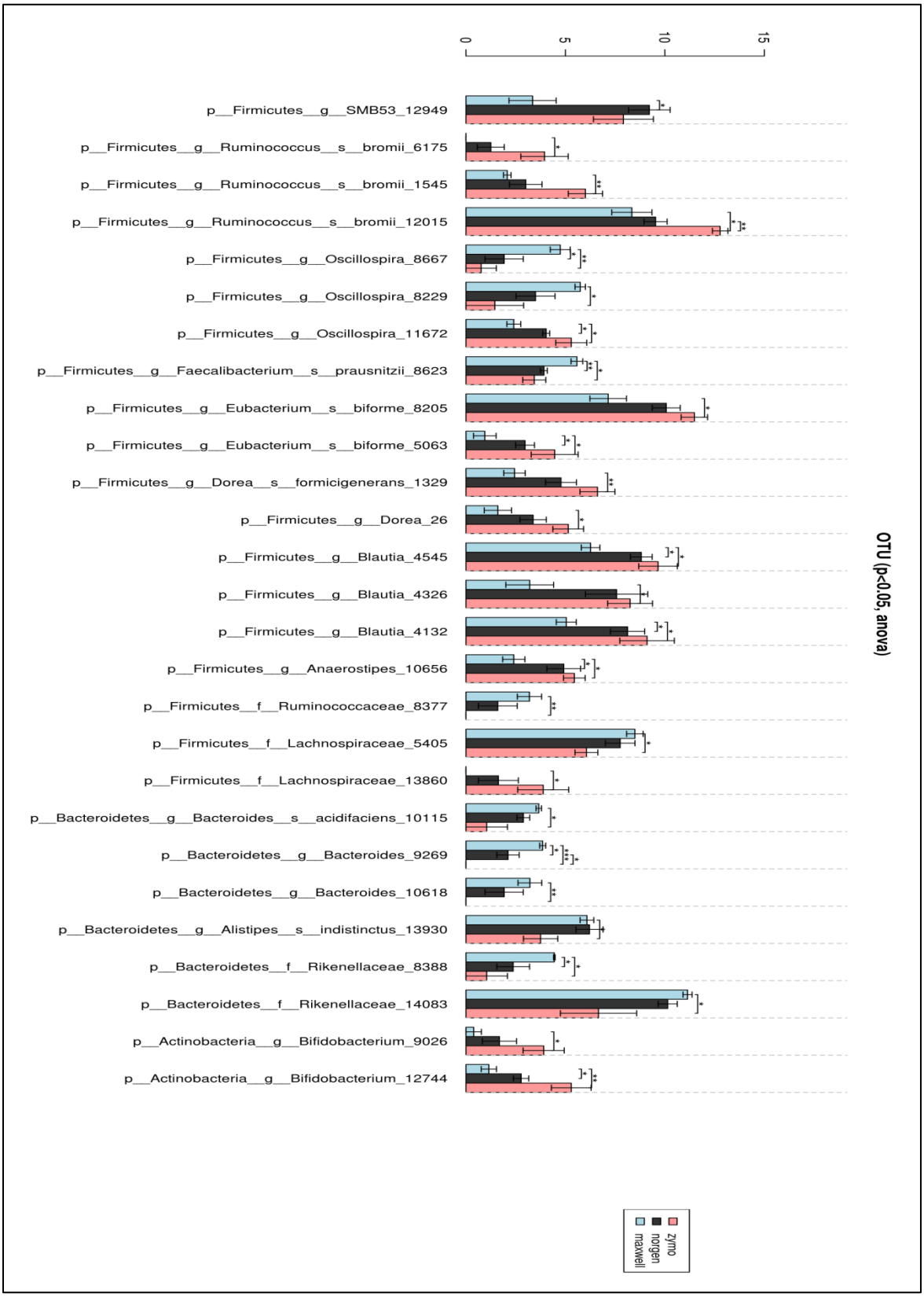
Slika 23. Stripchart alfa raznolikosti na temelju metode izolacije DNA

Slika 23 prikazuje ″Stripchart″ alfa raznolikosti te P vrijednost koja iznosi 0,373 (>0,05). Iz toga se može zaključiti kako upotreba različitih metoda izolacije DNA nema statistički značajan utjecaj na alfa raznolikost uzoraka. Također se može zaključiti kako metoda Maxwell daje konzistentnije rezultate za razlike od metoda Norgen te pogotovo Zymo koje pokazuju veći rasap procjene broja OTU-ova.

Pomoću Calypso aplikacije usporedili smo i raznolikost taksonomskih grupa ili OTU-a između uzoraka – beta raznolikost. Na slici 24 prikazana je usporedba zastupljenosti OTU-a između metoda izolacija. Oznaka *** označava p<0.001 odnosno najveću raznolikost između metoda. U ovom radu najveća je raznolikost između metoda Zymo i Maxwell pri identifikaciji za rod(genus) Bacteroides_9269. Oznaka ** označava p<0.01 odnosno značajnu raznolikost između metoda. Metode Zymo i Maxwell pokazuju značajnu raznolikost za rod Bifidobacterium_12744, rod Ruminococcus vrsta *bromii_1545*, rod Dorea vrsta formicigerans_1329, koljeno Ruminoicoccacea, rod Bacteroides_10618 te rod Oscillospira_8229. Metode Norgen i Maxwell pokazuju značajnu raznolikost za rod Faecilabacterium dok je značajna raznolikost između metoda Zymo i Norgen vidljiva kod roda Ruminococcus vrsta *bromii_1205*.

Usporedba raznolikosti na razini koljena pokazuje vrlo značajne razlike između metoda Zymo i Maxwell za koljeno Actinobacteria, Bacteroidetes i Firmicutes što se podudara s rezultatima

na slici 15. Značajna razlika između metoda Norgen i Maxwell je na razini koljena Firmicutes kao i između metoda Zymo i Norgen. Iz ove analize zaključuje se kako najveće razlike postoje između metoda Zymo i Maxwell što se može i potvrditi na slikama 22 i 23 koje prikazuju beta raznolikost pomoću PCoA analize u QIIME-u.



Slika 24. Beta raznolikost pomoću Calypso aplikacije

5. ZAKLJUČCI

Na temelju provedenih analiza, mogu se izvesti sljedeći zaključci:

1. Metode za izolaciju DNA iz uzoraka fecesa mogu imati značajan utjecaj na konačni rezultat određivanja sastava mikrobiote sekvencioniranjem 16S rRNA gena.
2. Korištenje dodataka za stabilizaciju uzorka ujednačuje raznolikost unutar svih uzoraka izoliranih sa trima korištenim metodama
3. Uzorci pod rednim brojem dva (s2m, s2n, s2z) imaju najmanju međusobnu sličnost u bakterijskoj raznolikosti od sve tri promatrane metode
4. U prosjeku, najveću konzistentnost odnosno sličnost sastava mikrobiote unutar testiranih uzoraka daje Maxwell metoda (Maxwell® 16 Tissue DNA Purification Kit), slijedi Norgen (Stool DNA Isolation Kit (Cat. 27600) Norgen Biotek Corp) i na kraju Zymo (Quick-DNA™ Fecal/Soil Microbe Miniprep Kit Zymo Research)
5. Statistička analiza potvrđuje značajno grupiranje uzoraka prema donoru (osobi) što nije slučaj za grupiranje uzoraka prema metodi izolacije DNA.
6. Najznačajnije razlike u raznolikosti između uzoraka pokazuju metode Maxwell i Zymo.
7. S obzirom na UniFrac udaljenosti istih uzoraka izoliranih različitim metodama za daljna istraživanja poželjno je koristiti jednu metodu kako bi se smanjila pogreška odnosno dobili sastavi mikrobiote koji su međusobno usporedivi.

6. LITERATURA

1. Xu, Z. i Knight, R., (2015) Dietary effects on human gut microbiome diversity, *Br. J. Nutr.* **113**,S1-S5
2. Keeney, K. M., i Finlay, B. B. (2011). Enteric pathogen exploitation of the microbiota-generated nutrient environment of the gut., *Curr Opin Microbiol* **14**(1), 92–98. http://doi.org/10.1016/j.mib.2010.12.012
3. Lozupone, C. A., i sur. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, **489**(7415), 220–230. http://doi.org/10.1038/nature11550
4. Yuan S, i sur., (2012) Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. *PLoS ONE* **7**(3): http://doi: 10.1371/journal.pone.0033865
5. BioDinaMik, < http://biodinamik.pbf.hr/opis-projekta/>, Pristupljeno 20.02.2017.
6. Zhou, L., i Foster, J. A. (2015). Psychobiotics and the gut–brain axis: in the pursuit of happiness, *Neuropsychiatr Dis and Treatt* **11**, 715–723. http://doi.org/10.2147/NDT.S61997
7. WHO (2017) World Health Organization, Fact Sheet, http://www.who.int/mediacentre/factsheets/fs311/en/, Pristupljeno 20.02.2017.
8. Foster, J.A., McVey Neufeld, K.A. (2013) Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci.*; **36**(5):305–312
9. Keeney M.K., Finlay B.B., (2011) Enteric pathogen exploitation of the microbiota-generated nutrient environment of the gut**.** *Curr Opin Microbiol*, **14**:92-98
10. Qin, J. i sur. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65.
11. Turnbaugh, P. J. i sur. (2009) A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484.
12. Arumugam, M. i sur. (2011) Enterotypes of the human gut microbiome. *Nature* **473**: 174–180
13. Knights, D. i sur., (2014) Rethinking Enterotypes. *Cell Host Microbe* **16(4)** 433-437.
14. Stoner, L., i Cornwall, J. (2014). Did the American Medical Association make the correct decision classifying obesity as a disease? *Australas Med J*, **7(11),** 462–464.
15. Backhed, F. i sur. (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl. Acad. Sci. USA* **101 (44),** 15718–15723.

16. Shen, J., Obin, M.,S., Zhao, L. (2012) The gut microbiota, obesity and insulin resistance, *Mol Aspects Med* **34(1)**:39-58
17. Arora, T., i sur. (2011) Propionate. Anti-obesity and satiety enhancing factor? *Appetite* **56**, 511–515
18. Berthoud, H.-R. (2008). The vagus nerve, food intake and obesity. *Regul Pept.*, **149(1-3)**, 15–25.
19. Den Besten, G. i sur., (2013) The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res*, **54(9)**, 2325–2340.
20. Aguirre, M. i Venema, K., 2015. Does the Gut Microbiota Contribute to Obesity? Going beyond the Gut Feeling. *Microorganisms*, **3(2)**, pp.213–235
21. Ley, R.E., i sur. (2006) Microbial ecology: Human gut microbes associated with obesity. *Nature* **444**, 1022–1023.
22. Turnbaugh, P.J. i sur (2009)., A core gut microbiome in obese and lean twins. *Nature* **457(7228),** 480–484
23. Walker, A.W. i Parkhill, J., (2013) Fighting obesity with bacteria. *Science*, **341**, 1069-1070
24. Kennedy N.A., i sur.,(2014) The Impact of Different DNA Extraction Kits and Laboratories upon the Assessment of Human Gut Microbiota Composition by 16S rRNA Gene Sequencing. PLOS ONE 9(2): e88982.
25. QIIME(2017), *beta_diversity.py – Calculate beta diversity (pairwise sample dissimilarity) on one or many otu tables*¶ QIIME – Quantitaive Insights Into Microbial Ecology <http://qiime.org/scripts/beta_diversity.html > Pristupljeno: 01.04.2017.
26. QIIME(2017) OTU picking strategies in QIIME Quantitaive Insights Into Microbial Ecology. <http://qiime.org/tutorials/otu_picking.html> Pristupljeno: 10.04.2017.
27. Werner, J., (2017): Alfa diversity, WernerLab.org http://www.wernerlab.org/teaching/qiime/overview/e Pristupljeno: 11.04.2017
28. Shimadzu Analytical and Measuring Instruments(2017), BioSpec-nano<http://www.shimadzu.com/an/molecular_spectro/uv/spc/bionano3.html> Pristupljeno: 01.04.2017.
29. QIIME (2017) *QIIME tutorials* QIIME – Quantitaive Insights Into Microbial Ecology <http://qiime.org/tutorials/index.html> Pristupljeno: 20.03.2017.
30. McOrist, A.L. i sur., 2002. A comparison of five methods for extraction of bacterial DNA from human faecal samples. *J Microbiol Methods*, **50(2)**, 131–139.

31. Liu, W. i sur. (2012) Popular Molecular Markers in Bacteria. *Mol. Genet., Microbiol. and Virol*. **27**, 103-107.
32. Roux, S. i sur. (2011) Comparison of 16S rRNA and protein-coding genes as molecular markers for assessing microbial diversity (*Bacteria* and *Archaea*) in ecosystems. *FEMS Microbiol. Ecol*. **78**, 617-628.
33. Clarridge, J.E., 3rd, (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin. Microbiol. Rev*, **17(2)** 840–862.
34. Anonymous,(2017) <https://www.google.hr/search?q=%3Chttp://www.alimetrics.net/en/index.php/dna-sequenceanalysis%3E.&source=lnms&tbm=isch&sa=X&ved=0ahUKEwjTwpXYwoDTAhXDchQKHXwWDfYQ_AUICCgB&biw=1280&bih=694#imgrc=rQbAYDP5DwxQkM> Pristupljeno: 27.03.2017.
35. Illumina (2016) *16S Metagenomic Sequencing library Preparation*, <http://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf>. Pristupljeno: 27.03.2017.
36. Illumina (2010) *Illumina Sequencing Technology*, <http://www.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf>. Pristupljeno: 27.03.3017.
37. Illumina (2017), <https://www.illumina.com/systems/sequencing-platforms/miseq.html> Pristupljeno: 27.03.2017.
38. UCONN (2016) *Reversible Dye Termination Sequencing by synthesis Technology*, UCONN – University of Connecticut, Connecticut, < http://cgi.uconn.edu/illumina-miseq/>. Pristupljeno: 28.03.2017.
39. Caporaso, J.G. i sur., (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods,* **7(5)**, 335–336.
40. Cock, P. J. A. i sur. (2010) The Sanger FASTQ file format for sequences with quality scores, and te Solexa/Illumina FASTQ variants, *Nucl. Acids Res*. **38**, 1767-1771.
41. Illumina (2017) *Paired-End Sequencing*. <http://www.illumina.com/technology/next-generation-sequencing/paired-end-sequencing_assay.html>. Pristupljeno 29.03.2017.
42. Anonymous (25. siječnja 2016), <http://www.yourgenome.org/facts/how-do-you-put-a-genome-back-together-after-sequencing> Pristupljeno: 29.03.2017.

43. Promega (2016) *Maxwell® 16 DNA Purification Kits*, <https://www.promega.com/~/media/files/resources/protocols/technical%20manuals/0/maxwell%2016%20dna%20purification%20kits%20protocol.pdf>. Pristupljeno: 30.03.2017.
44. Zymo Research (2017) *Quick-DNA™ Fecal/Soil Microbe Miniprep Kit,* <http://www.zymoresearch.com/downloads/dl/file/id/91/d6010i.pdf> Pristupljeno: 30.03.2017.
45. Norgen Biotek (2017) *Stool DNA Isolation Kit,* <https://norgenbiotek.com/sites/default/files/resources/Stool-DNA-Kit-Manual-PI27600-7.pdf> Pristupljeno: 30.03.2017.
46. Illumina (2016) *Multiplex Sequencing Assay*, <http://www.illumina.com/technology/next-generation-sequencing/multiplexing-sequencing-assay.html>. Pristupljeno 31.03.2017.
47. QIIME (2015) *Input Files*. QIIME – Quantitaive Insights Into Microbial Ecology, <http://qiime.org/documentation/file_formats.html#metadata-mapping-files>. Pristupljeno: 31.03.2017.
48. Virtual Box, (2017) <https://www.virtualbox.org/manual/ch02.html#installation_windows> Pristupljeno: 01.03.2017.
49. Edgar, R. C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinf.* **26**, 2460-2461.
50. Flint, H.J. i sur., (2012) The role of the gut microbiota in nutrition and health. *Nat. Rev. Gastroenterol. Hepatol.*, **9(10),** 577–589.
51. Kuczynski, J.,i sur., (2012). Using QIIME to Analyze 16S rRNA Gene Sequences from Microbial Communities. *Curr Protoc in Microbiol 0 1*, Unit–1E.5. http://doi.org/10.1002/9780471729259.mc01e05s27
52. Vazquez-Baeza, Y. i sur. (2013) EMPeror: a tool for visualizing high-throughput microbial community data, *GigaScience*. **26**, 1-4.
53. Calypso (2017) <http://cgenome.net/wiki/index.php/Calypso_File_Upload> Pristupljeno: 30.04.2017.