

# Funkcionalna analiza metagenoma bioreaktora s djelomičnom reakcijom nitritacije i anaerobne oksidacije amonijaka

---

Zvjerković, Dora

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:393684>

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

Download date / Datum preuzimanja: **2024-10-02**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek

Dora Zvjerković

**Funkcionalna analiza metagenoma  
bioreaktora s djelomičnom reakcijom  
nitritacije i anaerobne oksidacije amonijaka**

Diplomski rad

Zagreb, 2022.

University of Zagreb  
Faculty of Science  
Department of Biology

Dora Zvjerković

# **Functional metagenomic analysis of a partial nitritation-anammox bioreactor**

Master thesis

Zagreb, 2022.

Ovaj diplomski rad izrađen je u laboratoriju za računalnu biologiju pri Zavodu za molekularnu biologiju Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu pod vodstvom doc. dr. sc. Rose Karlić te je predan na ocjenu Biološkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu radi stjecanja zvanja magistra molekularne biologije.

## ZAHVALE

Zahvaljujem mentorici doc. dr. sc. Rosi Karlič na komentarima, savjetima i vodstvu koje se nastavilo čak i tijekom njezinoga porodiljnog dopusta. Zahvaljujem asistentici mag. mol. biol. Pauli Štancl na svakom susretljivom i entuzijastičnom mejlu kojim je odgovorila na svaki moj panični (bilo ih je), te na podršci i praktičnoj pomoći bez koje ovaj rad ne bih mogla ni započeti.

Hvala mami Gordani što me razumije, voli i podržava na svaki mogući način, i kad mi ide i kad ne ide. Zahvaljujem tati Borisu na podršci i životnim savjetima, i na tome što je vjerovao u mene. Hvala im što su me uzdržavali ovoliko dugo, neću im nikad moći vratiti sve što su mi pružili.

Hvala braći Robertu i Romanu što znam da su uvijek tu i da uvijek razumiju, i kad je lako i kad je teško, i hvala im što su jako dobri, zabavni i sposobni bilo što staviti u širu perspektivu. Zahvaljujem i svim prijateljima koji su me trpili sve ove godine i svejedno se htjeli družiti sa mnom, a posebno Matei, Jeleni, Ivanu Š. i Ivanu M.

Najveće hvala Nikoli, znat će on zašto.

## TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek

Diplomski rad

### **Funkcionalna analiza metagenoma bioreaktora s djelomičnom reakcijom nitritacije i anaerobne oksidacije amonijaka**

Dora Zvjerković

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Otpadne vode prije ispuštanja u okoliš potrebno je obraditi radi uklanjanja tvari koje bi u prirodnim vodama izazvale ekološki štetnu eutrofikaciju. Za uklanjanje dušikovih spojeva uglavnom se koriste bioreaktori različitih tipova. U ovom radu korištene su bioinformatičke metode za taksonomsku i funkcionalnu analizu metagenoma bioreaktora laboratorijske veličine koji koristi djelomičnu reakciju nitritacije i anaerobne oksidacije amonijaka za uklanjanje amonijaka iz otpadnih voda koji gotovo potpuno autotrofno, tj. bez potrebe za organskim ugljikom kao ishranom za aktivne bakterije, uklanja dušik iz otpadnih voda. Pokazano je da je taksonomski sastav bioreaktora raznolik te da se u njemu nalaze u velikom broju i heterotrofne skupine vezane uz druge puteve mikrobnog ciklusa dušika, no da su enzimi ciljanih reakcija (oksidacije amonijaka u nitrit i *anammox* reakcije) svejedno snažno eksprimirani i da je ova metoda uklanjanja dušika iz otpadnih voda vjerojatno stabilna i s prisustvom brzo rastućih heterotrofnih bakterija. Poželjna bi bila slična analiza ovog tipa bioreaktora, na industrijskoj veličini i pri različitim uvjetima.

(55 stranica, 15 slika, 1 tablica, 105 literaturnih navoda, izvorni jezik: hrvatski)

Voditeljica: Doc. dr. sc. Rosa Karlić

Ocjenitelji: Doc. dr. sc. Rosa Karlić

Prof. dr. sc. Biljana Balen

Doc. dr. sc. Nenad Malenica

Zamjena: Prof. dr. sc. Kristian Vlahoviček

## **BASIC DOCUMENTATION CARD**

University of Zagreb

Faculty of Science

Division of Biology

Master thesis

### **Functional metagenomic analysis of a partial nitrification-anammox bioreactor**

Dora Zvjerković

Roosevelt square 6, 10000 Zagreb, Hrvatska

Removal of various substances from wastewater before it is discharged into the environment is necessary to ameliorate the harmful ecological effects of eutrophication on natural water bodies. Different kinds of bioreactors are commonly used for removal of nitrogen from wastewater. In this master thesis, the taxonomic and functional properties of a laboratory scale partial nitrification-anammox bioreactor were analysed. This type of bioreactor removes nitrogen from wastewater almost completely autotrophically, leaving much of the organic carbon present in influent wastewater available for anaerobic digestion and biogas production. The analysis showed that the taxonomic composition of the active microbial community is diverse and contains many heterotrophic bacterial groups responsible for other pathways of the microbial nitrogen cycle. Still, enzymes included in nitrification and anammox were predicted to be strongly expressed, pointing to the stability of the partial nitrification-anammox process even with many fast-growing heterotrophic species present. In the future, similar analyses should be done on industrial scale bioreactors operating under variable conditions.

(55 pages, 15 figures, 1 table, 105 references, original in Croatian)

Supervisor: Doc. dr. sc. Rosa Karlić

Reviewers: Doc. dr. sc. Rosa Karlić

Prof. dr. sc. Biljana Balen

Doc. dr. sc. Nenad Malenica

Substitute: Prof. dr. sc. Kristian Vlahoviček

## *Sadržaj*

<b>1. UVOD.....</b>	<b>1</b>
1.1. Eutrofikacija .....	1
1.2. Obrada otpadnih voda.....	3
1.3. Uklanjanje dušika iz otpadnih voda.....	5
1.3.1. Nitrifikacija i denitrifikacija.....	6
1.3.2. Anaerobna oksidacija amonijaka .....	7
1.4. Metagenomika .....	9
1.5. Metagenom bioreaktora s djelomičnom nitritacijom i anaerobnom oksidacijom amonijaka.....	12
1.6. Ciljevi rada .....	13
<b>2. MATERIJALI I METODE.....</b>	<b>14</b>
2.1. Izvor podataka .....	14
2.2. Kontrola kvalitete .....	14
2.3. Filtriranje kontaminacije humanim sljedovima.....	15
2.4. Sastavljanje metagenoma .....	15
2.5. Poravnanje na proteinsku bazu podataka.....	16
2.6. Taksonomska klasifikacija .....	17
2.7. Predviđanje gena koji kodiraju proteine.....	18
2.8. Funkcionalna anotacija predviđenih gena koji kodiraju proteine.....	18
2.9. Analiza nejednake uporabe kodona.....	19
<b>3. REZULTATI.....</b>	<b>21</b>
3.1. Kontrola kvalitete i filtriranje kontaminacije .....	21
3.2. Sastavljanje metagenoma .....	24
3.3. Taksonomski sastav bioreaktora.....	25
3.4. Predviđanje gena koji kodiraju proteine i funkcionalna anotacija pomoću mapiranja na bazu ortologije KEGG .....	28
<b>4. RASPRAVA.....</b>	<b>35</b>
4.1. Kvaliteta očitanih sljedova i sastavljenoga metagenoma .....	35
4.2. Taksonomski sastav bioreaktora.....	36
4.3. Nejednaka uporaba kodona i funkcionalna analiza metagenoma.....	38



<b>5. ZAKLJUČAK .....</b>	<b>41</b>
<b>6. LITERATURA.....</b>	<b>42</b>
<b>7. PRILOZI.....</b>	<b>i</b>
PRILOG 1: Tablica KEGG identifikacijskih brojeva ribosomskih proteina.....	i
PRILOG 2: Relativna ekspresija gena.....	iv
PRILOG 3: Datoteka za unos u mrežnu aplikaciju IPATH3.....	v
<b>8. ŽIVOTOPIS .....</b>	<b>xxxv</b>

# 1. UVOD

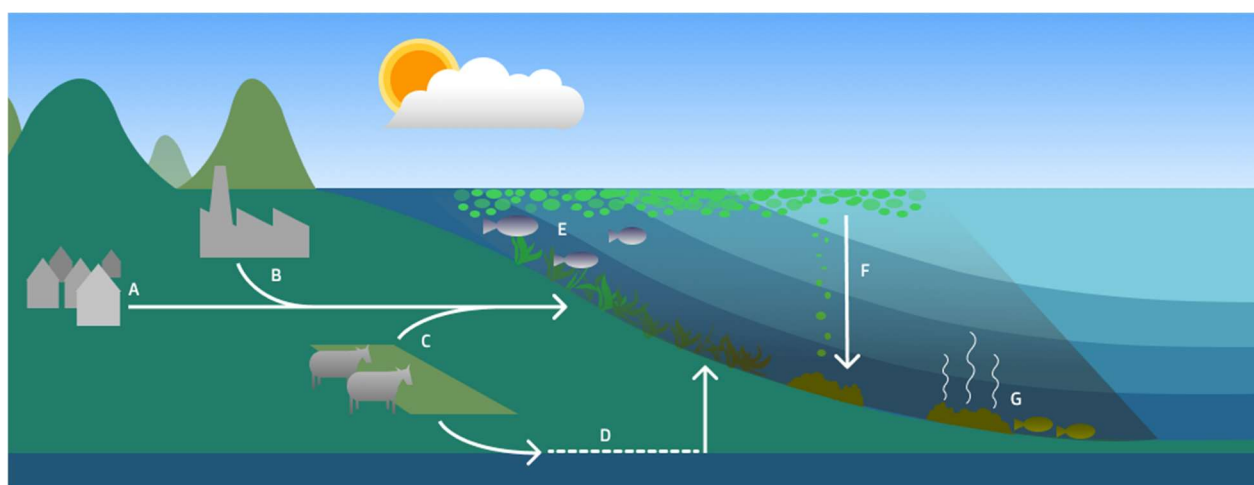
## 1.1. Eutrofikacija

Ispuštanje neprerađenih industrijskih i komunalnih otpadnih voda i otjecanje kišnice s poljoprivrednih površina zasićenih gnojivom u podzemne vode može značajno promijeniti kemijska svojstva prirodnih vodotoka. Naglo povećanje koncentracije organskih i anorganskih nutrijenata u vodenim ekosustavima, a posebice fosforovih i dušikovih spojeva, pogoduje bujanju algi i vodenih biljaka koje ograničavaju količinu prostora i svjetlosti dostupnih ostalim organizmima ekosustava. Nagli porast biomase algi, kolokvijalno nazivan cvjetanjem mora, koncentrira visoku razinu organske tvari u površinskim slojevima vode gdje dopire najviše sunčeve svjetlosti, a ta organska tvar s vremenom sedimentira na dno gdje je razgrađuju bentičke bakterije i drugi organizmi koji usput troše kisik i otpuštaju ugljikov dioksid.

Upravo metabolizam razgrađivača algi i vodenih biljaka uzrokuje nastanak hipoksičnih ili anoksičnih „mrtvih zona“ gdje se mogu dogoditi masovna odumiranja sesilnih i sporo krećućih organizama koji ne mogu dovoljno brzo migrirati iz zone hipoksije, a ta odumiranja služe kao izvor dodatnog organskog materijala za razgrađivače i ubrzavaju hipoksiju (Carpenter 2005; Lopez-Ridaura i sur. 2009). Opisani proces zove se eutrofikacija (Slika 1). Eutrofikacija je prirodna periodična pojava, no pogonjena ljudskim utjecajem (osim ispuštanja organskih tvari u prirodne vodotoke, i antropogenično globalno zatopljenje faktor je u eutrofikaciji zbog smanjenja topljivosti kisika s porastom temperature) događa se na mnogo većim vremenskim i prostornim rasponima. Možda najpoznatija hipoksična zona nalazi se u Baltičkom moru gdje se koncentracija kisika u vodenom stupu prati od početka 20. stoljeća. Površina baltičke zone hipoksije od početka mjerenja porasla je s oko 5000 km<sup>2</sup> na preko 60000 km<sup>2</sup> u nešto više od 100 godina, za što se uzrokom smatraju zatopljenje mora, pretjeran izlov ribe i ispuštanje otpadnih voda (Carstensen i sur. 2014).

Osim degradacije ekosustava, eutrofikacija može uzrokovati i kontaminaciju pitke vode fitotoksinima koje proizvode razne alge, kao i bioakumulaciju fitotoksina u organizmima koje čovjek koristi za hranu, posebno školjkašima (Heisler i sur. 2008). Degradacija ekosustava

uključuje i smanjenu brojnost i raznolikost kralježnjaka, od kojih su za ljude važne neke vrste riba koje se koriste za komercijalni izlov (Köster i sur. 2001). Dakle, eutrofikacija dovodi ne samo do degradacije vodenih ekosustava, već i do neuporabljivosti voda za piće i navodnjavanje te do manjka ili toksičnosti morske hrane. Uz ostale mjere smanjenja danas sve očitijeg čovjekova negativnog utjecaja na stabilnost globalnih ekosustava (smanjenje ugljičnog otiska, održiv izlov ribe i uzgoj ostale hrane), nužna je i stroga kontrola sastava otpadnih voda, posebno što se tiče koncentracije dušikovih i fosforovih spojeva (Vahtera i sur. 2007).



*Slika 1 Shematski prikaz procesa eutrofikacije. Ako se otpadne vode iz kućanstava (A), industrije (B), poljoprivrede i stočarstva (C) neobrađene ispuštaju u okoliš ili ispiranjem kišom dopijevaju u podzemne vode (D), značajno povećavaju razinu organskih tvari i anorganskih nutrijenata poput amonijaka i fosfata u vodenim ekosustavima. Povećava se biomasa vodenih biljaka i algi (E) koje ograničavaju prostor i svjetlost dostupne ostalim organizmima ekosustava. Kada alge i biljke počnu odumirati i tonuti na dno (F), koriste kao izvor hrane razgrađivačima koji troše kisik, što može rezultirati hipoksijom vode i umiranjem životinja (G).*

## 1.2. Obrada otpadnih voda

Iz navedenih razloga u mnogim državama postoje odredbe, ili barem preporuke, oko obrade otpadnih voda kao nusprodukata industrije, agrikulture ili kućanstava prije ponovne uporabe ili ispuštanja u okoliš (npr. Direktiva 91/271/EEC Europskog vijeća 1991; Smjernica EPA-821-R-21-003 Agencije za zaštitu okoliša SAD-a 2021). Otpadne vode stoga u posebnim postrojenjima prolaze uglavnom kroz tri, a ponekad i četiri stadija obrade (Gerba i Pepper 2009):

0. **Preliminarna obrada** uključuje prolazak otpadne vode kroz široke rešetke u kojima se zadržava krupni kruti materijal, npr. kućansko smeće, krupni komadi hrane, plastike, drveta, papira, itd.
1. **Primarna obrada** sastoji se od nekoliko sati sedimentacije otpadnih voda u velikim bazenima, tijekom kojih se materijal srednje veličine sliježe na dno ili nakuplja na površini vode ovisno o gustoći. Takav materijal otklanja se mehanički.
2. **Sekundarna obrada** uglavnom se odvija u bioreaktorima (Slika 2), umjetnim spremnicima u kojima masa živih mikroorganizama ili „aktivni mulj“ (eng. *activated sludge*) obavlja neku ciljanu biokemijsku pretvorbu ulaznog materijala; u ovom slučaju otpadnih voda. Važno je da aktivni mikroorganizmi flokuliraju, tj. da se nakupljaju u krupnije čestice kako bi mogli sedimentirati i tako se razdvojiti od pročišćene vode. Postoji mnogo vrsta bioreaktora koji se razlikuju po sastavu zajednice mikroorganizama i po mehaničkom dizajnu. Sekundarna obrada često uključuje prolazak vode kroz više različitih bioreaktora; npr. za neke procese je potrebna aktivnost i aerobnih i anaerobnih organizama kojima su potrebni različiti uvjeti.
3. **Tercijarna obrada** sastoji se od dodatnog pročišćavanja vode koristeći fizikalno-kemijske metode poput adsorpcije na aktivnom ugljenu i reverzne osmoze. Provodi se najčešće kada je voda namijenjena za piće, navodnjavanje poljoprivrednih površina ili ispuštanje u okoliš u kojemu ljudi češće provode rekreativno vrijeme (npr. kupališta, tereni za golf), namjene za koje je visoka čistoća vode važna za održavanje javnog zdravlja.

Tijekom sekundarne obrade otpadnih voda organska tvar uklonjena iz otpadne vode akumulira se kao biomasa mikroorganizama aktivnoga mulja (sekundarni mulj), i mora se konstantno ili

periodično odvoditi i obrađivati. Dio nastale biomase može se koristiti za inicijaciju novih bioreaktora, no većina se dodatno obrađuje za druge svrhe.

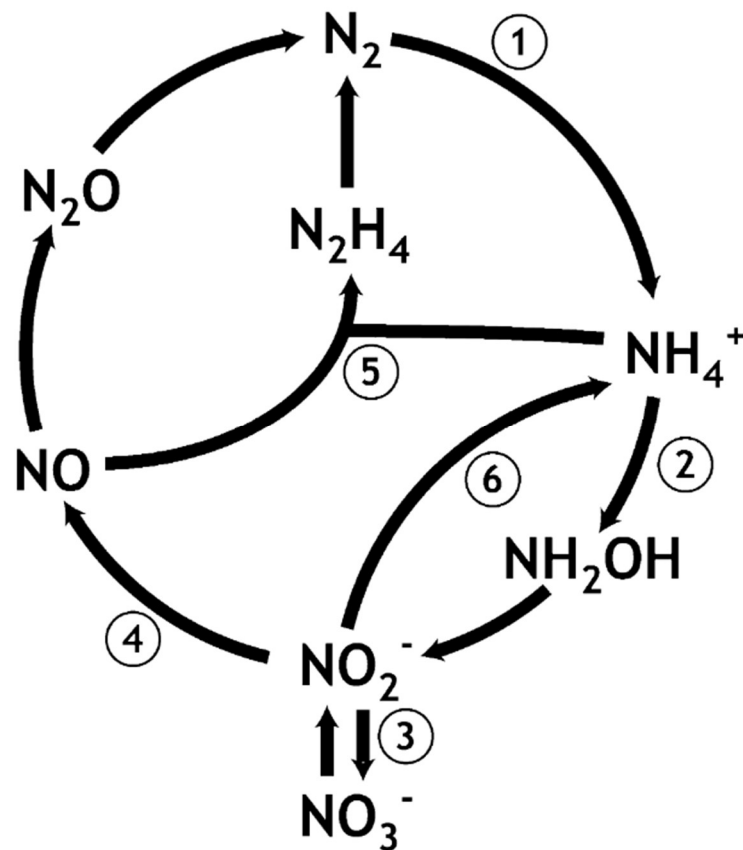
Tako nakupljeni sekundarni mulj može se dehidrirati i koristiti kao kompost, anaerobno digestirati s bioplinom (uglavnom metanom) kao produktom ili dehidrirati i koristiti direktno kao izvor energije putem spaljivanja. U dobro dizajniranim postrojenjima za preradu sekundarnog mulja moguće je proizvesti više energije no što strojevi za preradu troše, što ih čini energetski neutralnim ili čak pozitivnim. Dio hongkonške električne mreže tako je pogonjen lokalnim postrojenjem za preradu aktivnog mulja. Sekundarni mulj tamo se spaljuje, a proizvedena energija dovoljna je i za svakodnevni rad postrojenja i za struju oko 4000 kućanstava (Hong Kong EPD 2005).



*Slika 2 Fotografija bioreaktora Zagebačkih otpadnih voda (ZOV). Prikazan je bioreaktor curećeg tipa (eng. Trickle-down) u kojemu se otpadna voda prska na podlogu velike površine na kojoj raste biofilm aktivnih mikroorganizama. Preuzeto s mrežne stranice ZOV-a (<https://www.zov-zagreb.hr/hr/otnama/nasa-postrojenja.html>).*

### 1.3. Uklanjanje dušika iz otpadnih voda

Bakterije i drugi mikroorganizmi obavljaju važnu ulogu u ciklusu dušika u prirodi (Slika 3), a analogne uloge igraju i u sklopu bioreaktora za uklanjanje dušika iz otpadnih voda. Dušik se u otpadnim vodama nalazi uglavnom u obliku amonijaka i amonijevih iona koji su krajnji produkt razgradnje mnogih bioloških makromolekula koje sadrže amino skupine.



Slika 3 Poznati putevi mikrobnoga ciklusa dušika uključuju fiksaciju dušika (1), aerobnu oksidaciju amonijaka (2), aerobnu oksidaciju nitrita (3), denitrifikaciju (4), anaerobnu oksidaciju amonijaka (5), i disimilatornu redukciju nitrita i nitrata u amonijak (6). Preuzeto iz Jetten 2008.

### 1.3.1. Nitrifikacija i denitrifikacija

Kod klasičnog tipa biološkog uklanjanja dušika iz otpadnih voda koji je i danas u najširoj uporabi, dušik u obliku amonijaka je u aerobnim uvjetima prvo nitrificiran (što uključuje nitritaciju tj. oksidaciju amonijaka u nitrit te nitrataciju tj. oksidaciju nitrita u nitrat), a zatim denitrificiran (reduciran u dušikove okside i molekularni dušik) u anoksičnim uvjetima. Procese nitritacije i nitratacije provode dvije odvojene skupine bakterija: bakterije koje oksidiraju amonijak (eng. *Ammonium oxidizing bacteria*, AOB) i bakterije koje oksidiraju nitrit (eng. *Nitrite oxidizing bacteria*, NOB). *Nitrosomonas*, *Nitrosococcus*, *Nitrosospira*, *Nitrosovibrio* i *Nitrosolobus* poznati su rodovi AOB (Samoča i Prangnell 2019), a *Nitrobacter*, *Nitrococcus*, *Nitrospira* i *Nitrospina* neki su od poznatih rodova NOB. Denitrifikaciju provode bakterije razolikih rodova. Te bakterije uglavnom su fakultativni anaerobi koji u odsustvu kisika koriste nitrat kao akceptor elektrona kod staničnog disanja (Rajta i sur. 2020).

Nitrifikacijskim bakterijama potreban je kisik, pa je za prozračivanje većih bioreaktora u kojima se odvija nitrifikacija potrebno puno energije za pogon pumpi za kisik. Heterotrofnim denitrifikacijskim organizmima, s druge strane, potreban je izvor energije u obliku organske tvari (uglavnom se dodaje metanol) koju sagorijevaju, otpuštajući ugljikov dioksid (Jorgensen i Halling-Sorensen 1993). Osim toga, ti procesi rezultiraju brзом akumulacijom biomase aktivnog mulja koji je potrebno redovito uklanjati i prerađivati. Ovi faktori čine nitrifikacijsko-denitrifikacijske bioreaktore relativno skupima za održavanje i izvorom dodatnih stakleničkih plinova.

### 1.3.2. Anaerobna oksidacija amonijaka

Do tada nepoznati organizmi koji bi oksidirali amonijak u molekularni dušik u anaerobnim uvjetima teorijski su predviđeni 1977. godine evolucijskim analogijama i termodinamičkim izračunom povoljnosti takve reakcije (Broda 1977). Tek je 1994. godine, gotovo 20 godina nakon teorijske predikcije, proces anaerobne oksidacije amonijaka (eng. *Anaerobic ammonium oxidation, anammox*) empirijski dokazan u bioreaktoru za denitrifikaciju otpadnih voda (Mulder i sur. 1995), a nekoliko godina nakon otkriveni su organizmi odgovorni za taj proces - također pomoću uzgoja u bioreaktoru (Strous i sur. 1998, 1999).

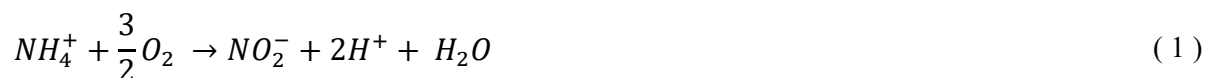
S vremenom udvostručivanja od preko 2 tjedna, *anammox* bakterije pokazale su se izazovnim za izolaciju i kultivaciju u uvjetima *in vitro*. Tek pročišćavanjem iz biofilma nastalog u bioreaktoru omogućena je izolacija i sekvenciranje njihova gena za 16S rRNA, nakon čega su klasificirane kao *Planctomycetes* s novootkrivenim redom unutar tog koljena nazvanim *Brocadiales* i prvom karakteriziranom vrstom nazvanom *Candidatus Brocadia anammoxidans* (Kuenen 2008).

U međuvremenu su *anammox* bakterije pronađene u raznolikim prirodnim vodenim staništima na granicama aerobne i anoksične zone. Ekspedicije u Crnom moru i u obalnim morima Kostarike procijenile su da je *anammox* proces zaslužan za preko 30% ukupnog obrata dušika u tim područjima (Dalsgaard i sur. 2003; Kuypers i sur. 2003). Otkriće ovog važnog ekološkog faktora bilo je moguće samo kultivacijom u bioreaktorima koji sadrže kompleksne zajednice mikroorganizama. Do danas je prepoznato 16 vrsta *anammox* bakterija svrstanih u pet rodova. Svi do sada otkriveni taksoni *anammox* bakterija nose titulu *Candidatus* zato što nisu uspješno kultivirani klasičnim metodama: *Ca. Anammoxoglobus*, *Ca. Brocadia*, *Ca. Jettenia*, *Ca. Kuenenia* i *Ca. Scalindua* (Sonthiphand, Hall, i Neufeld 2014).

Primjena *anammox* bakterija u uklanjanju dušika iz otpadnih voda vrlo je brzo prepoznata. Prvo istraživanje takve primjene objavljeno je 2001. godine i trajalo je tada već preko dvije godine, dokazujući stabilnost i održivost *anammox* bioreaktora (Van Dongen, Jetten, i Van Loosdrecht 2001). Proces se sastojao od SHARON (eng. *Single reactor system for high activity ammonium removal over nitrite*, sustav pojedinačnog bioreaktora visoke aktivnosti za uklanjanje amonijaka preko nitrita) bioreaktora čiji je efluent, s oko 53% amonijaka oksidiranog u nitrit, služio kao influent anaerobnom bioreaktoru s *anammox* bakterijama. Amonijak je bilo potrebno djelomično



nitritirati prije ulaza vode u *anammox* bioreaktor zato što *anammox* reakcija za oksidaciju amonijaka koristi nitrit kao akceptor elektrona, oksidirajući amonijak i reducirajući nitrit u molekularni dušik. Ispod su prikazane kemijske jednadžbe nitritacije (1), anaerobne oksidacije amonijaka (2) i zbroj tih dviju jednadžbi (3).



Od tada je razvijeno nekoliko različitih sustava za uklanjanje dušika koji koriste *anammox* reakciju (Kartal, Geerts, i Jetten 2011; Ni i sur. 2010; Schmidt i sur. 2003; Van Der Star i sur. 2008). Prvi upogonjeni bioreaktor industrijske veličine bio je upravo prvog opisanog tipa – kombinirani SHARON i *anammox* proces – u Rotterdamu (van der Star i sur. 2007). U revijalnom radu iz 2019. godine procijenjeno je da je preko 200 bioreaktora koji koriste djelomičnu nitritaciju i *anammox* (eng. *Partial nitrification and anammox*, PNA) tada bilo u pogonu; međutim, većinom su se koristili u sklopu sporednog toka otpadne vode u postrojenjima u kojima glavninu obradu još uvijek odrađuju konvencionalniji tipovi bioreaktora (Winkler i Straka 2019). Sporedni tok odnosi se na vodu iz različitih izvora iz samog postrojenja za obradu otpadne vode, npr. supernatant digestora sekundarnog aktivnog mulja, vodu preostalu nakon dehidracije mulja itd.

Prepreka efikasnom uklanjanju dušika iz glavnog toka otpadne vode je visok omjer masa organskog ugljika i dušika koji pogoduje preferencijalnom rastu heterotrofnih bakterija koje kompetiraju sa sporo rastućim *anammox* bakterijama (Cao, van Loosdrecht, i Daigger 2017). Uz to, glavni tok obično ima nižu koncentraciju amonijaka i nitrita od sporednog te je puno varijabilnije temperature u odnosu na sporedni, a aktivnost *anammox* bakterija opada dramatično s padom temperature (Lotti i sur. 2015). Održivost PNA procesa mora se, dakle, osloniti na poticanje rasta aerobnih nitritirajućih i *anammox* bakterija te na supresiju rasta heterotrofnih nitratirajućih bakterija, a posebno na hladnijim temperaturama. Također je potrebno selektivno zadržavanje *anammox* i nitritirajućih bakterija, a ispiranje neželjenih heterotrofnih bakterija (Agrawal i sur. 2018).

## 1.4. Metagenomika

Procjenjuje se da preko 99% vrsta mikroorganizama prisutnih u prirodi nije kultivirano, a da dio možda nije ni moguće kultivirati standardnim metodama *in vitro* (Hugenholtz, Goebel, i Pace 1998). Gotovo da nema područja Zemlje ili ekološke niše koja nije ispunjena raznolikim mikroorganizmima čiji je biodiverzitet nemoguće procijeniti klasičnom kultivacijom, a još je teže tako proučavati ekologiju tih mikrookoliša. Metagenomika, tj. izolacija, sekvenciranje i analiza genetskog materijala izoliranoga izravno iz okolišnih uzoraka zaobilazi te poteškoće, promatrajući ukupni genetski materijal mikrookoliša na neki način kao jedinstven genom.

Metagenomika je započela sa sekvenciranjem rRNA amplikona iz okolišnih uzoraka i njihovom filogenetskom analizom, a već rana istraživanja usredotočena na ekstremne okoliše s teško uzgojivim mikroorganizmima identificirala su amplikone brojnih do tada nepoznatih vrsta (Amann, Ludwig, i Schleifer 1995). Kao što je spomenuto, i *anammox* bakterije otkrivene su na ovaj način. U međuvremenu su razvoj metoda sekvenciranja druge i treće generacije, drastično smanjenje njihove cijene te razvoj računalnih metoda obrade ogromnih količina podataka omogućili sekvenciranje, sastavljanje i analizu gotovo cjelovitih metagenoma.

Suvremeni pristup sekvenciranju i analizi čitavog metagenoma analogan je pristupu sekvenciranja i analize genoma jednog organizma, uz neke bitne razlike. Cjelokupna DNA uzorka se izolira i konstruira se metagenomska DNA biblioteka koja se zatim sekvencira, u posljednje vrijeme najčešće pomoću neke od platformi proizvođača Illumina (Pérez-Cobas, Gomez-Valero, i Buchrieser 2020).

Sirovi podaci dobiveni sekvenciranjem prvo moraju proći kroz kontrolu kvalitete. Softveri koji pretvaraju grube izlazne podatke uređaja za sekvenciranje u slijed baza uključuju u svoje izlazne podatke, osim očitano slijeda, i podatak o sigurnosti u svaku pojedinu očitano bazu. Najrašireniji takav softver i po njemu nazvana mjera sigurnosti u pojedine baze slijeda je *phred* (Ewing i sur. 1998), gdje je mjera sigurnosti definirana kao  $Q$  u jednadžbi (4) gdje je  $P$  vjerojatnost pogreške kod očitavanja baze (Ewing i Green 1998). Phred vrijednost 20 koja odgovara 99% sigurnosti u očitano bazu često se uzima kao prag vjerodostojnosti slijeda.

$$Q = -10\log_{10} P \quad (4)$$

U datotekama koje sadržavaju nukleotidne sljedove formata FASTQ, *phred* mjera kvalitete sadržana je kod novijih podataka najčešće u obliku ASCII (od eng. *American Standard Code for Information Interchange*, američki standardni kod za razmjenu informacija) znakova kodiranih vrijednošću Q + 33 (Cock i sur. 2009). Pomoću softvera za kontrolu kvalitete sljedova sadržanih u FASTQ datotekama sljedovi se tako mogu ovisno o kvaliteti pojedinih nukleotida ili većih dijelova filtrirati (iz daljnje analize potpuno isključiti) ili rezati (odbaciti samo dijelove slijeda manje kvalitete ili dijelove koji odgovaraju tehničkim, a ne metagenomskim sljedovima).

Tako obrađeni sljedovi zatim se sastavljaju, tj. traženjem preklapanja među manjim sljedovima slažu se dulji sljedovi koji izvorno proizlaze iz iste vrste organizma. Što je veća pokrivenost (eng. *Coverage*) metagenoma tj. što su pojedine baze pokrivena u prosjeku s više očitanih sljedova (eng. *Reads*), moguće je sastaviti dulje neprekinute sljedove (eng. *Contiguous sequences*, *Contigs*). Ovdje nailazimo na jednu od bitnih razlika između genomike i metagenomike: pokrivenost metagenoma u pravilu je mnogo manja od pokrivenosti genoma, posebno kod kompleksnijih zajednica mikroorganizama gdje se može dogoditi da je pojedina vrsta predstavljena samo jednim očitanim slijedom (Kunin i sur. 2008). To znači da će se moguće pogreške kod očitavanja baza u tim sljedovima propagirati u nizvodne analize zbog nedostatka uprosječenja s drugim sljedovima koje pokrivaju istu regiju, a i da će sastavljanje metagenoma dovesti do prosječno kraćih neprekinutih sljedova. Metagenomi kompleksnih zajednica zato gotovo nikad nisu cjeloviti i nemaju preciznost usporedivu s genomima.

Druga važna razlika između genomike i metagenomike je korak klasifikacije (eng. *Binning*) složenih sljedova kod metagenomike. Složenim sljedovima metagenoma pridružuju se pripadajuće vrste ili više taksonomske razine poravnanjem s bazama anotiranih sljedova i tako se procjenjuje taksonomski sastav zajednice i udjeli pojedinih vrsta u njoj. Geni koji kodiraju proteine mogu se predvidjeti *ab initio*, iz neanotiranih složenih sljedova, analizom otvorenih okvira čitanja (eng. *Open reading frames*, ORF) u kojima se traže neke karakteristike poznatih gena, npr. blizina poznatih promotorskih sljedova, mjesta vezanja transkripcijskih faktora, određena duljina od *start* do *stop* kodona itd. (Mir i sur. 2012). Funkcija gena pronađenih u metagenomu može se predvidjeti klasifikacijom gena u neku od ortolognih skupina gena, tj. skupina gena pretpostavljene iste ili slične funkcije zbog evolucijske konzerviranosti (Fang i sur. 2010).

Sveukupno, metagenomski pristup dovodi do karakterizacije mikrookoliša na razini taksonomije, ekologije, funkcionalnog potencijala, a kod metagenoma s visokom pokrivenošću i strukture sastavnih genoma. U kontekstu bioreaktora, te informacije mogu dovesti do razumijevanja mikrobnih interakcija i metaboličkih mreža koje čine njihovu funkcionalnost i potencijalno pružiti temelj za optimizaciju uvjeta za efikasnije djelovanje i ekonomičnije održavanje.

## 1.5. Metagenom bioreaktora s djelomičnom nitritacijom i anaerobnom oksidacijom amonijaka

Prethodna istraživanja metagenoma aktivnog mulja bioreaktora koji koriste *anammox* reakciju pokazala su da su vrste ključne za djelovanje PNA sustava uz *anammox* bakterije rodova *Brocadia*, *Kuenenia* i *Jettenia* (Li i sur. 2009; Ma i sur. 2017; Park i sur. 2010; Vlaeminck i sur. 2010) nitritirajuće bakterije poput nekih *Proteobacteria*, posebno roda *Nitrosomonas* (Chu i sur. 2015; Gonzalez-Martinez i sur. 2015; Li i sur. 2009; Vlaeminck i sur. 2010). Uz njih, u bioreaktorima su često identificirane i heterotrofne bakterije kao što su *Chloroflexi*, *Chlorobi* i *Bacteroidetes* (Chu i sur. 2015) za koje se smatra da degradiraju ostatke umrlih *anammox* i drugih bakterija u bioreaktoru, sprječavajući nakupljanje tih ostataka (Kindaichi i sur. 2012).

Većina dosadašnjih istraživanja bila je usredotočena na taksonomsku analizu PNA bioreaktora, s rjeđim funkcionalnim analizama (Speth i sur. 2016; Wang i sur. 2020). Funkcionalna analiza metagenoma omogućuje identifikaciju gena i metaboličkih puteva potencijalno važnih za suživot i djelovanje mikrobioma. Uz metagenomiku, metatranskriptomika i metaproteomika su vrijedni alati za karakterizaciju aktivnih gena mikrobioma i za povezivanje gena s njihovom ekološkom funkcijom putem analize njihove ekspresije u različitim uvjetima. Te dvije discipline, međutim, imaju ograničenja, među njima i tehničku zahtjevnost i ponekad visoku cijenu izolacije mRNA i proteina.

Grublje informacije o potencijalnoj razini ekspresije gena u populaciji mikroorganizama koji nastanjuju neku ekološku nišu moguće je dobiti i iz njihovog genoma. Poznato je da se sinonimni kodoni aminiokiselina ne koriste jednoliko, već su u manjoj ili većoj mjeri optimizirani pod pritiskom raznih sila prirodne selekcije, npr. translacijskom efikasnošću (Sharp i Li 1987). Kako translacijska efikasnost najveći utjecaj ima na više eksprimirane gene, moguće je iz samoga slijeda genoma predvidjeti snažniju ekspresivnost tj. potencijal brže translacije gena čija je uporaba kodona više optimizirana. Kao referentna točka za procjenu optimizacije uporabe kodona obično se uzimaju poznato visoko eksprimirani ribosomski geni. Pokazano je i da genomi različitih vrsta istoga mikrobioma mogu pokazati pristranost prema istim kodonima pod utjecajem uvjeta u zajedničkoj ekološkoj niši (Roller i sur. 2013). Zahvaljujući tome, moguće je

analizom samog metagenoma dobiti procjenu relativne ekspresivnosti gena i metaboličkih puteva važnih za život mikrobioma.

## 1.6. Ciljevi rada

U ovom radu analizirala sam PNA bioreaktor laboratorijske veličine (Mardanov i sur. 2019). Cilj analize je proučiti strukturu zajednice mikroorganizama zaslužnih za djelovanje bioreaktora na razini taksonomske raznolikosti i funkcionalnih svojstava. U tu svrhu sekvencirani metagenom prošao je kroz kontrolu kvalitete, zatim sam ga sastavila, a složene sljedove taksonomski klasificirala putem poravnanja s proteinskom bazom podataka. Gene koji kodiraju proteine predvidjela sam analizom ORF-ova. Predviđenim genima pridružila sam vjerojatnu funkciju pomoću klasifikacije u bazu ortolognih skupina gena KEGG. Također sam procijenila i razinu ekspresivnosti gena analizom nejednake uporabe kodona kako bih identificirala metaboličke puteve ključne za djelovanje bioreaktora.

## 2. MATERIJALI I METODE

### 2.1. Izvor podataka

Bioreaktor čiji metagenom sam analizirala u ovom radu je PNA bioreaktor volumena 9 litara. Bioreaktor je sekvencijalno serijskog tipa (eng. *Sequencing batch reactor*, SBR) što znači da mu je radni ciklus podijeljen na faze: od ukupno 6 sati trajanja ciklusa, 30 minuta odvojeno je za sedimentaciju, 30 minuta za unos umjetne otpadne vode te 5 sati i 10 minuta za miješanje otpadne vode i aktivnog mulja (Mardanov i sur. 2019). Pročišćena voda odvodi se istovremeno s unosom otpadne vode. Bioreaktor je inokuliran aktivnim muljem iz prethodno opisanoga PNA bioreaktora (Mardanov i sur. 2016).

Izolacija DNA, priprema DNA biblioteke za sekvenciranje s obje strane fragmenata (eng. *Paired-end sequencing*) te samo sekvenciranje opisani su u radu Mardanova i sur. iz 2016. Nukleotidni slijed metagenoma pristupnog koda SRR9831403 dohvatila sam s baze podataka SRA (NCBI/SRA 2021).

### 2.2. Kontrola kvalitete

Početnu kvalitetu očitanih sljedova provjerila sam pomoću softverskog alata *FastQC* (Anders 2010). Izrezivanje i filtriranje sljedova manje kvalitete napravila sam pomoću programa *Trimmomatic* (Bolger, Lohse, i Usadel 2014). Izrezivanje tehničkih sljedova (Illumina adaptera Truseq2, Illumina 2015) obavljeno je poravnavanjem očitanih sljedova metagenoma sa sljedovima Illumina adaptera, dopuštajući pogrešno poravnanje dvije baze slijeda metagenoma sa sljedovima Illumina adaptera. Očitani sljedovi s *phred* +33 ocjenom kvalitete manjom od 20 (što odgovara 99% sigurnosti u očitane baze) izrezani su koristeći klizeći prozor veličine 20, tj. uzeti su prosjeci kvalitete za svakih 20 uzastopnih baza počevši od početka slijeda, a sljed je izrezan na poziciji pada izračunatoga prosjeka ispod 20. Očitane sljedove kojima je nakon izrezivanja duljina bila manja od 50 sam odbacila. Sljedove koji su nakon obrade ostali bez svoga para

sekvenciranog sa suprotne strane izdvojila sam i nisu korišteni u daljnjim analizama pošto većina kasnije korištenih alata može kao ulazne podatke uzeti ili sparene ili nesporene sljedove, ali ne i mješavinu sparenih i nesparenih. Kvalitetu tako obrađenih sljedova ponovno sam provjerila i vizualizirala pomoću programa *FastQC*.

### 2.3. Filtriranje kontaminacije humanim sljedovima

Potencijalnu kontaminaciju metagenoma humanim sljedovima uklonila sam pomoću programa *Bowtie2* (Langmead i Salzberg 2012). Metagenom sam poravnala s 38. verzijom sastavljenog humanog genoma Konzorcija za referentne genome (eng. *Homo sapiens genome assembly hg38 of the Genome Reference Consortium*, GRCh38, NCBI 2019), ne dopuštajući poravnanje samo jednoga među sparenim očitanim sljedovima (parametar *-no-mixed*). Sljedove koji su uspješno poravnati s humanim genomom sam odbacila.

### 2.4. Sastavljanje metagenoma

Očitane sljedove metagenoma sastavila sam u prekinute sljedove (eng. *Scaffolds*) pomoću programa *MetaSPAdes* koristeći osnovne postavke i parametre (Nurk i sur. 2017). Sastavljanje prekinutih sljedova metagenoma uključuje slaganje očitanih sljedova u neprekinute sljedove (eng. *Contigs*) traženjem preklapanja među očitanim sljedovima i slaganjem sljedova u neprekinuti niz nukleotida, te slaganje prekinutih sljedova sastavljenih od neprekinutih sljedova poznatoga poretka i orijentacije razmaknutih razmacima nepoznatoga slijeda i duljine. Algoritam kojim program *MetaSPAdes* sastavlja neprekinute sljedove temelji se na konstrukciji de Bruijnovih grafova u kojima su svi mogući podnizovi očitanih sljedova duljine  $k - 1$  ( $k-1$ -meri) predstavljeni vrhovima grafa, a podnizovi duljine  $k$  ( $k$ -meri) predstavljeni su usmjerenim bridovima koji spajaju dva  $k-1$ -mera tog  $k$ -mera. Algoritam traži Eulerov put kroz graf, tj. put koji prolazi svakim bridom točno jednom. *MetaSPAdes* konstruira de Bruijnovu grafove s  $k$  vrijednostima 21, 33 i 55, a za sastavljanje neprekinutih sljedova kombinira rezultate triju grafova.



U programskom jeziku R izračunala sam broj, ukupnu duljinu, medijan i prosjek duljina prekinutih sljedova, te vrijednost N50. N50 je česta statistika za procjenu kvalitete sastavljenoga genoma ili metagenoma, a označava minimalnu duljinu prekinutoga slijeda takvu da je 50% ukupne duljine sastavljenoga metagenoma pokriveno prekinutim sljedovima te ili veće duljine.

Radi procjene količine očitanih sljedova koji su uspješno sastavljeni u prekinute sljedove metagenoma, očitane sljedove poravnala sam sa sastavljenim prekinutim sljedovima pomoću programa *Bowtie2*. Korištene su osnovne postavke i parametri.

## 2.5. Poravnanje na proteinsku bazu podataka

Sastavljeni metagenom poravnala sam na neredundantnu proteinsku bazu podataka NCBI-nr (NCBI 2019b) pomoću modula *blastx* programa *DIAMOND* (Buchfink, Reuter, i Drost 2021). Program *DIAMOND* služi poravnanju sljedova s bazama podataka - kod modula *blastx* s proteinskim bazama podataka. Program koristi algoritam srodan algoritmu *BLAST* pretrage (Altschul i sur. 1990), tj. algoritam osnovan na principu traženja lokalnih poravnanja visoke sličnosti (eng. *High-scoring pair*) s preinakama koje ga čine bržim i prikladnijim za brojne skupove sljedova kao što su skupovi metagenomskih prekinutih sljedova.

Korišteni su sljedeći parametri:

*-F 15*: Zadana vrijednost kazne za pogrešno poravnanje koje pomiče okvir čitanja (eng. *Frameshift penalty*) bila je 15, prema preporukama autora programa kod poravnavanja nukleotidnih sljedova s proteinskim bazama podataka.

*--range-culling*: Kod kreiranja izlazne datoteke koja sadrži poredak najboljih „pogodaka“ u proteinskoj bazi podataka za svaki prekinuti slijed metagenoma, automatski je zadano da program prijavljuje samo 25 proteina s najboljom ocjenom poravnanja za svaki prekinuti slijed. Parametar *--range-culling* mijenja tu postavku tako da se slabije ocijenjena poravnanja uklanjaju iz izlazne datoteke samo ako 25 jednako ili bolje ocijenjenih poravnanja pokriva barem 50% njihovoga raspona u slijedu. Drugim riječima, ako je slijed nekog proteina iz baze podataka poravnat s dijelom slijeda metagenoma u kojemu je mali broj bolje ocijenjenih pogodaka, bit će zadržan u

izlaznoj datoteci. U suprotnom, dio slijeda metagenoma koji može biti odlično poravnat s mnogim proteinima iz baze podataka (npr. evolucijski vrlo konzervirani sljedovi) može zauzeti svih 25 mjesta izlazne datoteke kod automatski zadanih parametara.

--top 10: Poravnanja su uklonjena iz izlazne datoteke samo ako imaju najmanje 10% lošiju ocjenu od drugih poravnanja kojima se raspon preklapa s 50% raspona tog proteina.

-f 100: Format izlazne datoteke bio je *DIAMOND alignment archive* (.daa).

## 2.6. Taksonomska klasifikacija

Izlaznu datoteku programa *DIAMOND* koja sadrži informacije o poravnanju prekinutih sljedova s proteinskom bazom podataka NCBI-nr unijela sam u program *Megan* (Huson i sur. 2007) koji svakom od poravnanja pridružuje taksonomsku klasifikaciju organizma u kojemu je protein iz baze podataka identificiran. Za taksonomsku klasifikaciju čitavih prekinutih sljedova odabrala sam *Megan* algoritam posljednjeg zajedničkog pretka prilagođen za duge sljedove (eng. *Interval-Union Last Common Ancestor*, IU-LCA). Algoritam IU-LCA definira značajno poravnanje kao ono kojemu je *bitscore* do 10% lošiji od najbolje ocijenjenog poravnanja koji pokriva dio njegovoga intervala ili čitav njegov interval, uz uvjete da je *bitscore* 50 ili više i *evaluate* 0,01 ili manje. Zatim svakoj taksonomskoj skupini identificiranoj među značajnim poravnanjima pridružuje uniju intervala odgovarajućih poravnanja, a skupinama viših taksonomskih razina pridružuje uniju intervala njihovih potomaka. Prekinuti slijed je zatim anotiran kao taksonomska skupina najniže razine čija je unija intervala duljine barem 80% ukupne duljine poravnatih dijelova prekinutih sljedova.

Tako dobivene informacije o taksonomskoj klasifikaciji vizualizirala sam na lokalnom računalu u programu *Megan*. Kako bi kvantifikacija taksonomske klasifikacije bila neovisna o varijaciji u duljinama skela, *Megan* kvantificira klasifikaciju prema broju poravnatih baza, a ne prema broju prekinutih sljedova poravnatih na svaku taksonomsku kategoriju. U suprotnome bi taksonomija duljih prekinutih sljedova, sastavljenih od mnoštva očitanih sljedova, bila podzastupljena. Informacije o taksonomskoj klasifikaciji transformirala sam u tablični format pomoću

*DIAMOND* modula *-view* radi unosa u programski paket *R* (verzija 3.5.2) gdje sam klasifikaciju vizualizirala.

## 2.7. Predviđanje gena koji kodiraju proteine

Program *Metagenemark* (Zhu, Lomsadze, i Borodovsky 2010) koristila sam za predviđanje ORF-ova koji kodiraju proteine. Koristila sam parametre *MetaGeneMark\_v1*, a zadani format rezultata predviđanja bio je generalni format za značajke (eng. *General feature format*, GFF). Tako sam dobila koordinate predviđenih gena. U programskom paketu *R* pomoću paketa *biostrings* (Pagés i sur. 2020) i *GenomicRanges* (Lawrence i sur. 2013) iz originalne datoteke s prekinutim sljedovima pomoću dobivenih koordinata gena izvukla sam sljedove predviđenih gena.

## 2.8. Funkcionalna anotacija predviđenih gena koji kodiraju proteine

Predviđene gene koji kodiraju proteine mapirala sam na bazu ortolognih skupina proteina KEGG pomoću programa *DIAMOND* i njegovog modula *blastx*. S obzirom da su u bazi nedostajala dva enzima vezana uz *anammox* reakcije, hidrazin sintaza i hidrazin dehidrogenaza, sljedove tih proteina dohvatila sam s mrežne stranice *Protein data bank* (<https://www.rcsb.org/>) s pristupnim kôdovima 5C2V i 6HIF. Sljedovima proteina u *fasta* formatu ručno sam pridodala njihove KEGG kôdove: K20932 za podjedinicu  $\alpha$  hidrazin sintaze; K20933 za podjedinicu  $\beta$  hidrazin sintaze; K20934 za podjedinicu  $\gamma$  hidrazin sintaze; K20935 za hidrazin dehidrogenazu. Od sljedova proteina napravila sam referentnu baza podataka pomoću *DIAMOND* modula *madeb*, a zatim sam na tu bazu ponovno poravnala predviđene gene metagenoma pomoću *DIAMOND* modula *blastx* i rezultate sam spojila s prvim poravnanjem na KEGG pomoću programskog jezika *R* jednostavnim spajanjem tablica koje sadrže rezultate ta dva mapiranja. Daljnje korake radila sam u programskom jeziku *R*.

Najbolje poravnanje za svaki predviđeni gen odredila sam dvama parametrima: isključila sam sva poravnanja s *E* vrijednosti većom od  $10^{-4}$ , a zatim sam svakom predviđenom genu pridružila poravnanje s najvećom vrijednosti *bitscore*. Iz daljnje analize izbacila sam poravnanja bez KEGG identifikacijskog broja.

## 2.9. Analiza nejednake uporabe kodona

Pomoću paketa *Biostrings* programskog jezika R izračunala sam udio GC baza u predviđenim genima. Uporabu kodona prvo sam kvantificirala pomoću programskog jezika R računanjem relativne uporabe sinonimnih kodona (eng. *Relative Synonymous Codon Usage*, RSCU) za svaki kodon, što je mjera definirana kao omjer opservirane i očekivane učestalosti uporabe kodona.

RSCU sam prvo izračunala za cijeli skup predviđenih gena definirajući očekivanu učestalost uporabe kodona kao jednaku uporabu sinonimnih kodona za svaku aminokiselinu. Zatim sam izračunala RSCU za ribosomske gene koje sam izdvojila iz skupa predviđenih gena pomoću njihovih KEGG identifikacijskih brojeva. Popis KEGG identifikacijskih brojeva ribosomskih proteina nalazi se u Prilogu 1 i preuzela sam ga iz KEGG baze kōdova ribosomskih proteina (<https://www.genome.jp/kegg/annotation/br01610.html>). Očekivanu učestalost uporabe kodona za ribosomske gene definirala sam kao prosječnu uporabu sinonimnih kodona za svaku aminokiselinu cijeloga skupa predviđenih gena. Dobivene vrijednosti unijela sam u softver za proračunske tablice *Excel* (Microsoft Corporation bez dat.), gdje sam napravila tablicu povezanosti kodona i aminokiselina koje kodiraju, s kodonima grupiranim po prvoj i drugoj nukleotidnoj bazi. Svakom kodonu pridružila sam izračunatu učestalost uporabe i RSCU vrijednost.

Za analizu nejednake uporabe kodona koristila sam paket *coRdon* programskog jezika R (Elek, Kuzman, i Vlahoviček 2019). Pomoću tog paketa prvo sam izračunala frekvencije uporabe kodona za svaki anotirani gen, a zatim i mjeru udaljenosti uporabe kodona neovisnu o duljini gena (eng. *Measure Independent of Length and Composition*, MILC, Supek i Vlahoviček 2005) i mjeru razine ekspresivnosti osnovanu na MILC vrijednosti (eng. *MILC-based Expression Level Predictor*, MELP) s genima koji su anotirani kao ribosomske podjedinice korištenim kao referentnim genima visoke optimizacije uporabe kodona. Grafove procijenjene relativne ekspresivnosti gena u odnosu na ribosomske gene napravila sam u programskom jeziku R. Gene, tj. njihove KEGG brojeve, također sam povezala i s pripadajućim brojevima KEGG modula (Kanehisa 2018) i izračunala sam ekspresivnost i relativnu ekspresivnost KEGG modula. Relativnu ekspresivnost za svaki gen  $x$  izračunala sam kao  $E_x$  prema formuli (5) gdje je  $G_x$  broj

predviđenih gena  $x$  s MELP vrijednosti većom od 1,  $A_x$  ukupni broj predviđenih gena  $x$ ,  $G_t$  ukupni broj predviđenih gena s MELP vrijednosti većom od 1, a  $A_t$  ukupni broj predviđenih gena.

$$E_x = \frac{(G_x + 1) - (A_x * \frac{G_t}{A_t} + 1)}{(A_x * \frac{G_t}{A_t} + 1)} * 100 \quad (5)$$

Za računanje relativne ekspresivnosti KEGG modula  $G_x$  i  $A_x$  definirani su kao zbrojevi vrijednosti  $G_x$  i  $A_x$  pripadajućih gena. S obzirom da se *anammox* metabolički put ne nalazi u listi KEGG modula, ručno sam ga pridodala skupu podataka zbrajanjem vrijednosti  $G_x$  i  $A_x$  za gene anotirane KEGG brojevima K20932, K20933, K20934 i K20935.

Graf procijenjene relativne ekspresivnosti gena uključenih u metabolizam dušika napravila sam izdvajanjem tih gena pomoću njihovih KEGG brojeva iz tablice koja sadržava procijenjenu ekspresivnost (MELP) svih predviđenih gena. KEGG brojeve gena vezanih uz metabolizam dušika pronašla sam u KEGG opisu metaboličkog puta MAP00910 ([https://www.genome.jp/kegg-bin/show\\_pathway?map00910](https://www.genome.jp/kegg-bin/show_pathway?map00910)).

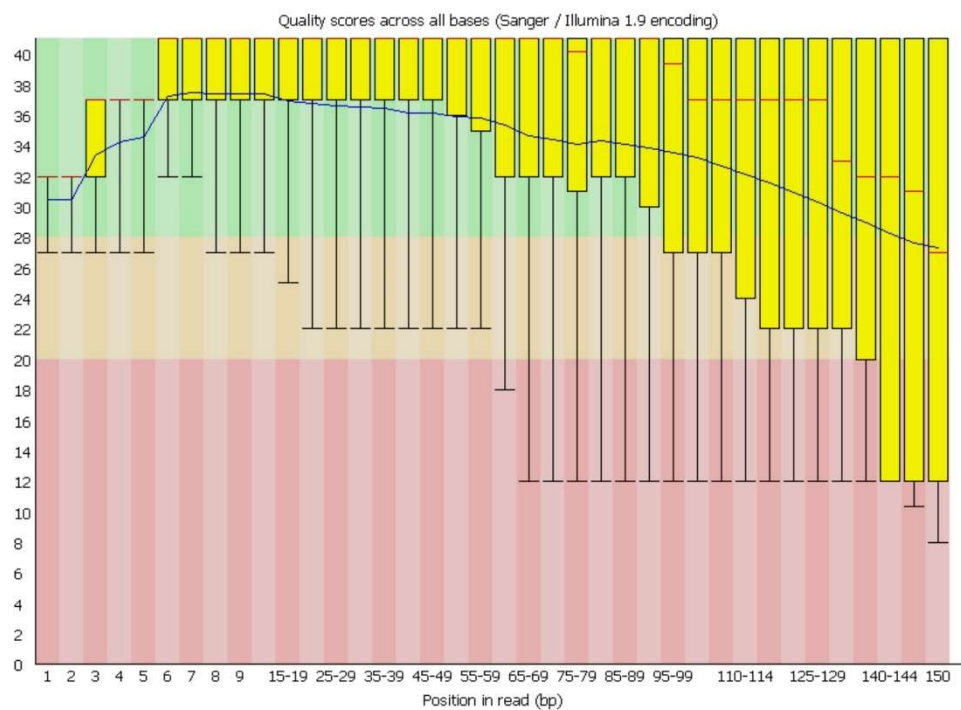
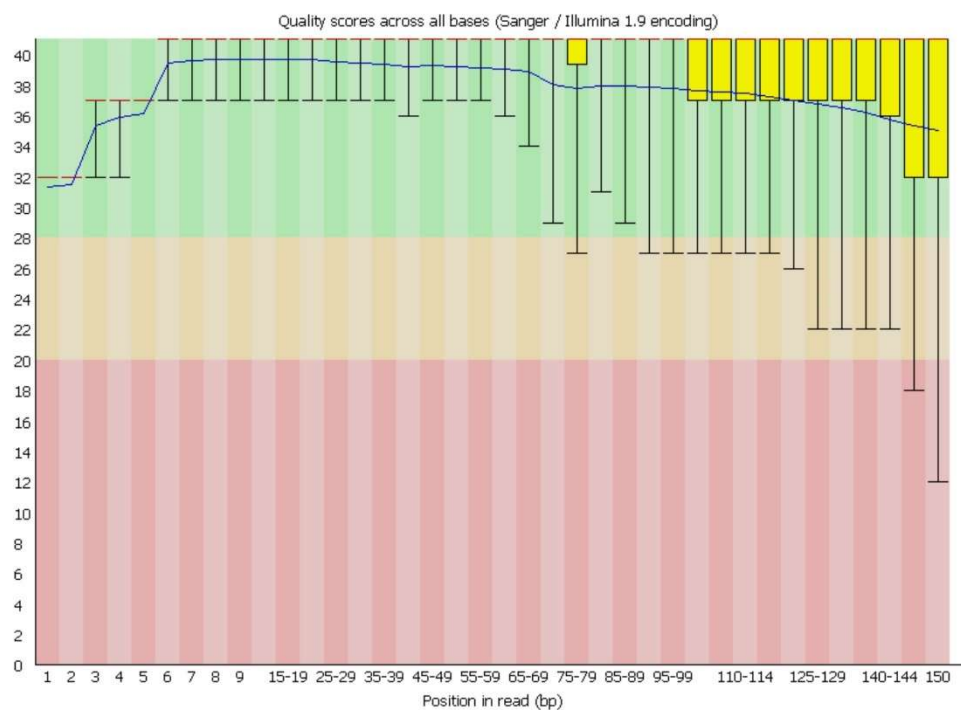
Rezultate sam zatim prilagodila za unos u mrežnu aplikaciju IPATH3 (Darzi i sur. 2018) radi vizualizacije predviđene relativne ekspresivnosti na shematskim prikazima metaboličkih puteva (eng. *Metabolic pathways*) i mikrobnog metabolizma u raznolikim okolišima (eng. *Microbial metabolism in diverse environments*). Kreirala sam tekstualnu datoteku takvu da je svakom broju KEGG proteina ili modula pridružena boja (crvena za predviđenu negativnu relativnu ekspresivnost, plava za predviđenu pozitivnu relativnu ekspresivnost) i širina linije ili točke kojom će protein ili modul biti nacrtan. Širine odgovaraju apsolutnoj vrijednosti relativne ekspresivnosti, a normalizirala sam ih na maksimalnu dopuštenu širinu u aplikaciji IPATH3 (30 piksela). Uzete su samo vrijednosti s  $p < 0,05$ .

## 3. REZULTATI

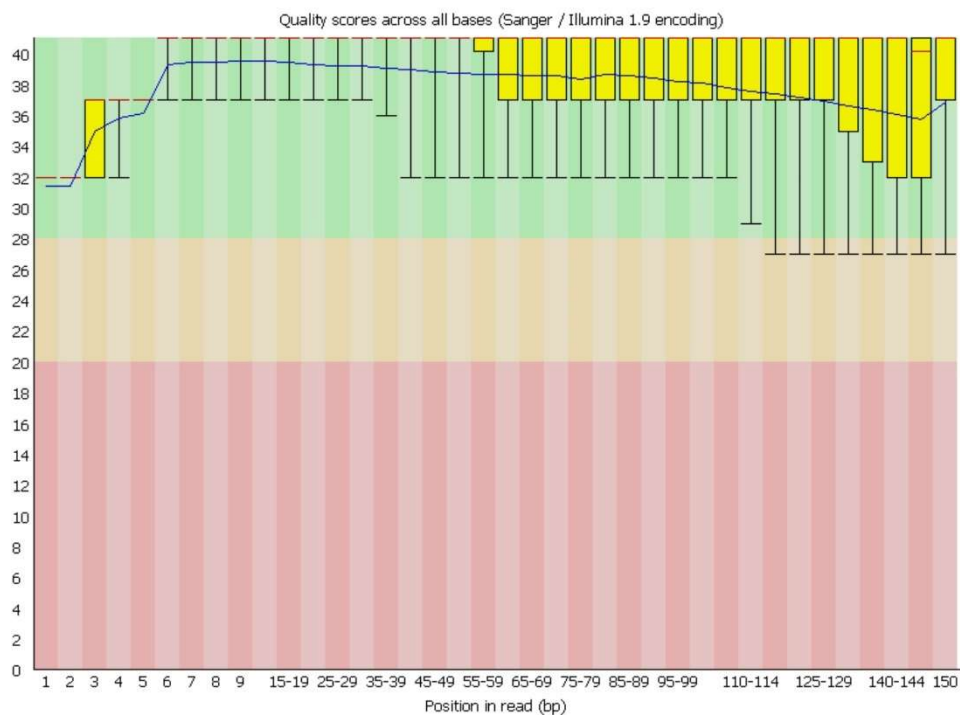
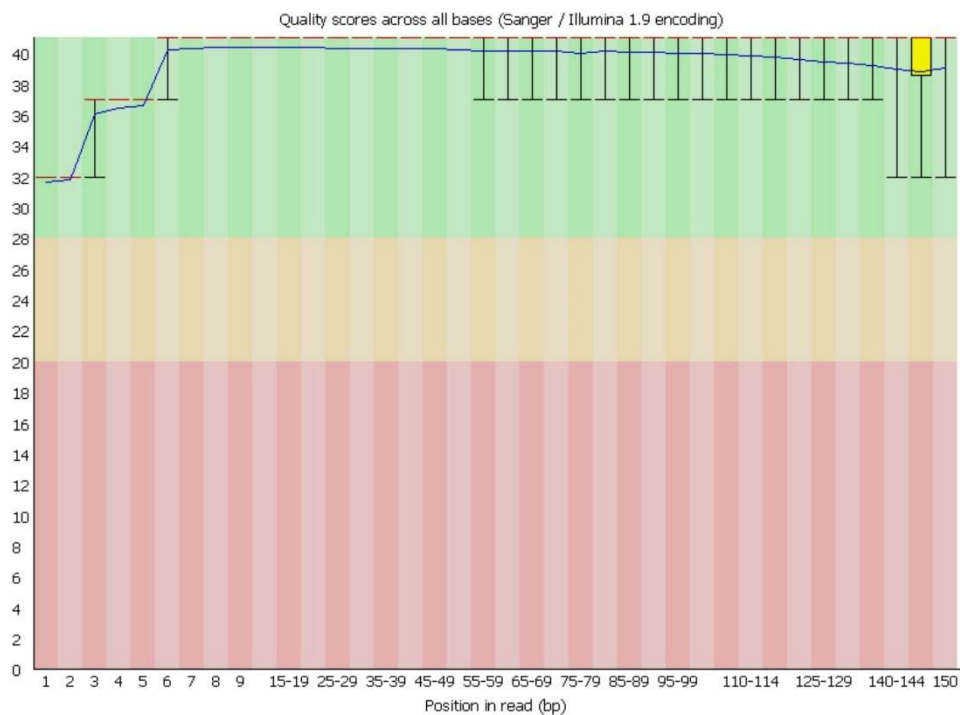
### 3.1. Kontrola kvalitete i filtriranje kontaminacije

Sirovi podaci sastojali su se od dvije datoteke formata fastq ukupne veličine 18706243 KB u kojima je sadržano ukupno 49999055 očitanih uparenih sljedova (eng. *Paired-end reads*) duljine 150 baza. Provjera kvalitete očitanih sljedova vizualizacijom pomoću programa *FastQC* pokazala je da niti jedan nije označen kao slijed loše kvalitete, ali da su sljedovi fragmenata sekvencirani s prvog kraja fragmenta (P1) bolje kvalitete u odnosu na sljedove drugog kraja fragmenta (P2), a posebno prema kraju očitanih sljedova (Slika 4). Sljedovi *Illumina TruSeq* adaptera, korištenog kod sekvenciranja ovoga metagenoma, identificirani su u 3,86% sljedova P1 i 3,7% sljedova P2. Nakon izrezivanja i filtriranja pomoću programa *Trimmomatic* preostalo je 37273626 ili 74,55% uparenih sljedova. Slika 5 prikazuje grafove prosječne kvalitete očitane baze po poziciji u slijedu. Nakon izrezivanja i filtriranja sljedovi *Illumina TruSeq* adaptera više se ne pojavljuju u skupu podataka.

Program *bowtie2* uspješno je poravnao 204817 ili 0,55% sljedova filtriranoga metagenoma s ljudskim genomom verzije GRCh38 te ih uklonio iz izlaznih datoteka. Nakon uklanjanja humanih sljedova iz metagenoma preostalo je 37068809 ili 99,45% sljedova filtriranoga metagenoma.



Slika 4 Grafovi prosječne kvalitete očitane baze po poziciji u skupu sljedova sekvenciranih s prve (gore) i druge (dolje) strane izražene u phred vrijednosti  $Q$  prije filtriranja i izrezivanja. Crne linije označavaju raspon između 10. i 90. percentila, žuti pravokutnici raspon između 1. i 3. kvartila, crvene linije označavaju medijanu, a modra linija prosjek kvaliteta.



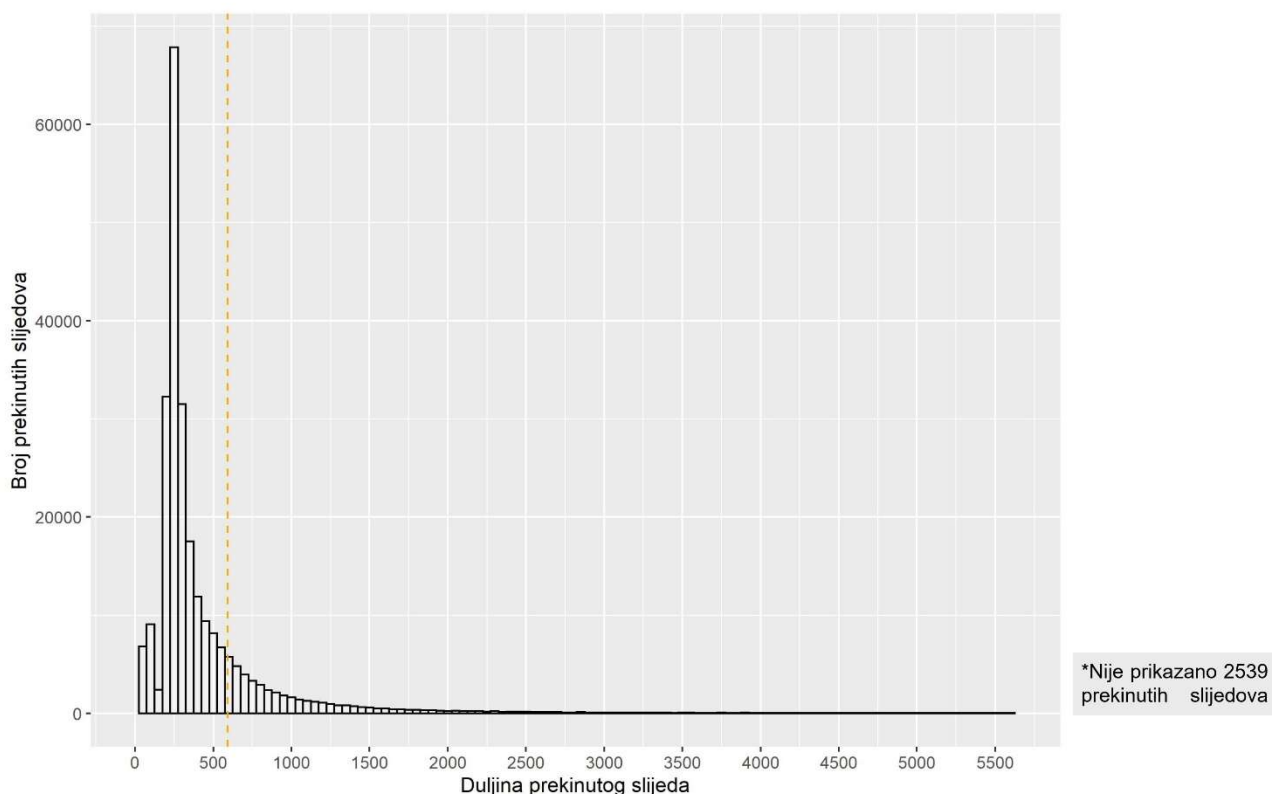
Slika 5 Grafovi prosječne kvalitete očitane baze po poziciji u skupu sljedova sekvenciranih s prve (gore) i druge (dolje) strane izražene u phred vrijednosti  $Q$  nakon filtriranja i izrezivanja prema kvaliteti. Crne linije označavaju raspon između 10. i 90. percentila, žuti pravokutnici raspon između 1. i 3. kvartila, crvene linije označavaju medijan, a modra linija prosjek kvaliteta.



### 3.2. Sastavljanje metagenoma

Očitani sljedovi metagenoma sastavljeni su u 253808 metagenomskih prekinutih sljedova ukupne duljine od 150548980 nukleotida. Prosječna duljina prekinutog slijeda bila je 593,16 nukleotida, a medijan duljine 283 nukleotida. Vrijednost N50 bila je 912 nukleotida. Na Slici 6 prikazan je histogram duljina sastavljenih prekinutih sljedova. Prikazane su duljine do 99. percentila, do duljine od 5633,93 nukleotida. Duljih od toga bilo je 2539 prekinutih sljedova koji nisu prikazani na grafu. Najdulji sastavljeni prekinuti slijed bilo je duljine 181189 nukleotida.

Ukupno 88,33% očitanih sljedova uspješno je poravnato sa sastavljenim prekinutim sljedovima. Parova sljedova iz kojih su oba slijeda poravnata u točnoj orijentaciji i udaljenosti bilo je 67,83%. Ostatak poravnanja čine očitani sljedovi koji su uspješno poravnati s prekinutim sljedovima, ali čiji parovi nisu.



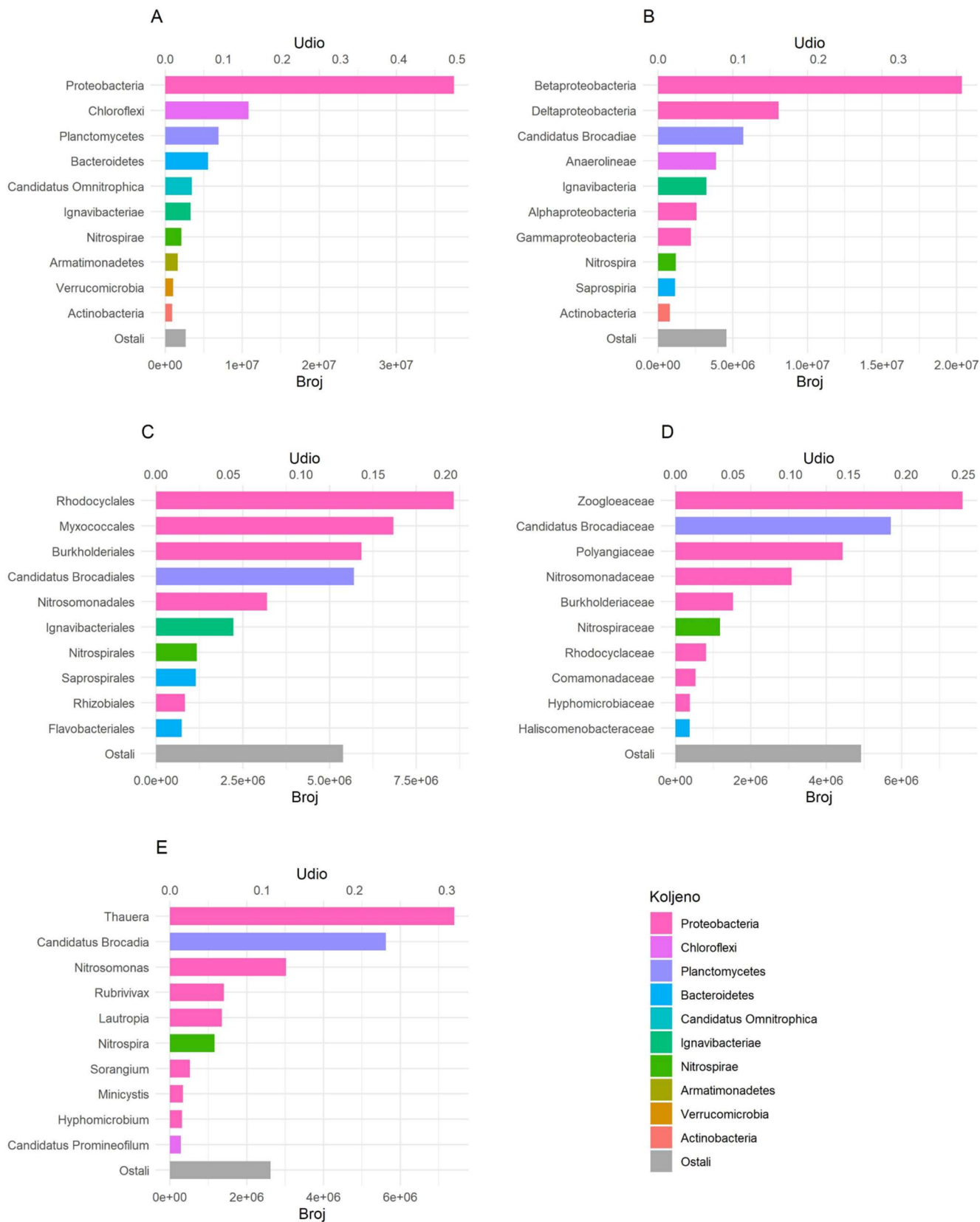
Slika 6 Histogram duljina sastavljenih metagenomskih prekinutih sljedova. Prikazan je 99. percentil duljina. Narandčastom iscrtkanom linijom označena je prosječna duljina prekinutoga slijeda.

### 3.3. Taksonomski sastav bioreaktora

Program *Megan* uspješno je klasificirao 95594816 od 113890864 nukleotidne baze odnosno njih 83,94%. Od toga 98,51% klasificirano je u domenu *Bacteria*, 0,51% u *Archaea*, a 0,99% u *Eukaryota*. Na razini koljena klasificirane su 75788402 nukleotidne baze. Među njima dominiraju *Proteobacteria* s 49,47% nukleotidnih baza, a slijede *Chloroflexi*, *Planctomycetes* i *Bacteroidetes* s redom 14,28%, 9,11% i 7,3% baza klasificiranih kao ta koljena (Slika 7A). Nukleotida klasificiranih u neki razred je 53691382. Od toga najviše nukleotidnih baza anotirano je kao razredi *Betaproteobacteria* i *Deltaproteobacteria* koljena *Proteobacteria*, *Ca. Brocadiaceae* koljena *Planctomycetes* i *Anaerolineae* koljena *Chloroflexi* s redom 37,89%, 15,04%, 10,62% i 7,2% anotiranih baza (Slika 7B). Kao red anotirano je ukupno 41737098 baza, od toga najviše kao *Rhodocyclales* (razred *Betaproteobacteria*), *Myxococcales* (razred *Deltaproteobacteria*), *Burkholderiales* (razred *Betaproteobacteria*), *Candidatus Brocadiaceae* (razred *Ca. Brocadiaceae*), *Nitrosomonadales* (razred *Betaproteobacteria*) i *Ignavibacteriales* (koljeno *Ignavibacteriae*, razred *Ignavibacteriae*) s redom 20,54%, 16,39%, 14,19%, 13,67%, 7,65% i 5,33% anotiranih baza (Slika 7C). Kao porodica anotirano je 30564426 baza, od toga najviše kao *Zoogloeaceae* (*Rhodocyclales*), *Ca. Brocadiaceae* (*Ca. Brocadiaceae*), *Polyangiaceae* (*Myxococcales*), *Nitrosomonadaceae* (*Nitrosomonadales*) i *Burkholderiaceae* (*Burkholderiales*) s redom 24,9%, 18,66%, 14,51%, 10,1% i 5% anotiranih baza (Slika 7D). Od 24047692 baza uspješno klasificiranih u rod, prevladavaju *Thauera* (*Zoogloeaceae*), *Ca. Brocadia* (*Ca. Brocadiaceae*), *Nitrosomonas* (*Nitrosomonadaceae*), *Rubrivivax* (*Burkholderiales genera incertae sedis*) i *Lautropia* (*Burkholderiaceae*) s redom 30,79%, 23,4%, 12,59%, 5,84% i 5,61% anotiranih baza (Slika 7E).

Identificirane su četiri vrste ili soja roda *Thauera*: *T. phenylacetatica*, *Thauera* sp. SWB20, *T. aminoaromatica*, i *Thauera* sp. MZ1T s redom 70,15%, 14,05%, 9,50%, i 4,31% baza klasificiranih do razine vrste unutar roda *Thauera*; četiri vrste ili soja roda *Ca. Brocadia*: *Ca. B. sapporoensis*, *Ca. B. fulgida*, *Ca. Brocadia* sp. UTAMX2 i *Ca. Brocadia* sp. UTAMX1 s redom 47,53%, 19,99%, 17,80% i 14,68% baza klasificiranih do razine vrste unutar roda *Ca. Brocadia*; dvije vrste roda *Nitrosomonas*: *N. europaea* i *N. eutropha* (87,86% i 12,14% roda *Nitrosomonas*); tri vrste roda *Nitrospira*: *Nitrospira* sp. OLB3, *N. moscoviensis*, *N. japonica* (41,52%, 35,42%, 23,06%). Kod ostalih šest od deset najzastupljenijih rodova identificirana je samo po jedna vrsta:

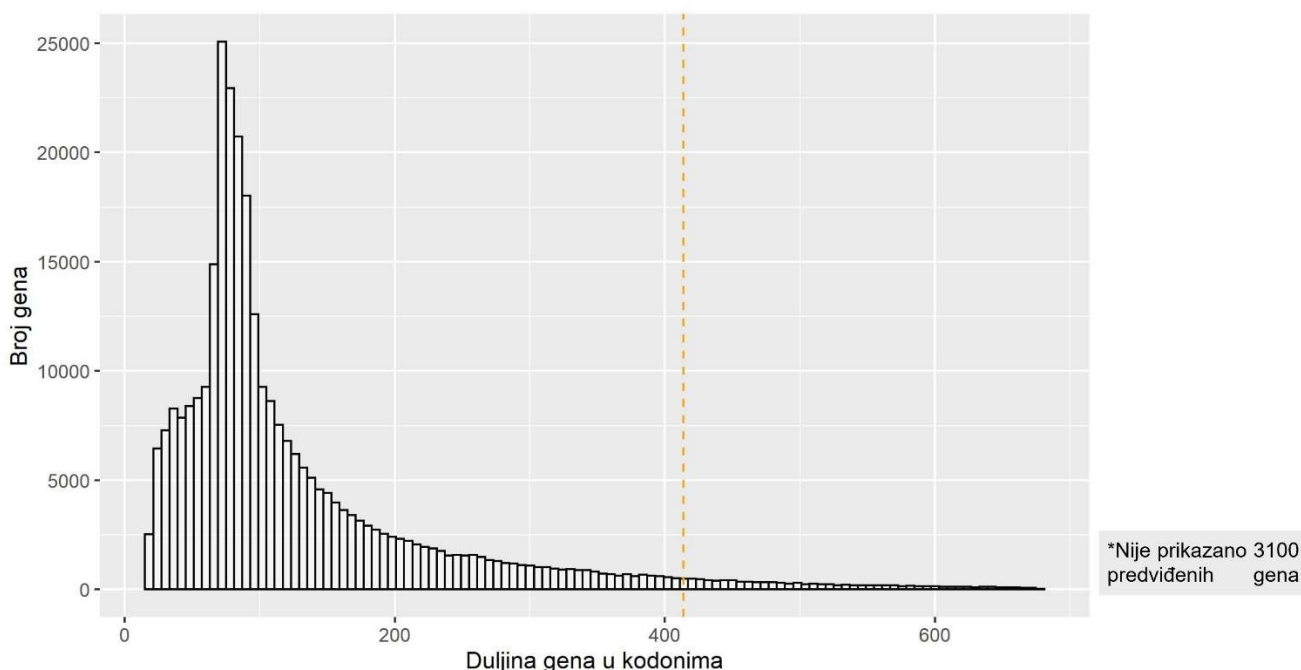
*Lautropia sp.* SCN 69-89, *Rubrivivax sp.* SCN 71-131, *Sorangium cellulosum*, *Minicystis rosea*,  
*Ca. Promineofilum breve*, i *Hyphomicrobium sp.*



Slika 7 Taksonomski sastav bioreaktora na razini (A) koljena, (B) razreda, (C) reda, (D) porodice i (E) roda prema algoritmu UI-LCA programa Megan. Broj na x osi odgovara ukupnom broju baza prekinutih sljedova klasificiranih u pojedinu taksonomsku skupinu.

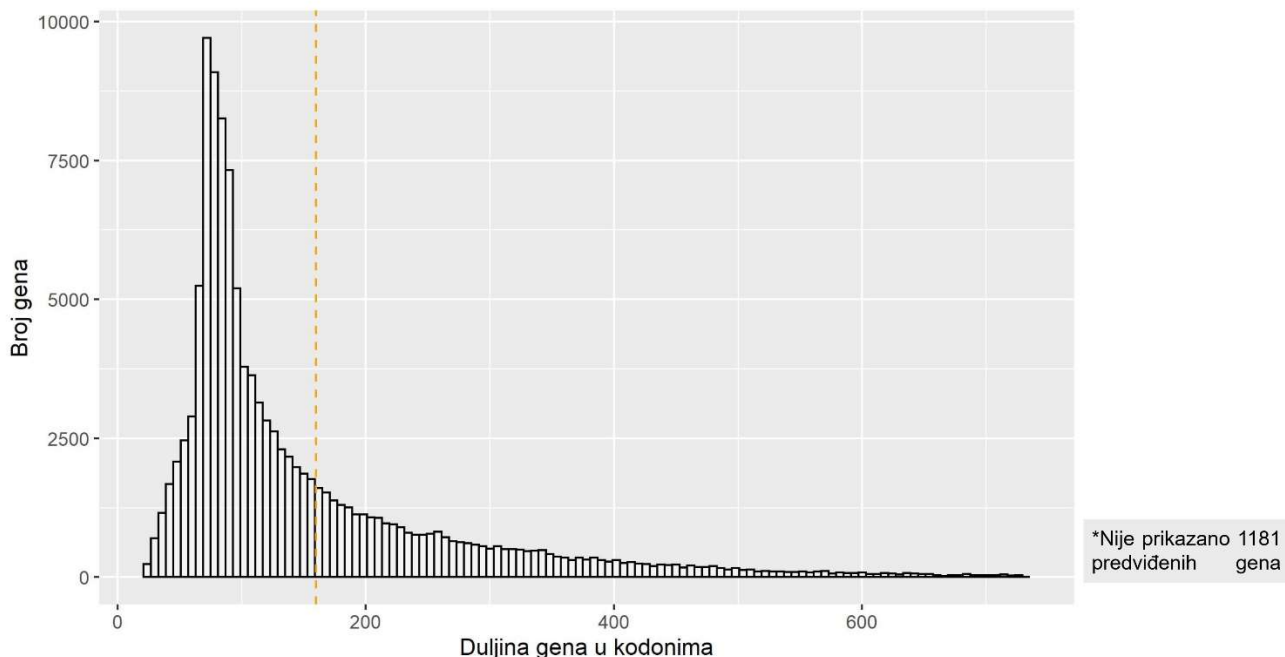
### 3.4. Predviđanje gena koji kodiraju proteine i funkcionalna anotacija pomoću mapiranja na bazu ortologije KEGG

Program *MetaGeneMark* predvidio je 310487 gena koji kodiraju proteine. Prosječna duljina predviđenog gena je 413,58, a medijan 276 kodona. Na Slici 9 prikazan je histogram duljina predviđenih gena u kodonima. Prikazane su duljine do 99. percentila tj. do 678 kodona; 3100 gena je dulje i nije prikazano na grafu.



*Slika 8 Histogram duljina gena (u kodonima) koji su predviđeni iz skela metagenoma pomoću programa MetaGeneMark; prikazan je 99. percentil duljina. Veličina koraka histograma je 6. Narančastom iscrtkanom linijom označena je prosječna duljina predviđenih gena.*

Na KEGG bazu poravnato je 203291 gena tj. 65,48% predviđenih, a isključujući poravnanja bez KEGG identifikacijskog broja poravnato je 118640 gena, tj. 38,21% predviđenih. Prosječna duljina poravnatih gena s KEGG brojem je 159,66, a medijan 105 kodona. Slika 10 prikazuje histogram duljina gena čiji rezultati poravnanja na KEGG bazu sadrže KEGG identifikacijski broj. Prikazane su duljine do 99. percentila tj. do 731 kodona; 1181 gena je dulje i nije prikazano na grafu.



*Slika 8 Histogram duljina (u kodonima) predviđenih gena koji su uspješno poravnati na KEGG bazu i čiji rezultati poravnanja sadrže KEGG identifikacijski broj proteina; prikazan je 99. percentil duljina. Veličina koraka histograma je 6. Narančastom iscrtkanom linijom označena je prosječna duljina poravnatih gena.*

Slika 10 prikazuje udio uporabe pojedinih kodona skupa predviđenih gena s pridodanim KEGG identifikacijskim brojem i RSCU vrijednosti za svaki kodon s jednolikom uporabom uzetom kao očekivanom. Kodon CGC pokazuje najveći pozitivni odmak od očekivane uporabe s 45,31% arginina kodiranih tim kodonom naspram očekivanih 16,67% i RSCU vrijednosti 2,72; slijede kodoni CTG s RSCU vrijednosti 2,18, GGC s RSCU vrijednosti 2,08 i CTC s RSCU vrijednosti 1,91. Najveći negativni odmak od očekivane uporabe pokazuje kodon CTA koji kodira samo 1,36% leucina s RSCU vrijednosti 0,08, a slijede kodoni TTA s RSCU vrijednosti 0,17, AGA s RSCU vrijednosti 0,25, te AGG s RSCU vrijednosti 0,29. Izračunat je udio GC baza od 60,35% među predviđenim genima s pridodanim KEGG identifikacijskim brojem.

Među predviđenim genima s pridruženim KEGG brojevima pronađeno je 2150 ribosomskih gena. Slika 11 prikazuje uporabu kodona ribosomskih gena s RSCU vrijednostima izračunatim s prosječnom uporabom kodona skupa gena (Slika 10) kao očekivanom. Najveći pozitivni odmak od očekivane uporabe pokazuju kodoni CTT s RSCU vrijednosti 2,02, AGA s RSCU vrijednosti 1,98, TCT s RSCU vrijednosti 1,75, te TTA s RSCU vrijednosti 1,59. Slika 12 prikazuje prosječni udio pojedine baze prema poziciji u kodonu kod svih predviđenih i ribosomskih gena.

Prikazana je također i očekivana raspodjela prema udjelu pojedine baze u predviđenim genima (A: 19,73%; C: 30,13%; T: 19,93%; G: 30,22%).

Prva baza	Druga baza												Treća baza				
	T			C			A			G							
T	TTT	0.3582	0.7164	Phe	TCT	0.0867	0.5201	Ser	TAT	0.4042	0.8084	Tyr	TGT	0.2928	0.5857	Cys	T
	TTC	0.6418	1.2836		TCC	0.2424	1.4546		TAC	0.5958	1.1916		TGC	0.7072	1.4143		C
	TTA	0.0289	0.1732	TCA	0.0674	0.4046	TAA		0.403	1.2091	Stop	TGA	0.4024	1.2073	Stop	A	
	TTG	0.1039	0.6235	TCG	0.2248	1.3490	TAG		0.1946	0.5837	Stop	TGG	1	1.0000	Trp	G	
C	CTT	0.1732	1.0390	Leu	CCT	0.0947	0.3790	Pro	CAT	0.3999	0.7999	His	CGT	0.1361	0.8169	Arg	T
	CTC	0.3183	1.9101		CCC	0.3169	1.2676		CAC	0.6001	1.2001		CGC	0.4531	2.7183		C
	CTA	0.0138	0.0827		CCA	0.1758	0.7030		CAA	0.221	0.4421	CGA	0.1078	0.6465	A		
	CTG	0.3619	2.1716		CCG	0.4126	1.6504		CAG	0.779	1.5579	Gln	CGG	0.2125	1.2751		G
A	ATT	0.2618	0.7855	Ile	ACT	0.1147	0.4587	Thr	AAT	0.4672	0.9343	Asn	AGT	0.0998	0.5988	Ser	T
	ATC	0.5843	1.7528		ACC	0.4529	1.8116		AAC	0.5328	1.0657		AGC	0.2788	1.6729		C
	ATA	0.1539	0.4617		ACA	0.1065	0.4261		AAA	0.4658	0.9315	Lys	AGA	0.0415	0.2488	A	
	ATG	1	1.0000	Met	ACG	0.3259	1.3036		AAG	0.5342	1.0685	AGG	0.0491	0.2943	G		
G	GTT	0.1832	0.7326	Val	GCT	0.1123	0.4490	Ala	GAT	0.4808	0.9616	Asp	GGT	0.1939	0.7757	Gly	T
	GTC	0.3865	1.5460		GCC	0.419	1.6760		GAC	0.5192	1.0384		GGC	0.5179	2.0717		C
	GTA	0.14	0.5600		GCA	0.0899	0.3597		GAA	0.4586	0.9172	GGA	0.1193	0.4774	A		
	GTG	0.2903	1.1613		GCG	0.3788	1.5152		GAG	0.5414	1.0828	Glu	GGG	0.1688	0.6752		G
	Kodon	Udio uporabe kodona	RSCU	Aminokiselina	Kodon	Udio uporabe kodona	RSCU	Aminokiselina	Kodon	Udio uporabe kodona	RSCU	Aminokiselina	Kodon	Udio uporabe kodona	RSCU	Aminokiselina	

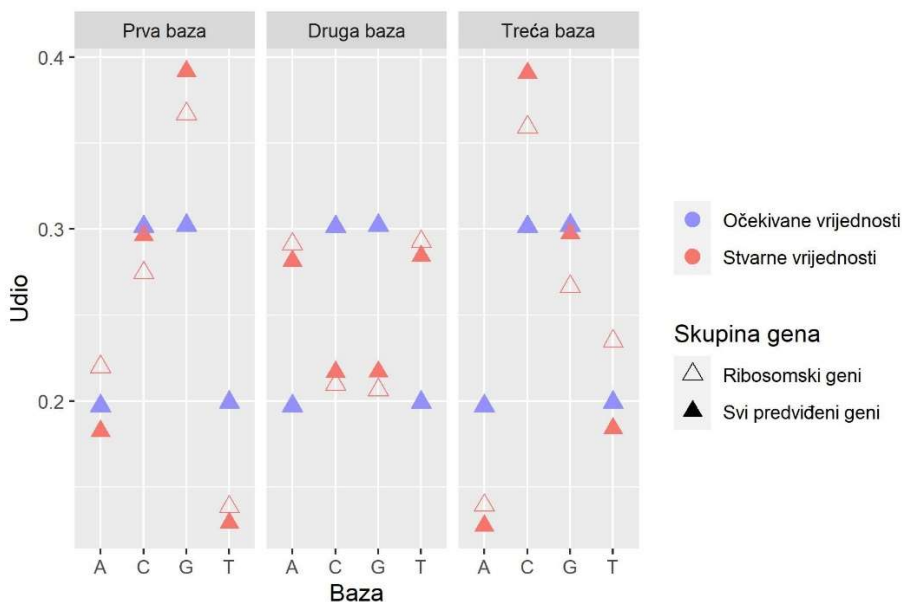
Slika 10 Prikaz uporabe kodona kod skupa predviđenih gena s pridodanim KEGG identifikacijskim brojem. Brojevi neposredno desno uz svaki kodon označavaju udio uporabe tog kodona i RSCU (izračunat s jednolikom uporabom kodona kao očekivanom) redom. Neposredno desno uz brojeve zapisana je i aminokiselina koju kodon kodira standardnim troslovnim simbolom.

Analiza MELP vrijednosti, tj. procjena relativne ekspresivnosti za predviđene gene grupirane u kategorije KEGG modula prikazana je na Slici 13. Od modula vezanih uz metabolizam dušika visoka relativna ekspresivnost predviđena je za gene povezane s nitritacijom, za *anammoX* gene, gene vezane uz asimilaciju nitrata i gene vezane uz denitrifikaciju. Među modulima s predviđenom malom relativnom ekspresivnosti nalazi se i asimilatorna redukcija nitrata u amonijak, tj. proces konverzije nitrata u biomasu putem amonijaka i sinteze biomolekula s amino skupinama. U Prilogu 2 se nalazi i graf procijenjene relativne ekspresivnosti za pojedine gene. Slika 13 prikazuje procjenu relativne ekspresivnosti za gene uključene u metabolizam dušika. Interaktivnom prikazu procijenjenih ekspresivnosti za sve analizirane module i gene moguće je pristupiti pomoću datoteke koja se nalazi u Prilogu 3.

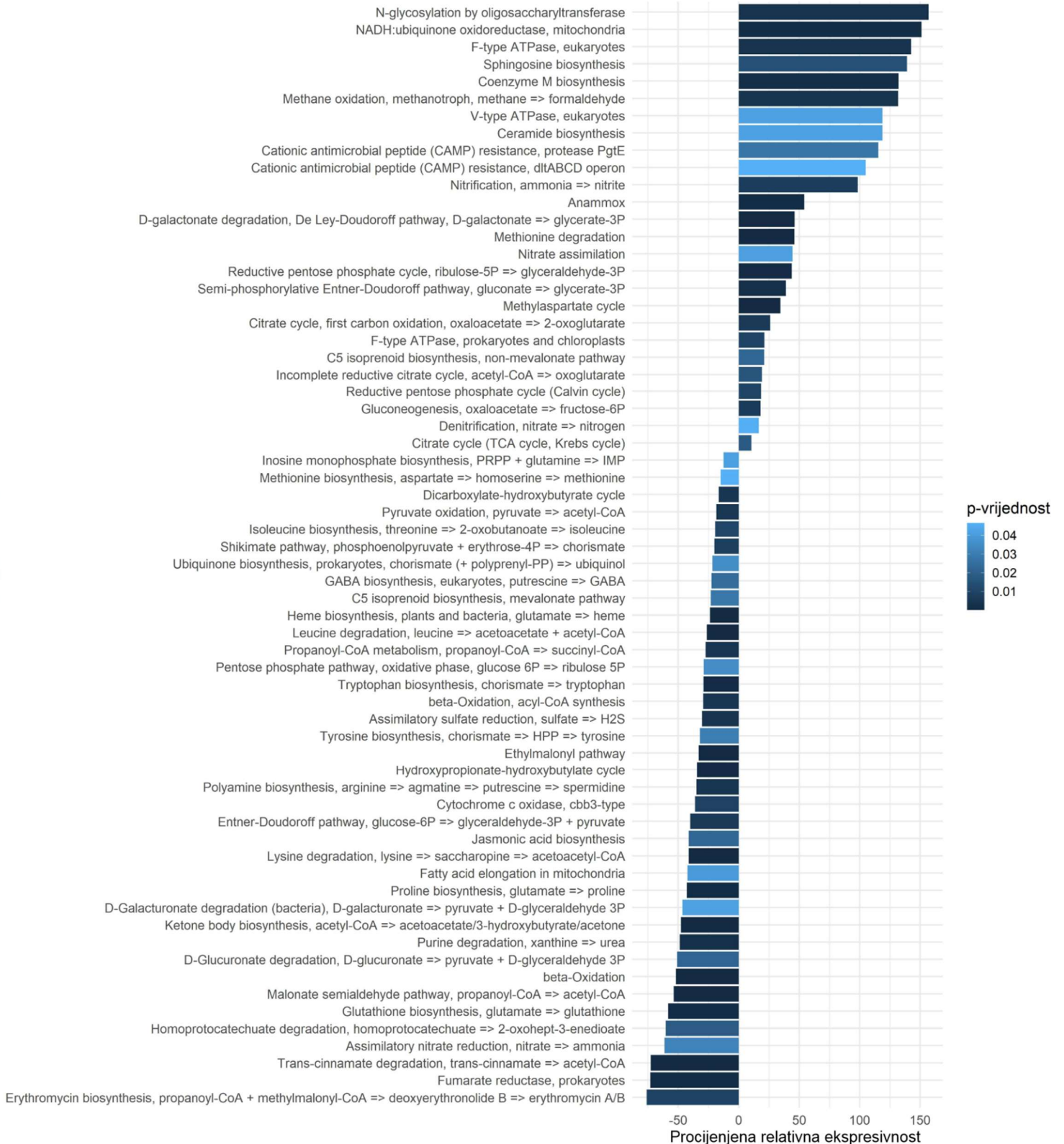


Prva baza	Druga baza												Treća baza				
	T			C			A			G							
T	TTT	0.47632	1.3338	Phe	TCT	0.1515	1.7467	Ser	TAT	0.40354	1.0002	Tyr	TGT	0.32416	1.1133	Cys	T
	TTC	0.52368	0.8146		TCC	0.22191	0.9149		TAC	0.59646	0.9998		TGC	0.67584	0.9534		C
	TTA	0.04567	1.5906		TCA	0.07427	1.1027		TAA	0.46679	1.1540		TGA	0.3607	0.8989		A
	TTG	0.09143	0.8870		TCG	0.1859	0.8249		TAG	0.17251	0.8883		TGG	1	1.0000		G
C	CTT	0.34973	2.0197	Leu	CCT	0.14716	1.5560	Pro	CAT	0.4259	1.0652	His	CGT	0.16771	1.2348	Arg	T
	CTC	0.24273	0.7612		CCC	0.29346	0.9224		CAC	0.5741	0.9565		CGC	0.40738	0.8959		C
	CTA	0.01603	1.1803		CCA	0.15596	0.9013		CAA	0.25418	1.1607		CGA	0.11306	1.0511		A
	CTG	0.25441	0.7016		CCG	0.40343	0.9739		CAG	0.74582	0.9549		CGG	0.16909	0.7983		G
A	ATT	0.31214	1.1938	Ile	ACT	0.12158	1.0687	Thr	AAT	0.45087	0.9647	Asn	AGT	0.10965	1.1065	Ser	T
	ATC	0.53207	0.9074		ACC	0.44909	0.9880		AAC	0.54913	1.0310		AGC	0.25676	0.9206		C
	ATA	0.15579	1.0239		ACA	0.10307	0.9760		AAA	0.41088	0.8861		AGA	0.08199	1.9819		A
	ATG	1	1.0000		ACG	0.32626	1.0006		AAG	0.58912	1.0985		AGG	0.06077	1.2469		G
G	GTT	0.24485	1.3429	Val	GCT	0.13456	1.2058	Ala	GAT	0.49353	1.0262	Asp	GGT	0.24701	1.2748	Gly	T
	GTC	0.35097	0.9065		GCC	0.41002	0.9769		GAC	0.50647	0.9758		GGC	0.48029	0.9246		C
	GTA	0.14341	1.0324		GCA	0.08473	0.9479		GAA	0.50903	1.1120		GGA	0.13805	1.1615		A
	GTG	0.26077	0.8944		GCG	0.37069	0.9773		GAG	0.49097	0.9054		GGG	0.13465	0.8019		G

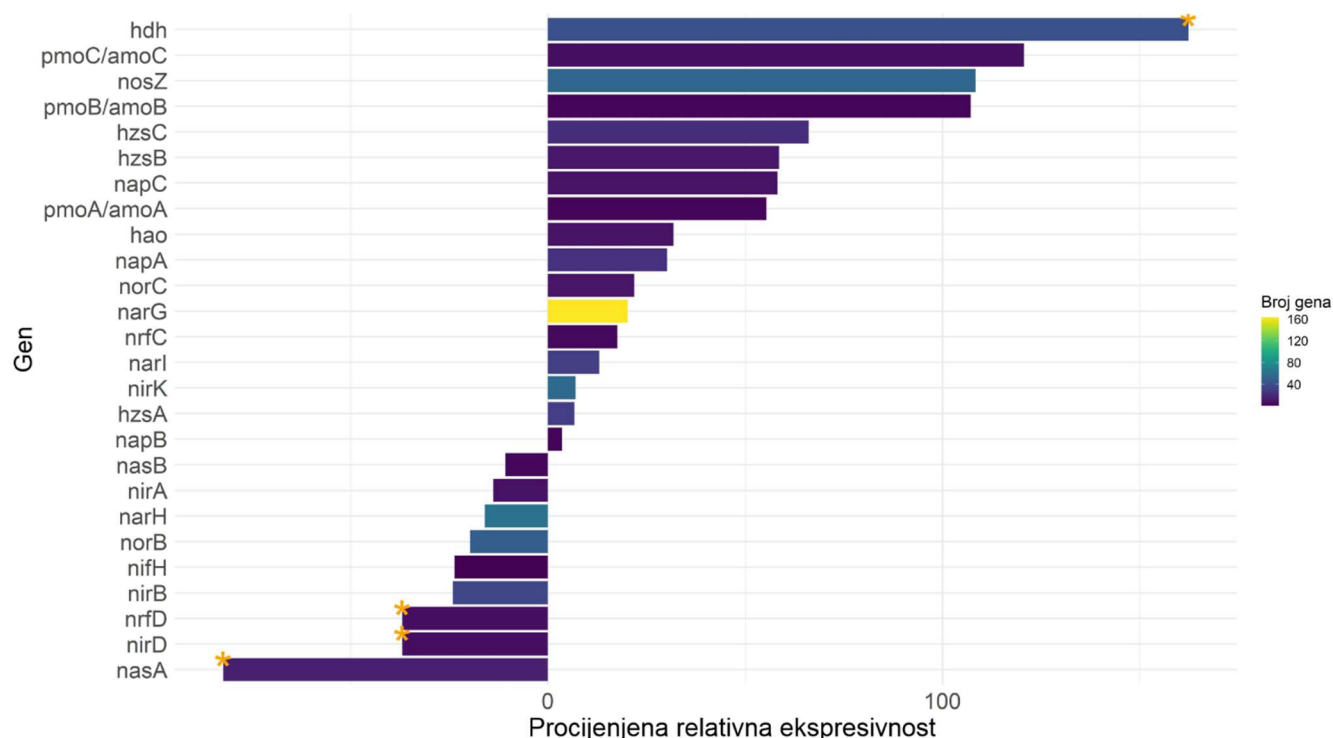
Slika 11 Prikaz uporabe kodona ribosomskih proteina s RSCU izračunatim uzimajući prosječnu uporabu kodona skupa predviđenih gena kao očekivanu. Brojevi neposredno desno uz svaki kodon označavaju udio uporabe tog kodona i RSCU redom. Neposredno desno uz brojeve zapisana je i aminokiselina koju kodon kodira standardnim troslovnim simbolom.



Slika 92 Prikaz raspodjele pojedinih baza prema poziciji u kodonu kod ribosomskih gena (prazni crveni trokuti) i svih predviđenih gena (puni crveni trokuti). Plavi trokuti označavaju udio pojedine baze među predviđenim genima bez obzira na poziciju u kodonima.



Slika 13 Graf relativne ekspresivnosti kategorija KEGG modula gena grupiranih po kategoriji. Relativna ekspresivnost za svaki modul  $x$  definirana je kao  $E_x$  u jednažbi (5). Prikazani su samo rezultati s  $p < 0,05$ .



Slika 104 Graf relativne ekspresivnosti svih gena uključenih u metabolizam dušika pronađenih među predviđenim genima: hidrazin dehidrogenaza (*hdh*); podjedinice A, B i C metan/amonijak monooksigenaze (*pmo/amoABC*); dušikov (I) oksid reduktaza (*nosZ*); podjedinice A, B i C hidrazin sintaze (*hzsABC*); podjedinice A, B i C periplazmatske nitrit reduktaze (*napABC*); hidroksilamin dehidrogenaza (*hao*); podjedinice A, B i C dušikov (II) oksid reduktaze (*norABC*), podjedinice  $\alpha$ ,  $\beta$  i  $\gamma$  respiratorne nitrat reduktaze / nitrit oksidoreduktaze (*narGHI*); podjedinice C i D citokrom c-552 nitrit reduktaze (*nrfCD*), nitrit reduktaza koja proizvodi dušikov (II) oksid (*nirK*); podjedinice A i B asimilatorne nitrat reduktaze (*nasAB*); feredoksin-nitrat reduktaza (*nirA*); podjedinica H nitrogenaze (*nifH*); podjedinice B i D NADH-nitrit reduktaze (*nirBD*). Relativna ekspresivnost za svaki gen  $x$  definirana je kao  $E_x$  u jednadžbi (5). Boje na grafu odgovaraju broju pojedinog gena pronađenom među predviđenim genima. Zbog relativno malog broja pronađenih pojedinih gena među predviđenim genima, za mnoge vrijednosti relativne ekspresivnosti  $p$  vrijednost je visoka. Narančastom zvjezdicom označeni su rezultati  $s p < 0,05$ .

## 4. RASPRAVA

### 4.1. Kvaliteta očitanih sljedova i sastavljenoga metagenoma

Pad kvalitete očitanih sljedova prema kraju fragmenata i lošija kvaliteta P2 očitanih sljedova česte su i očekivane pojave kod *Illumina* tehnologije sekvenciranja, iako je razlika u kvaliteti sljedova P1 i P2 slabo opisana u literaturi (Tan i sur. 2019). Kvaliteta pada prema kraju najčešće zbog kumulativne asinkronizacije sinteze molekula DNA s povećanjem broja ciklusa sekvenciranja, uglavnom zbog neuspješnog otcjepljivanja terminacijskih dijelova obilježenih nukleotida (Dohm i sur. 2008). Očitani sljedovi P2 vjerojatno su ukupno lošije kvalitete zato što se oni sintetiziraju i očitavaju tek nakon svih 150 ciklusa sekvenciranja sljedova P1, tijekom kojih je moguća degradacija dijela molekula DNA kalupa. Prosječna kvaliteta nefiltriranih sljedova P2 ni do 150. nukleotida ipak nije pala ispod *phred* vrijednosti 25, što odgovara preko 99,99% sigurnosti u očitane baze.

Moguće je da su humani sljedovi pronađeni u metagenomu potekli od kontaminacije genetskim materijalom samih istraživača koji su pripremali metagenomsku biblioteku, no moguće je i da je izvor sintetička voda koja je unošena u bioreaktor. Oko 10% volumena te sintetičke otpadne vode čini prava komunalna otpadna voda (Boeije i sur. 1999) u kojoj izvor humanih sljedova mogu biti urin, fekalije, kožne stanice i ostalo.

Prosječna duljina prekinutog slijeda od 593,16 nukleotida i medijan od 283 nukleotida opisuju vrlo fragmentiran metagenom; kod analize metagenoma često se u obzir uzimaju samo prekinuti sljedovi dulji od 500 nukleotida. Vrijednost N50 od 912 nukleotida nešto je prihvatljivija; vrijednosti iznad 1000 nukleotida povezuju se s vrlo visokom kvalitetom složenih metagenoma (van der Walt i sur. 2017). Postotak očitanih sljedova od kojih su složeni prekinuti sljedovi od 88,33% vrlo je visok. Izbor da se iz analize ne isključe prekinuti sljedovi kraći od 500 nukleotida bio je potaknut težnjom da u taksonomskom sastavu budu predstavljene i skupine manje brojnosti za koje su kratki prekinuti sljedovi i mala pokrivenost očekivani. Kratki prekinuti sljedovi mogu kod poravnanja na baze podataka dovesti do pogrešne klasifikacije zbog višestrukih poravnanja visoke kvalitete, no zbog UI-LCA algoritma programa *Megan* takva višestruka poravnanja na

proteine različitih taksonomskih skupina dovode do vjerojatnije točne klasifikacije u taksonomsku skupinu više razine koja je zajednička višestrukim poravnanjima.

## 4.2. Taksonomski sastav bioreaktora

Do sada analizirani taksonomski sastavi bioreaktora koji koriste reakciju *anammox* identificirali su, uz *anammox* bakterije (razreda *Ca. Brocadia*), i raznolike druge skupine bakterija. Kao i ovdje, *anammox* bakterije često su jedini ili dominantni predstavnici koljena *Planctomycetes*, uz njih se uvijek javljaju i *Proteobacteria* i *Chloroflexi*, a gotovo uvijek *Bacteroidetes*, *Acidobacteria* i *Chlorobi* (Pereira i sur. 2017). Rod *Ca. Brocadia* jedini je identificirani rod *anammox* bakterija u analiziranom metagenomu, što je bilo i očekivano s obzirom da je to također bio jedini rod *anammox* bakterija identificiran u inokulumu analiziranoga bioreaktora (Mardanov i sur. 2016).

Koljeno *Proteobacteria* koje prevladava u ovome bioreaktoru čini jednu od najvećih skupina bakterija i dijeli se na preko 200 rodova heterogene morfologije i metabolizma (Gupta 2000). Među njima su i svi u uvodu spomenuti rodovi AOB, od kojih se među deset najzastupljenijih rodova ovdje nalazi samo *Nitrosomonas* razreda *Betaproteobacteria*. Bakterije roda *Nitrosomonas* stoga vjerojatno obavljaju većinu parcijalne nitracije u ovome bioreaktoru, što je konzistentno s nekim do sada analiziranim bioreaktorima koji koriste *anammox* reakciju (Guo i sur. 2016; Jia i sur. 2021; Wang i sur. 2016, 2019).

Međutim, najbrojniji rod *Proteobacteria* ovdje je *Thauera*, za čije je predstavnike poznato da mogu obavljati heterotrofnu denitrifikaciju. I rodovi *Proteobacteria Rubrivivax* i *Lautropia*, u ovom bioreaktoru među najzastupljenijima, obavljaju denitrifikaciju (Persson i sur. 2017).

Iako denitrifikacijske bakterije mogu kompetirati s *anammox* bakterijama za nitrit, za neke vrste roda *Thauera* (npr. ovdje najbrojniju *T. phenylacetatica*) poznato je da u prisustvu nitrata ne koriste nitrit kao krajnji akceptor elektrona u procesu staničnog disanja (Liu i sur. 2013). S obzirom da u bioreaktorima tipa PNA metabolizam AOB, *anammox* i drugih bakterija ipak proizvodi manje količine nitrata (Shourjeh i sur. 2021) kojega bakterije roda *Thaurea* reduciraju u nitrit pri staničnom disanju, moguće je da te bakterije dijelom pridonose uklanjanju nitrata i

akumulaciji nitrita potrebnog za *anammox* reakcije. U odsustvu nitrata, te bakterije ipak koriste nitrit kao krajnji akceptor elektrona i mogu kompetirati s *anammox* bakterijama za nitrit te svojim bržim rastom potencijalno ugroziti efikasnost bioreaktora i povećati produkciju sekundarnog mulja. Međutim, denitrificirajuće bakterije redovito se pronalaze u sastavu bioreaktora tipa PNA koji efikasno djeluju u različitim uvjetima, što bi moglo značiti da obavljaju važne uloge u ekologiji zajednice mikroorganizama bioreaktora PNA (Persson i sur. 2017).

*Chloroflexi* se uvijek pronalaze u značajnom broju u bioreaktorima koji koriste *anammox* reakciju. Njihova točna uloga u bioreaktorima nije posve razjašnjena. Smatra se da zajedno s drugim heterotrofima sudjeluju u degradaciji biomase rastućeg aktivnog mulja i da su dijelom odgovorni za nakupljanje mulja u granule ili flokule (Gonzalez-Martinez, Francisco Osorio, i sur. 2015). Također je poznato da mnoge skupine mogu obavljati disimilatornu redukciju nitrita u amonijak, ometajući uklanjanje amonijaka iz otpadne vode (Guo i sur. 2016). Unatoč tome što se smatra da su *Chloroflexi* jedna od skupina koje čine jezgru zajednice mikroorganizama u bioreaktorima koji koriste *anammox* reakciju, zajedno s drugim heterotrofima vjerojatno također predstavljaju rizik za stabilnu funkcionalnost bioreaktora industrijske veličine gdje je unošena otpadna voda varijabilnija po sastavu, posebno po količini organskog ugljika koji potiče rast heterotrofa (Keren i sur. 2020). Stoga je kod takvih bioreaktora preporučljivo redovito praćenje i stroža kontrola udjela *Chloroflexi* i drugih heterotrofnih bakterija.

Rod NOB *Nitrospira* također je značajno zastupljen u ovome bioreaktoru. Održavanje količine NOB u bioreaktorima tipa PNA relativno niskom u odnosu na AOB i *anammox* bakterije je važno za djelovanje takvih bioreaktora radi sprječavanja akumulacije nitrata; NOB su u direktnoj kompeticiji s AOB za kisik i s *anammox* bakterijama za nitrit koji *anammox* bakterijama služi kao akceptor, a NOB kao donor elektrona (Shourjeh i sur. 2021). Međutim, s obzirom na prisustvo denitrificirajućih bakterija u ovom bioreaktoru koje preferencijalno koriste nitrat kao akceptor elektrona, moguće je da te skupine zajedno ne pridonose manjku nitrita koji je potreban *anammox* bakterijama za uklanjanje amonijaka iz otpadne vode.

### 4.3. Nejednaka uporaba kodona i funkcionalna analiza metagenoma

Struktura prosječne uporabe kodona čitavog metagenoma pokazuje da su za svaku aminokiselinu gotovo svi kodoni kojima je treća baza G ili C češće korišteni od svih kojima je treća baza A ili T (uz samo tri slabije korištene iznimke: TTG, AGG i GGG). Taj efekt se vjerojatno može dijelom objasniti visokim udjelom GC baza od 60,35% u predviđenim genima. Visok udio GC baza kod bakterija povezan je s višom optimalnom temperaturom rasta (Musto i sur. 2006), ali i preferencijalnim supstitucijama u G ili C pri rekombinaciji koje ne donose nužno povišenje fitnesa bakterije (Lassalle i sur. 2015). S druge strane, kodoni obogaćeni kod ribosoma, za koje se pretpostavlja da su na neki način optimizirani za transkripciju ili translaciju, slabije prate to pravilo u odnosu na prosjek ostalih predviđenih gena i češće imaju na svakoj poziciji baze A ili T, a pogotovo na trećoj poziciji. Preferencija za A ili T na trećoj poziciji kodona, posebno kod aminokiselina kodiranih s više od dva kodona, prethodno je povezana s većom efikasnosti translacije i brzinom produkcije proteina (Ran i Higgs 2010) potvrđujući valjanost analize nejednake uporabe kodona kao metode procjene ekspresije gena.

Neobično visoka relativna ekspresivnost predviđena je za gene svrstane u module biosinteza sfingozina i biosinteza ceramida s obzirom da je prisustvo sfingolipida dokazano u malom broju skupina bakterija (Harrison, Dunn, i Campopiano 2018). Među tim skupinama nalaze se i neke *Bacteroidetes* čiji su srodnici identificirani u ovome bioreaktoru, iako su njihovi sfingolipidi istraživani samo u kontekstu crijevne mikrobiote gdje se pretpostavlja da igraju ulogu u interakciji bakterije s imunološkim sustavom domaćina i u staničnoj adheziji (An i sur. 2014). Uloga sfingolipida kod zajednica bakterija bioreaktora za obradu otpadnih voda neistražena je i potencijalno važna za međustaničnu komunikaciju bakterija ili prijanjanje za stanične agregate (granule, flokule ili biofilm).

Visoka je relativna ekspresivnost predviđena i za module vezane uz nitritaciju, *anammox* i asimilaciju dušika (modul koji uključuje uglavnom transportne proteine za nitrat i nitrit, [https://www.genome.jp/module/vgo\\_M00615](https://www.genome.jp/module/vgo_M00615)). Geni vezani uz denitrifikaciju također su iznadprosječne procijenjene relativne ekspresivnosti, ali u manjoj mjeri nego oni vezani uz nitritaciju, *anammox* i asimilaciju dušika. S obzirom na visok broj klasificiranih denitrificirajućih bakterija, može se očekivati da je razina denitrifikacije u bioreaktoru visoka. Predložene

strategije ispiranja heterotrofnih bakterija, uključujući i denitrificirajuće bakterije i NOB, vezane su uz imobilizaciju željenih AOB i *anammox* bakterija u obliku granula ili biofilma (Liu i sur. 2017; Park i sur. 2015), ali i uz sastav otpadnih voda za kojega je poželjno da ima manji udio organskog ugljika koji omogućuje rast heterotrofa (Zhang i sur. 2019).

Geni vezani uz asimilatornu redukciju nitrata, tj. prevođenje nitrata u amonijak koji se ugrađuje u biomolekule, slabe su procijenjene relativne ekspresivnosti. Enzimi ovog modula pronađeni su kod heterotrofnih bakterija uključujući denitrificirajuće (Luque-Almagro i sur. 2011), dakle u ovom bioreaktoru vjerojatno je donekle uspješno suprimirano prevođenje amonijaka u biomasu aktivnog mulja umjesto u molekularni dušik, što je jedan od ciljeva bioreaktora PNA radi efikasnijeg prevođenja organskog ugljika prisutnog u otpadnoj vodi u bioplin (metan) (Siegrist i sur. 2008).

Značajno je manja od prosjeka relativna ekspresivnost mnogih modula vezanih uz biosintezu i degradaciju masnih kiselina. Uzrok tome je nepoznat, no moguće je da je smanjena ekspresivnost modula vezanih uz degradaciju masnih kiselina povezana s činjenicom da se masti često uklanjaju iz otpadnih voda prije ulaza u bioreaktor, tijekom primarne obrade otpadne vode.

Među izdvojenim pojedinim genima vezanim uz metabolizam dušika, daleko najbrojniji pronađen je gen za podjedinicu  $\alpha$  respiratorne nitrat reduktaze (*narG*), enzima denitrifikacije, sa 160 predviđenih gena klasificiranih u tu ortolognu skupinu (iduća po brojnosti je ortologna skupina podjedinice  $\beta$  respiratorne nitrat reduktaze, *narH*, s 62 predviđena gena). Međutim, jedini geni čija je vrijednost relativne ekspresivnosti statistički značajna ( $p < 0,05$ ) su hidrazin dehidrogenaza (*hdh*, 41 među predviđenim genima) s procijenjenom vrlo visokom relativnom ekspresivnosti, te podjedinica D citokrom c-552 nitrit reduktaze (*nrfD*, 7 pronađenih), podjedinica D NADH-nitrit reduktaze (*nirD*, 7 pronađenih), i podjedinica A asimilatorne nitrat reduktaze (*nasA*, 15 pronađenih) s procijenjenom niskom relativnom ekspresivnosti. Produkt gena *hdh* oksidira hidrazin ( $N_2H_4$ ) s molekularnim dušikom kao produktom i krajnji je enzim *anammox* puta. Njegova visoka relativna ekspresivnost zajedno s visokom brojnošću tih gena u bioreaktoru govori da je *anammox* put vrlo aktivan i da se vjerojatno, unatoč prisustvu mnogih denitrificirajućih bakterija i NOB, značajan dio amonijaka otpadne vode prevodi u molekularni dušik autotrofnim putem. Dvije skupine nitrit reduktaza *nirD* i *nasA* vezane su uz asimilatornu redukciju nitrata (<https://www.genome.jp/module/M00531+K00372>) i, kao i pripadajući modul,



niske su relativne ekspresivnosti, ponovno potvrđujući efikasnost autotrofnog uklanjanja dušika iz otpadne vode u ovome bioreaktoru.

## 5. ZAKLJUČAK

U ovom radu analiziran je taksonomski sastav i funkcionalni potencijal zajednice mikroorganizama u bioreaktoru sekvencijalno serijalnog tipa laboratorijske veličine s djelomičnom reakcijom nitritacije i anaerobne oksidacije amonijaka (PNA) u kojemu se uklanjanje dušika iz sintetičke otpadne vode izvršava istovremenim djelovanjem AOB i *anammox* bakterija. Pokazano je da su u bioreaktoru prisutne mnoge skupine heterotrofnih bakterija za koje se inače smatra da mogu omesti ravnotežu procesa PNA, no da su enzimi metaboličkih puteva nitritacije i *anammox* reakcija vjerojatno visoke ekspresivnosti unatoč njihovom prisustvu. Buduća istraživanja trebala bi analizirati bioreaktore ovog tipa industrijske veličine koji djeluju pod različitim i varijabilnim uvjetima kakvi su mogući kod obrade stvarnih otpadnih voda.

## 6. LITERATURA

- Agrawal, Shelesh, Dries Seuntjens, Pieter De Cocker, Susanne Lackner, i Siegfried E. Vlaeminck. 2018. „Success of mainstream partial nitrification/anammox demands integration of engineering, microbiome and modeling insights“. *Current Opinion in Biotechnology* 50:214–21. doi: 10.1016/J.COPBIO.2018.01.013.
- Altschul, Stephen F., Warren Gish, Webb Miller, Eugene W. Myers, i David J. Lipman. 1990. „Basic local alignment search tool“. *Journal of Molecular Biology* 215(3):403–10. doi: 10.1016/S0022-2836(05)80360-2.
- Amann, R. I., W. Ludwig, i K. H. Schleifer. 1995. „Phylogenetic identification and in situ detection of individual microbial cells without cultivation“. *Microbiological Reviews* 59(1):143–69. doi: 10.1128/MR.59.1.143-169.1995.
- An, Dingding, Sungwhan F. Oh, Torsten Olszak, Joana F. Neves, Fikri Y. Avci, Deniz Erturk-Hasdemir, Xi Lu, Sebastian Zeissig, Richard S. Blumberg, i Dennis L. Kasper. 2014. „Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells“. *Cell* 156(0):123. doi: 10.1016/J.CELL.2013.11.042.
- Anders, Simon. 2010. „Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data“. *Soil* 5(1). Preuzeto 19. kolovoz 2021. (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).
- „KEGG MODULE: M00531“. Preuzeto 21. ožujak 2022.a. (<https://www.genome.jp/module/M00531+K00372>).
- „KEGG MODULE: M00615“. Preuzeto 21. ožujak 2022.b. ([https://www.genome.jp/module/vgo\\_M00615](https://www.genome.jp/module/vgo_M00615)).
- Boeije, Geert, Ronald Corstanje, André Rottiers, i Diederik Schowanek. 1999. „Adaptation of the CAS test system and synthetic sewage for biological nutrient removal: Part I: Development of a new synthetic sewage“. *Chemosphere* 38(4):699–709. doi: 10.1016/S0045-6535(98)00311-7.

- Bolger, Anthony M., Marc Lohse, i Bjoern Usadel. 2014. „Trimmomatic: A flexible trimmer for Illumina sequence data“. *Bioinformatics* 30(15):2114–20. doi: 10.1093/bioinformatics/btu170.
- Broda, E. 1977. „Two kinds of lithotrophs missing in nature“. *Zeitschrift für allgemeine Mikrobiologie* 17(6):491–93. doi: 10.1002/jobm.3630170611.
- Buchfink, Benjamin, Klaus Reuter, i Hajk-Georg Drost. 2021. „Sensitive protein alignments at tree-of-life scale using DIAMOND“. *Nature Methods* 2021 18:4 18(4):366–68. doi: 10.1038/s41592-021-01101-x.
- Cao, Yeshe, Mark C. M. van Loosdrecht, i Glen T. Daigger. 2017. „Mainstream partial nitrification–anammox in municipal wastewater treatment: status, bottlenecks, and further studies“. *Applied Microbiology and Biotechnology* 101(4):1365–83. doi: 10.1007/s00253-016-8058-7.
- Carpenter, Stephen R. 2005. „Eutrophication of aquatic ecosystems: Bistability and soil phosphorus“. *Proceedings of the National Academy of Sciences of the United States of America* 102(29):10002–5. doi: 10.1073/pnas.0503959102.
- Carstensen, Jacob, Jesper H. Andersen, Bo G. Gustafsson, i Daniel J. Conley. 2014. „Deoxygenation of the Baltic Sea during the last century“. *Proceedings of the National Academy of Sciences of the United States of America* 111(15):5628. doi: 10.1073/PNAS.1323156111.
- Chu, Zhao rui, Ke Wang, Xiang kun Li, Meng ting Zhu, Liu Yang, i Jie Zhang. 2015. „Microbial characterization of aggregates within a one-stage nitrification-anammox system using high-throughput amplicon sequencing“. *Chemical Engineering Journal* 262:41–48. doi: 10.1016/j.cej.2014.09.067.
- Cock, Peter J. A., Christopher J. Fields, Naohisa Goto, Michael L. Heuer, i Peter M. Rice. 2009. „The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants“. *Nucleic Acids Research* 38(6):1767–71. doi: 10.1093/nar/gkp1137.
- Dalsgaard, Tage, Donald E. Canfield, Jan Petersen, Bo Thamdrup, i Jenaro Acuña-González. 2003. „N<sub>2</sub> production by the anammox reaction in the anoxic water column of Golfo Dulce,

- Costa Rica“. *Nature* 422(6932):606–8. doi: 10.1038/nature01526.
- Darzi, Youssef, Ivica Letunic, Peer Bork, i Takuji Yamada. 2018. „IPath3.0: Interactive pathways explorer v3“. *Nucleic Acids Research* 46(W1):W510–13. doi: 10.1093/nar/gky299.
- Dohm, Juliane C., Claudio Lottaz, Tatiana Borodina, i Heinz Himmelbauer. 2008. „Substantial biases in ultra-short read data sets from high-throughput DNA sequencing“. *Nucleic Acids Research* 36(16). doi: 10.1093/NAR/GKN425.
- Van Dongen, U., M. S. M. Jetten, i M. C. M. Van Loosdrecht. 2001. „The SHARON®-Anammox® process for treatment of ammonium rich wastewater“. *Water Science and Technology* 44(1):153–60. doi: 10.2166/wst.2001.0037.
- Anamaria Elek, Maja Kuzman i Kristian Vlahoviček. 2020. "coRdon: Codon Usage Analysis and Prediction of Gene Expressivity". R package version 1.8.0. <https://github.com/BioinfoHR/coRdon>
- Environmental Protection Department of the Hong Kong Special Administrative Region. bez dat. „Energy Recovery“. Preuzeto 17. listopad 2021. ([https://www.epd.gov.hk/epd/english/environmentinhk/waste/prob\\_solutions/Energy.html](https://www.epd.gov.hk/epd/english/environmentinhk/waste/prob_solutions/Energy.html)).
- Ewing, Brent, i Phil Green. 1998. „Base-calling of automated sequencer traces using phred. II. Error probabilities“. *Genome Research* 8(3):186–94. doi: 10.1101/gr.8.3.186.
- Ewing, Brent, La Deana Hillier, Michael C. Wendl, i Phil Green. 1998. „Base-calling of automated sequencer traces using phred. I. Accuracy assessment“. *Genome Research* 8(3):175–85. doi: 10.1101/gr.8.3.175.
- Fang, Gang, Nitin Bhardwaj, Rebecca Robilotto, i Mark B. Gerstein. 2010. „Getting Started in Gene Orthology and Functional Analysis“. *PLoS Computational Biology* 6(3). doi: 10.1371/JOURNAL.PCBI.1000703.
- Gerba, Charles P., i Ian L. Pepper. 2009. „Wastewater Treatment and Biosolids Reuse“. *Environmental Microbiology* 503. doi: 10.1016/B978-0-12-370519-8.00024-9.
- Gonzalez-Martinez, Alejandro, F. Osorio, Alejandro Rodriguez-Sanchez, Maria Victoria Martinez-Toledo, Jesus Gonzalez-Lopez, Tommaso Lotti, i M. C. M. van Loosdrecht. 2015.

- „Bacterial community structure of a lab-scale anammox membrane bioreactor“. *Biotechnology Progress* 31(1):186–93. doi: 10.1002/btpr.1995.
- Gonzalez-Martinez, Alejandro, Francisco Osorio, Jose A. Morillo, Alejandro Rodriguez-Sanchez, Jesus Gonzalez-Lopez, Ben A. Abbas, i Mark C. M. van Loosdrecht. 2015. „Comparison of bacterial diversity in full scale anammox bioreactors operated under different conditions“. *Biotechnology Progress* 31(6):1464–72. doi: 10.1002/BTPR.2151.
- Guo, Jianhua, Yongzhen Peng, Lu Fan, Liang Zhang, Bing Jie Ni, Boran Kartal, Xin Feng, Mike S. M. Jetten, i Zhiguo Yuan. 2016. „Metagenomic analysis of anammox communities in three different microbial aggregates“. *Environmental Microbiology* 18(9):2979–93. doi: 10.1111/1462-2920.13132.
- Gupta, Radhey S. 2000. „The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes“. *FEMS Microbiology Reviews* 24(4):367–402. doi: 10.1111/J.1574-6976.2000.TB00547.X.
- Harrison, Peter J., Teresam Dunn, i Dominic J. Campopiano. 2018. „Sphingolipid biosynthesis in man and microbes“. *Natural Product Reports* 35(9):921. doi: 10.1039/C8NP00019K.
- Heisler, J., P. Glibert, J. Burkholder, D. Anderson, W. Cochlan, W. Dennison, C. Gobler, Q. Dortch, C. Heil, E. Humphries, A. Lewitus, R. Magnien, H. Marshall, K. Sellner, D. Stockwell, D. Stoecker, i M. Suddleson. 2008. „Eutrophication and Harmful Algal Blooms: A Scientific Consensus“. *Harmful Algae* 8(1):3. doi: 10.1016/J.HAL.2008.08.006.
- Hugenholtz, Philip, Brett M. Goebel, i Norman R. Pace. 1998. „Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity“. *Journal of Bacteriology* 180(18):4765. doi: 10.1128/jb.180.18.4765-4774.1998.
- Huson, Daniel H., Alexander F. Auch, Ji Qi, i Stephan C. Schuster. 2007. „MEGAN analysis of metagenomic data“. *Genome Research* 17(3):377. doi: 10.1101/GR.5969107.
- Illumina. 2015. „Illumina | Sequencing and array-based solutions for genetic research“. Preuzeto 31. listopad 2021. (<http://www.illumina.com/>).
- Jensen, Lars Juhl, Philippe Julien, Michael Kuhn, Christian von Mering, Jean Muller, Tobias Doerks, i Peer Bork. 2008. „eggNOG: automated construction and annotation of orthologous

- groups of genes“. *Nucleic acids research* 36(Database issue). doi: 10.1093/NAR/GKM796.
- Jetten, Mike S. M. 2008. „The microbial nitrogen cycle“. *Environmental Microbiology* 10(11):2903–9. doi: 10.1111/j.1462-2920.2008.01786.x.
- Jia, Fangxu, Yongzhen Peng, Jianwei Li, Xiyao Li, i Hong Yao. 2021. „Metagenomic prediction analysis of microbial aggregation in anammox-dominated community“. *Water Environment Research* 93(11):2549–58. doi: 10.1002/WER.1529.
- Jorgensen, S. E., i B. Halling-Sorensen. 1993. „Biological Nitrification and Denitrification“. Str. 43–53 u *Studies in Environmental Science, Volume 54*.
- Kanehisa, Minoru. 2018. „KEGG MODULE Database“. *Kanehisa Laboratories*. Preuzeto 19. veljača 2022. (<https://www.kegg.jp/kegg/module.html>).
- Kartal, Boran, Wim Geerts, i Mike S. M. Jetten. 2011. „Cultivation, Detection, and Ecophysiology of Anaerobic Ammonium-Oxidizing Bacteria“. *Methods in Enzymology* 486(C):89–108. doi: 10.1016/B978-0-12-381294-0.00004-3.
- Keren, Ray, Jennifer E. Lawrence, Weiqin Zhuang, David Jenkins, Jillian F. Banfield, Lisa Alvarez-Cohen, Lijie Zhou, i Ke Yu. 2020. „Increased replication of dissimilatory nitrate-reducing bacteria leads to decreased anammox bioreactor performance“. *Microbiome* 8(1):1–21. doi: 10.1186/S40168-020-0786-3/FIGURES/9.
- Kindaichi, Tomonori, Shota Yuri, Noriatsu Ozaki, i Akiyoshi Ohashi. 2012. „Ecophysiological role and function of uncultured Chloroflexi in an anammox reactor“. *Water Science and Technology* 66(12):2556–61. doi: 10.2166/wst.2012.479.
- Köster, Friedrich W., Hans-Harald Hinrichsen, Michael A. St. John, Dietrich Schnack, Brian R. MacKenzie, Jonna Tomkiewicz, i Maris Plikshs. 2001. „Developing Baltic cod recruitment models. II. Incorporation of environmental variability and species interaction“. *Canadian Journal of Fisheries and Aquatic Sciences* 58(8):1534–56. doi: 10.1139/CJFAS-58-8-1534.
- Kuenen, J. Gijss. 2008. „Anammox bacteria: from discovery to application“. *Nature Reviews Microbiology* 2008 6:4 6(4):320–26. doi: 10.1038/nrmicro1857.
- Kunin, Victor, Alex Copeland, Alla Lapidus, Konstantinos Mavromatis, i Philip Hugenholtz.

2008. „A Bioinformatician’s Guide to Metagenomics“. *Microbiology and Molecular Biology Reviews* 72(4):557–78. doi: 10.1128/MMBR.00009-08/ASSET/02282DF5-A19B-4A48-8246-C820E0A0FABA/ASSETS/GRAPHIC/ZMR0040821930007.JPEG.
- Kuypers, Marcel M. M., A. Olav Silekers, Gaute Lavik, Markus Schmid, Bo Barker Jørgensen, J. Gijs Kuenen, Jaap S. Sinninghe Damsté, Marc Strous, i Mike S. M. Jetten. 2003. „Anaerobic ammonium oxidation by anammox bacteria in the Black Sea“. *Nature* 422(6932):608–11. doi: 10.1038/nature01472.
- Langmead, Ben, i Steven L. Salzberg. 2012. „Fast gapped-read alignment with Bowtie 2“. *Nature Methods* 9(4):357–59. doi: 10.1038/nmeth.1923.
- Lassalle, Florent, Séverine Périan, Thomas Bataillon, Xavier Nesme, Laurent Duret, i Vincent Daubin. 2015. „GC-Content Evolution in Bacterial Genomes: The Biased Gene Conversion Hypothesis Expands“. *PLOS Genetics* 11(2):e1004941. doi: 10.1371/JOURNAL.PGEN.1004941.
- Lawrence, Michael, Wolfgang Huber, Hervé Pagès, Patrick Aboyoun, Marc Carlson, Robert Gentleman, Martin T. Morgan, i Vincent J. Carey. 2013. „Software for Computing and Annotating Genomic Ranges“. *PLoS Computational Biology* 9(8):e1003118. doi: 10.1371/journal.pcbi.1003118.
- Li, Xiao Ran, Bing Du, Hai Xia Fu, Ruo Fan Wang, Jin Huan Shi, Yue Wang, Mike S. M. Jetten, i Zhe Xue Quan. 2009. „The bacterial diversity in an anaerobic ammonium-oxidizing (anammox) reactor community“. *Systematic and Applied Microbiology* 32(4):278–89. doi: 10.1016/j.syapm.2009.03.002.
- Liu, Binbin, Yuejian Mao, Linda Bergaust, Lars R. Bakken, i Åsa Frostegård. 2013. „Strains in the genus *Thauera* exhibit remarkably different denitrification regulatory phenotypes“. *Environmental Microbiology* 15(10):2816–28. doi: 10.1111/1462-2920.12142.
- Liu, Yuan, Qigui Niu, Shaopo Wang, Jiayuan Ji, Yu Zhang, Min Yang, Toshimasa Hojo, i Yu You Li. 2017. „Upgrading of the symbiosis of *Nitrosomanas* and anammox bacteria in a novel single-stage partial nitrification–anammox system: Nitrogen removal potential and Microbial characterization“. *Bioresour. Technol.* 244:463–72. doi: 10.1016/J.BIORTECH.2017.07.156.



- Lopez-Ridaura, Santiago, Hayo van der Werf, Jean Marie Paillat, i Bertrand Le Bris. 2009. „Environmental evaluation of transfer and treatment of excess pig slurry by life cycle assessment“. *Journal of Environmental Management* 90(2):1296–1304. doi: 10.1016/j.jenvman.2008.07.008.
- Lotti, T., R. Kleerebezem, J. M. Abelleira-Pereira, B. Abbas, i M. C. M. van Loosdrecht. 2015. „Faster through training: The anammox case“. *Water Research* 81:261–68. doi: 10.1016/J.WATRES.2015.06.001.
- Luque-Almagro, Víctor M., Andrew J. Gates, Conrado Moreno-Vivián, Stuart J. Ferguson, David J. Richardson, i M. Dolores Roldán. 2011. „Bacterial nitrate assimilation: gene distribution and regulation“. *Biochemical Society Transactions* 39(6):1838–43. doi: 10.1042/BST20110688.
- Ma, Bin, Wenting Qian, Chuansheng Yuan, Zhiguo Yuan, i Yongzhen Peng. 2017. „Achieving Mainstream Nitrogen Removal through Coupling Anammox with Denitratation“. *Environmental Science and Technology* 51(15):8405–13. doi: 10.1021/acs.est.7b01866.
- Mancosu, Gianmaria, Massimiliano Cosso, Francesca Marras, Cesare Cappio Borlino, Giuseppe Ledda, Teresa Manias, Mauro Adamo, Donatella Serra, Paola Melis, i Mario Pirastu. 2005. „Browsing isolated population data“. *BMC Bioinformatics* 6(SUPPL.4). doi: 10.1186/1471-2105-6-S1-S17.
- Mardanov, A. V., A. V. Beletskii, A. Yu. Kallistova, R. Yu. Kotlyarov, Yu. A. Nikolaev, M. V. Kevbrina, A. M. Agarev, N. V. Ravin, i N. V. Pimenov. 2016. „Dynamics of the composition of a microbial consortium during start-up of a single-stage constant flow laboratory nitritation/anammox setup“. *Microbiology 2016* 85:6 85(6):681–92. doi: 10.1134/S002626171606014X.
- Mardanov, Andrey V., Roman V. Kotlyarov, Alexey V. Beletsky, Yury A. Nikolaev, Anna Yu Kallistova, Vladimir A. Grachev, Yulia Yu Berestovskaya, Nikolai V. Pimenov, i Nikolai V. Ravin. 2019. „Metagenomic data of the microbial community of lab-scale nitritation-anammox sequencing-batch bioreactor performing nitrogen removal from synthetic wastewater“. *Data in Brief* 27. doi: 10.1016/j.dib.2019.104722.
- Mardanov, Andrey V, Roman V Kotlyarov, Alexey V Beletsky, Nikolai V Pimenov, Yury A.

- Nikolaev, Vladimir A. Grachev, i Yulia Yu. 2019. „Metagenomic data of the microbial community of lab-scale nitrification-anammox sequencing- batch bioreactor performing nitrogen removal from synthetic wastewater“. *Data in Brief* (November):104722. doi: 10.1016/j.dib.2019.104722.
- Microsoft Corporation. bez dat. „Softver za proračunske tablice Microsoft Excel“. Preuzeto 14. veljača 2022. (<https://www.microsoft.com/hr-hr/microsoft-365/excel>).
- Mir, Katharina, Klaus Neuhaus, Siegfried Scherer, Martin Bossert, i Steffen Schober. 2012. „Predicting Statistical Properties of Open Reading Frames in Bacterial Genomes“. *PLOS ONE* 7(9):e45103. doi: 10.1371/JOURNAL.PONE.0045103.
- Mulder, A., A. A. Van De Graaf, L. A. Robertson, i J. G. Kuenen. 1995. „Anaerobic ammonium oxidation discovered in a denitrifying fluidized bed reactor“. *MICROBIOLOGY ECOLOGY FEMS Microbiology Ecology* 16:177–84.
- Musto, Héctor, Hugo Naya, Alejandro Zavala, Héctor Romero, Fernando Alvarez-Valín, i Giorgio Bernardi. 2006. „Genomic GC level, optimal growth temperature, and genome size in prokaryotes“. *Biochemical and biophysical research communications* 347(1):1–3. doi: 10.1016/J.BBRC.2006.06.054.
- NCBI/SRA. 2021. „Home - SRA - NCBI“. Preuzeto 19. kolovoz 2021. (<https://www.ncbi.nlm.nih.gov/sra>).
- NCBI. 2019a. „GRCh38 - hg38 - Genome - Assembly - NCBI“. Preuzeto 01. studeni 2021. ([https://www.ncbi.nlm.nih.gov/assembly/GCF\\_000001405.26/](https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/)).
- NCBI. 2019b. „National Center for Biotechnology Information“. 1723–1723. Preuzeto 01. studeni 2021. (<https://www.ncbi.nlm.nih.gov/>).
- Ni, Bing Jie, Bao Lan Hu, Fang Fang, Wen Ming Xie, Boran Kartal, Xian Wei Liu, Guo Ping Sheng, Mike Jetten, Ping Zheng, i Han Qing Yu. 2010. „Microbial and physicochemical characteristics of compact anaerobic ammonium-oxidizing granules in an upflow anaerobic sludge blanket reactor“. *Applied and Environmental Microbiology* 76(8):2652–56. doi: 10.1128/AEM.02271-09.
- Nurk, Sergey, Dmitry Meleshko, Anton Korobeynikov, i Pavel A. Pevzner. 2017. „MetaSPAdes:

- A new versatile metagenomic assembler“. *Genome Research* 27(5):824–34. doi: 10.1101/gr.213959.116.
- Overbeek, Ross, Robert Olson, Gordon D. Pusch, Gary J. Olsen, James J. Davis, Terry Disz, Robert A. Edwards, Svetlana Gerdes, Bruce Parrello, Maulik Shukla, Veronika Vonstein, Alice R. Wattam, Fangfang Xia, i Rick Stevens. 2014. „The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST)“. *Nucleic Acids Research* 42(Database issue). doi: 10.1093/NAR/GKT1226.
- Pagés, H., P. Aboyou, R. Gentleman, i S. DebRoy. 2020. „Biostrings: Efficient manipulation of biological strings“. *R package version 2.58.0*.
- Park, Hongkeun, Alex Rosenthal, Krish Ramalingam, John Fillos, i Kartik Chandran. 2010. „Linking community profiles, gene expression and n-removal in anammox bioreactors treating municipal anaerobic digestion reject water“. *Environmental Science and Technology* 44(16):6110–16. doi: 10.1021/es1002956.
- Park, Hongkeun, Suneethi Sundar, Yiwei Ma, i Kartik Chandran. 2015. „Differentiation in the microbial ecology and activity of suspended and attached bacteria in a nitrification-anammox process“. *Biotechnology and Bioengineering* 112(2):272–79. doi: 10.1002/BIT.25354.
- Pereira, Alyne Duarte, Angela Cabezas, Claudia Etchebehere, Carlos Augusto de Lemos Chernicharo, i Juliana Calábria de Araújo. 2017. „Microbial communities in anammox reactors: a review“. *Environmental Technology Reviews* 6(1):74–93. doi: 10.1080/21622515.2017.1304457.
- Pérez-Cobas, Ana Elena, Laura Gomez-Valero, i Carmen Buchrieser. 2020. „Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses“. *Microbial Genomics* 6(8):1–22. doi: 10.1099/MGEN.0.000409.
- Persson, Frank, Carolina Suarez, Malte Hermansson, Elzbieta Plaza, Razia Sultana, i Britt Marie Wilén. 2017. „Community structure of partial nitrification-anammox biofilms at decreasing substrate concentrations and low temperature“. *Microbial Biotechnology* 10(4):761–72. doi: 10.1111/1751-7915.12435.
- Rajta, A., R. Bhatia, H. Setia, i P. Pathania. 2020. „Role of heterotrophic aerobic denitrifying

- bacteria in nitrate removal from wastewater“. *Journal of Applied Microbiology* 128(5):1261–78. doi: 10.1111/JAM.14476.
- Ran, Wenqi, i Paul G. Higgs. 2010. „The influence of anticodon-codon interactions and modified bases on codon usage bias in bacteria“. *Molecular Biology and Evolution* 27(9):2129–40. doi: 10.1093/MOLBEV/MSQ102.
- Roller, Maša, Vedran Lucić, István Nagy, Tina Perica, i Kristian Vlahoviček. 2013. „Environmental shaping of codon usage and functional adaptation across microbial communities“. *Nucleic Acids Research* 41(19):8842–52. doi: 10.1093/nar/gkt673.
- Samocha, Tzachi M., i David I. Prangnell. 2019. „System treatment and preparation“. Str. 119–31 u *Sustainable Biofloc Systems for Marine Shrimp*. Academic Press.
- Schmidt, Ingo, Olav Sliemers, Markus Schmid, Eberhard Bock, John Fuerst, J. Gijs Kuenen, Mike S. M. Jetten, i Marc Strous. 2003. „New concepts of microbial treatment processes for the nitrogen removal in wastewater“. *FEMS Microbiology Reviews* 27(4):481–92. doi: 10.1016/S0168-6445(03)00039-1.
- Sharp, Paul M., i Wen-Hsiung Li. 1987. „The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications“. *Nucleic Acids Research* 15(3):1281–95. doi: 10.1093/NAR/15.3.1281.
- Shourjeh, Mehdi Sharif, Przemysław Kowal, Xi Lu, Li Xie, i Jakub Drewnowski. 2021. „Development of strategies for AOB and NOB competition supported by mathematical modeling in terms of successful deammonification implementation for energy-efficient WWTPs“. *Processes* 9(3). doi: 10.3390/PR9030562.
- Siegrist, H., D. Salzgeber, J. Eugster, i A. Joss. 2008. „Anammox brings WWTP closer to energy autarky due to increased biogas production and reduced aeration energy for N-removal“. *Water Science and Technology* 57(3):383–88. doi: 10.2166/wst.2008.048.
- Sonthiphand, Puntipar, Michael W. Hall, i Josh D. Neufeld. 2014. „Biogeography of anaerobic ammonia-oxidizing (anammox) bacteria“. *Frontiers in Microbiology* 5(AUG):399. doi: 10.3389/FMICB.2014.00399/ABSTRACT.
- Speth, Daan R., Michiel H. in ’t Zandt, Simon Guerrero-Cruz, Bas E. Dutilh, i Mike S. M. Jetten.

2016. „Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system“. *Nature Communications* 2016 7:1 7(1):1–10. doi: 10.1038/ncomms11172.
- van der Star, Wouter R. L., Wiebe R. Abma, Dennis Blommers, Jan Willem Mulder, Takaaki Tokutomi, Marc Strous, Cristian Picioreanu, i Mark C. M. van Loosdrecht. 2007. „Startup of reactors for anoxic ammonium oxidation: Experiences from the first full-scale anammox reactor in Rotterdam“. *Water Research* 41(18):4149–63. doi: 10.1016/j.watres.2007.03.044.
- Van Der Star, Wouter R. L., Andreea I. Miclea, Udo G. J. M. Van Dongen, Gerard Muyzer, Cristian Picioreanu, i Mark C. M. Van Loosdrecht. 2008. „The membrane bioreactor: A novel tool to grow anammox bacteria as free cells“. *Biotechnology and Bioengineering* 101(2):286–94. doi: 10.1002/BIT.21891.
- Strous, M., J. J. Heijnen, J. G. Kuenen, i M. S. M. Jetten. 1998. „The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms“. *Applied Microbiology and Biotechnology* 50(5):589–96. doi: 10.1007/s002530051340.
- Strous, Marc, John A. Fuerst, Evelien H. M. Kramer, Susanne Logemann, Gerard Muyzer, Katinka T. van de Pas-Schoonen, Richard Webb, J. Gijs Kuenen, i Mike S. M. Jetten. 1999. „Missing lithotroph identified as new planctomycete“. *Nature* 1999 400:6743 400(6743):446–49. doi: 10.1038/22749.
- Supek, Frank, i Kristian Vlahoviček. 2005. „Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity“. *BMC Bioinformatics* 6. doi: 10.1186/1471-2105-6-182.
- Tan, Ge, Lennart Opitz, Ralph Schlapbach, i Hubert Rehrauer. 2019. „Long fragments achieve lower base quality in Illumina paired-end sequencing“. *Scientific Reports* 9(1). doi: 10.1038/S41598-019-39076-7.
- The European Union. 1991. „Council Directive of 21 May 1991 concerning urban waste water treatment“. *Official Journal of the European Communities* L 135:40–52. Preuzeto 16. listopad 2021. (<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31991L0271>).

- United States Environmental Protection Agency. 2021. „Preliminary Effluent Guidelines Program Plan 15; September 2021“.
- Vahtera, Emil, Daniel J. Conley, Bo G. Gustafsson, Harri Kuosa, Heikki Pitkänen, Oleg P. Savchuk, Timo Tamminen, Markku Viitasalo, Maren Voss, Norbert Wasmund, i Fredrik Wulff. 2007. „Internal Ecosystem Feedbacks Enhance Nitrogen-fixing Cyanobacteria Blooms and Complicate Management in the Baltic Sea“. *A Journal of the Human Environment* 36(2):186–94. doi: 10.1579/0044-7447(2007)36[186:IEFENC]2.0.CO;2.
- Vlaeminck, Siegfried E., Akihiko Terada, Barth F. Smets, Haydée De Clippeleir, Thomas Schaubroeck, Selin Bolea, Lien Demeestere, Jan Mast, Nico Boon, Marta Carballa, i Willy Verstraete. 2010. „Aggregate size and architecture determine microbial activity balance for one-stage partial nitrification and anammox“. *Applied and Environmental Microbiology* 76(3):900–909. doi: 10.1128/AEM.02337-09.
- van der Walt, Andries Johannes, Marc Warwick van Goethem, Jean Baptiste Ramond, Thulani Peter Makhalanyane, Oleg Reva, i Don Arthur Cowan. 2017. „Assembling metagenomes, one community at a time“. *BMC Genomics* 18(1). doi: 10.1186/S12864-017-3918-9.
- Wang, Bo, Yongzhen Peng, Yuanyuan Guo, Mengyue Zhao, i Shuying Wang. 2016. „Illumina MiSeq sequencing reveals the key microorganisms involved in partial nitrification followed by simultaneous sludge fermentation, denitrification and anammox process“. *Bioresource Technology* 207:118–25. doi: 10.1016/J.BIORTECH.2016.01.072.
- Wang, Depeng, Qi Zheng, Kailong Huang, Dirk Springael, i Xu Xiang Zhang. 2020. „Metagenomic and metatranscriptomic insights into the complex nitrogen metabolic pathways in a single-stage bioreactor coupling partial denitrification with anammox“. *Chemical Engineering Journal* 398:125653. doi: 10.1016/j.cej.2020.125653.
- Wang, Yulin, Qigui Niu, Xu Zhang, Lei Liu, Yubo Wang, Yiqiang Chen, Mishty Negi, Daniel Figeys, Yu You Li, i Tong Zhang. 2019. „Exploring the effects of operational mode and microbial interactions on bacterial community assembly in a one-stage partial-nitrification anammox reactor using integrated multi-omics“. *Microbiome* 7(1):1–15. doi: 10.1186/S40168-019-0730-6/FIGURES/6.
- Winkler, Mari KH, i Levi Straka. 2019. „New directions in biological nitrogen removal and

recovery from wastewater“. *Current Opinion in Biotechnology* 57:50–55. doi: 10.1016/J.COPBIO.2018.12.007.

Zagrebačke Otpadne Vode. bez dat. „Naše postrojenje“. Preuzeto 17. listopad 2021. (<https://www.zov-zagreb.hr/hr/o-nama/nasa-postrojenja.html>).

Zhang, Meng, Siyu Wang, Bin Ji, i Yu Liu. 2019. „Towards mainstream deammonification of municipal wastewater: Partial nitrification-anammox versus partial denitrification-anammox“. *Science of The Total Environment* 692:393–401. doi: 10.1016/J.SCITOTENV.2019.07.293.

Zhu, Wenhan, Alexandre Lomsadze, i Mark Borodovsky. 2010. „Ab initio gene identification in metagenomic sequences“. *Nucleic Acids Research* 38(12):e132–e132. doi: 10.1093/NAR/GKQ275.

## 7. PRILOZI

### PRILOG 1: Tablica KEGG identifikacijskih brojeva ribosomskih proteina

Tablica 1. KEGG identifikacijski brojevi ribosomskih proteina

Tablica 1.a KEGG identifikacijski brojevi proteina male podjedinice ribosoma

Protein	Kloroplast		Bakterije	Arheje	Eukarioti
		Mitochondrij			
S1		K02945			
S2	S Ae	K02967			K02998
S3	S3e	K02982			K02985
S4	S9e	K02986			K02997
S5	S2e	K02988			K02981
S6		K02990			
S7	S5e	K02992			K02989
S8	S15A e	K02994			K02957
S9	S16e	K02996			K02960
S10	S20e	K02946			K02969
S11	S14e	K02948			K02955
S12	S23e	K02950			K02973
S13	S18e	K02952			K02964
S14	S29e	K02954			K02980
S15	S13e	K02956			K02953
S16		K02959			
S17	S11e	K02961			K02949
S18		K02963			
S19	S15e	K02965			K02958
S20		K02968			
S21		K02970			
	S3Ae			K02984	
	S4e			K02987	
	S6e			K02991	
	S7e				K02993
	S8e			K02995	
	S10e				K02947
	S12e				K02951
	S17e			K02962	
	S19e			K02966	
	S21e				K02971

Tablica 1.b KEGG identifikacijski brojevi proteina velike podjedinice ribosoma

Protein	Kloroplast		Bakterije	Arheje	Eukarioti
		Mitochondrij			
L1	L10Ae	K02863			K02865
L2	L8e	K02886			K02938
L3	L3e	K02906			K02925
L4	L4e	K02926		K02930	
L5	L11e	K02931			K02868
L6	L9e	K02933			K02940
L7/L12		K02935			
L9		K02939			
L10	P0	K02864			K02941
L11	L12e	K02867			K02870
L12	P1			K02869	K02942
	P2				K02943
L13	L13Ae	K02871			K02872
L14	L23e	K02874			K02894
L15	L27Ae	K02876			K02900
L16	L10e	K02878		K02866	
L17		K02879			
L18	L5e	K02881			K02932
L19		K02884			
L20		K02887			
L21		K02888			
L22	L17e	K02890			K02880
L23	L23Ae	K02892			K02893
L24	L26e	K02895			K02898
L25					
L27		K02899			
L28		K02902			
L29	L35e	K02904			K02918
L30	L7e	K02907			K02937
L31		K02909			
L32		K02911			

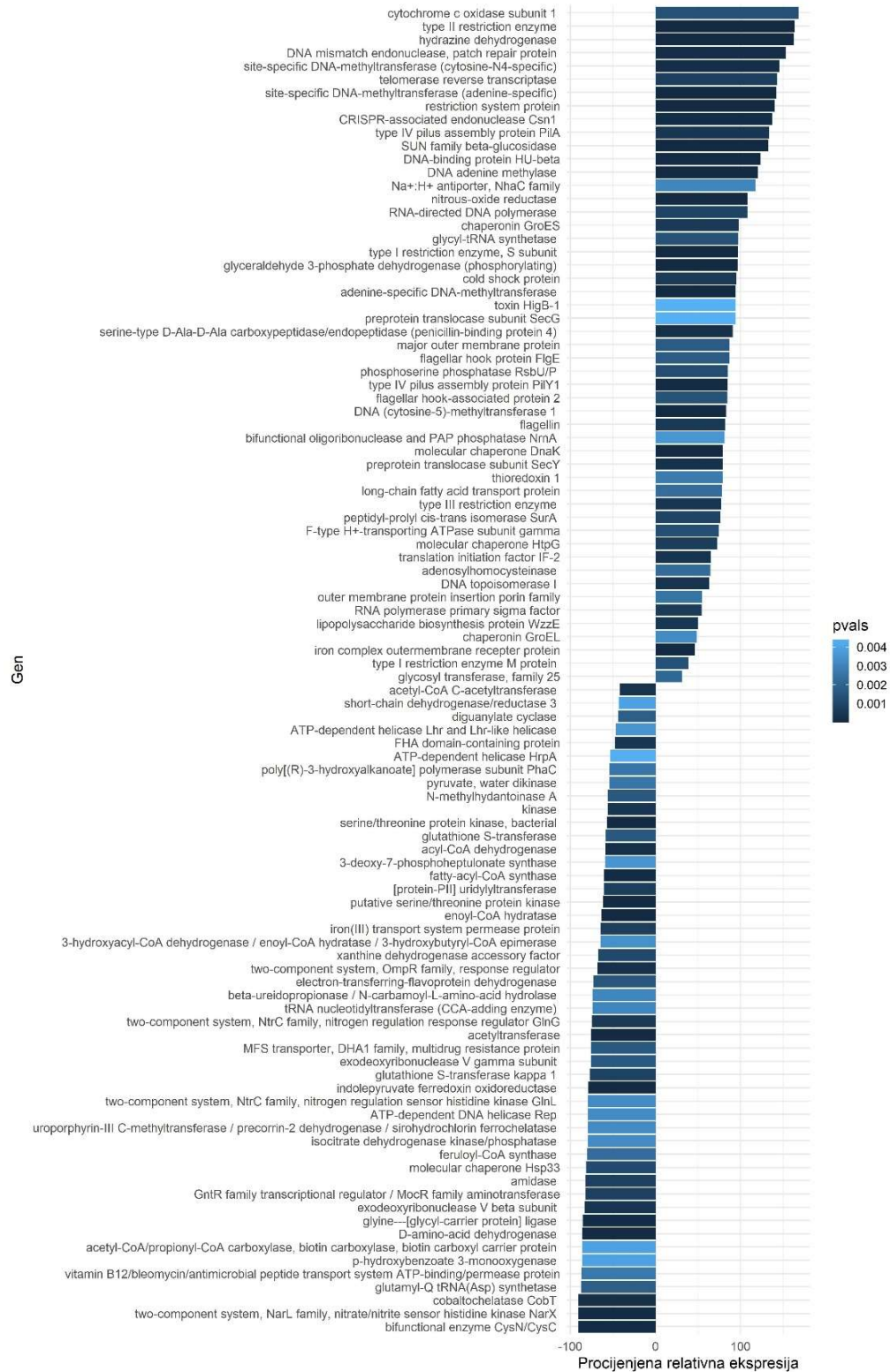


	S24e			K02974
	S25e			K02975
	S26e			K02976
	S27e			K02978
	S27A e			K02977
	S28e			K02979
	S30e			K02983
MRPS2 1		K17415		
MRPS2 2		K17401		
MRPS2 3		K17402		
MRPS2 4		K17403		
MRPS2 5		K17404		
MRPS2 6		K17405		
MRPS2 7		K17406		
MRPS2 8		K17407		
MRPS2 9		K17408		
MRPS3 0		K17409		
MRPS3 1		K17410		
MRPS3 3		K17411		
MRPS3 4		K17412		
MRPS3 5		K17413		
MRPS3 6		K17414		
YMR31		K17417		
PSRP3		K19032		
PSRP4		K19033		

L33		K02913		
L34		K02914		
L35		K02916		
L36		K02919		
L6e				K02934
L7A	L7Ae		K07590	K02936
	L13e			K02873
	L14e			K02875
	L15e			K02877
	L18e			K02883
	LX			K02944
	L18Ae			K02882
	L19e			K02885
	L21e			K02889
	L22e			K02891
	L24e			K02896
	L27e			K02901
	L28e			K02903
	L29e			K02905
	L30e			K02908
	L31e			K02910
	L32e			K02912
	L34e			K02915
	L35Ae			K02917
	L36e			K02920
	L37e			K02922
	L37Ae			K02921
	L38e			K02923
	L39e			K02924
	L40e			K02927
	L41e			K02928
	L44e (L36Ae )			K02929
MRPL3 7		K17418		
MRPL3 8		K17419		

MRPL3 9		K17420			
MRPL4 0		K17421			
MRPL4 1		K17422			
MRPL4 2		K17423			
MRPL4 3		K17424			
MRPL4 4		K17425			
MRPL4 5		K17426			
MRPL4 6		K17427			
MRPL4 7		K17428			
MRPL4 8		K17429			
MRPL4 9		K17430			
MRPL5 0		K17431			
MRPL5 1		K17432			
MRPL5 2		K17433			
MRPL5 3		K17434			
MRPL5 4		K17435			
MRPL5 5		K17436			
MRPL1 5		K17437			
MRPL2 5		K17438			
MRPL3 5		K17439			
MRP49		K17440			
PSRP5		K19034			
PSRP6		K19035			

## PRILOG 2: Relativna ekspresija gena



Slika 15 Graf relativne ekspresije predviđenih gena. Relativna ekspresija za svaki modul definirana je kao  $E_x$  u jednadžbi (5). Prikazano je 100 rezultata s najmanjom vrijednosti  $p$  radi preglednosti.

### **PRILOG 3: Datoteka za unos u mrežnu aplikaciju IPATH3**

category colour width

M00003 #002CE1 W2.92862474515965

M00006 #FF1660 W5.17097204788547

M00008 #FF1660 W6.42243507354912

M00009 #002CE1 W1.78797226957096

M00010 #002CE1 W4.48685790836585

M00013 #FF1660 W9.37030802337125

M00015 #FF1660 W6.62841535776186

M00017 #FF1660 W2.52196730702218

M00022 #FF1660 W3.08669897962935

M00023 #FF1660 W4.83903582889359

M00025 #FF1660 W4.88946944455488

M00032 #FF1660 W5.9033805756472

M00035 #002CE1 W8.0515746109441

M00036 #FF1660 W3.74119846505836

M00048 #FF1660 W2.19008891616348

M00061 #FF1660 W8.1973998209108

M00072 #002CE1 W25.9331045368147

M00085 #FF1660 W6.76968964515542

M00086 #FF1660 W5.1707850011151

M00087 #FF1660 W8.58569772861057

M00088 #FF1660 W7.53419074336897

M00094 #002CE1 W19.5841189215082

M00095 #FF1660 W4.31286419457458

M00096 #002CE1 W3.34021617696438

M00099 #002CE1 W22.9562279805705  
M00113 #FF1660 W7.01346940474962  
M00118 #FF1660 W9.80062413502995  
M00121 #FF1660 W3.95945238986369  
M00122 #FF1660 W4.99496242481326  
M00133 #FF1660 W5.54159841298769  
M00135 #FF1660 W3.62476323134507  
M00150 #FF1660 W10.1002382742109  
M00156 #FF1660 W6.15686279694813  
M00157 #002CE1 W3.80746209976552  
M00158 #002CE1 W23.5922850404223  
M00160 #002CE1 W19.5841189215082  
M00165 #002CE1 W3.06611564883083  
M00166 #002CE1 W7.43591547963339  
M00174 #002CE1 W21.831180225134  
M00176 #FF1660 W4.96679435346316  
M00177 #002CE1 W30  
M00178 #002CE1 W15.4124760827142  
M00179 #002CE1 W17.0573696676169  
M00183 #002CE1 W3.63659853305273  
M00188 #FF1660 W5.3596753524972  
M00190 #FF1660 W8.14718890123711  
M00212 #FF1660 W7.26378430630245  
M00237 #FF1660 W7.49208602306596  
M00239 #FF1660 W4.381515486117  
M00240 #FF1660 W7.4485282017761

M00258 #FF1660 W3.7529647595445  
M00261 #002CE1 W19.5841189215082  
M00262 #002CE1 W13.8252640049399  
M00264 #002CE1 W19.0917569635953  
M00285 #002CE1 W22.5997947801876  
M00300 #FF1660 W10.6071780664397  
M00307 #FF1660 W2.99103864410594  
M00308 #002CE1 W6.64097020599506  
M00331 #FF1660 W3.32943882930563  
M00332 #FF1660 W9.18913815247773  
M00334 #FF1660 W5.15480716652801  
M00335 #002CE1 W6.12518935297592  
M00337 #002CE1 W19.9680720096473  
M00340 #002CE1 W20.9611086357926  
M00341 #002CE1 W16.3438589788724  
M00352 #002CE1 W21.712388253323  
M00354 #002CE1 W20.9611086357926  
M00355 #002CE1 W18.4559289746386  
M00358 #002CE1 W13.1438130500062  
M00359 #002CE1 W1.49089345994474  
M00360 #002CE1 W1.34891035570972  
M00373 #FF1660 W5.27121853827419  
M00374 #FF1660 W2.73265423587521  
M00375 #FF1660 W5.7941265446118  
M00400 #002CE1 W15.9203696151781  
M00403 #002CE1 W15.9203696151781

M00427 #002CE1 W22.9562279805705  
M00444 #002CE1 W19.0917569635953  
M00445 #FF1660 W13.5854030754698  
M00455 #FF1660 W13.0986217544993  
M00471 #FF1660 W13.2060668772044  
M00493 #FF1660 W13.2704941540843  
M00497 #FF1660 W12.9927307339579  
M00501 #FF1660 W10.5849531047225  
M00507 #FF1660 W6.84996134230319  
M00516 #FF1660 W3.96987346560551  
M00528 #002CE1 W16.3526051198015  
M00529 #002CE1 W3.72226124221876  
M00530 #002CE1 W3.02754098553858  
M00531 #FF1660 W9.89325394330502  
M00545 #FF1660 W11.7870258107737  
M00546 #FF1660 W7.67536794085583  
M00552 #002CE1 W7.8254958841447  
M00570 #FF1660 W3.18714510697972  
M00615 #002CE1 W7.44978283385815  
M00620 #002CE1 W3.16733178480932  
M00631 #FF1660 W7.7475662406433  
M00647 #FF1660 W6.03276296400832  
M00722 #FF1660 W11.6433933284035  
M00723 #002CE1 W19.0917569635953  
M00724 #002CE1 W19.0917569635953  
M00740 #002CE1 W5.89121587346713

M00741 #FF1660 W4.46766252984494

M00744 #002CE1 W19.0917569635953

M00760 #FF1660 W13.0986217544993

M00774 #FF1660 W12.0268609866548

K00003 #FF1660 W6.55342624962495

K00021 #002CE1 W12.8062006329185

K00022 #FF1660 W6.30241832088774

K00031 #002CE1 W7.78095449557239

K00119 #FF1660 W11.9112747652015

K00128 #FF1660 W3.69972405871861

K00134 #002CE1 W14.2041068901532

K00135 #FF1660 W6.10183917294871

K00154 #FF1660 W11.9112747652015

K00161 #FF1660 W5.90419054146868

K00162 #FF1660 W6.33198998396954

K00163 #FF1660 W5.85662468458383

K00175 #002CE1 W6.50466103063759

K00177 #002CE1 W18.265732425729

K00210 #FF1660 W9.30748980881625

K00241 #002CE1 W7.49847413460866

K00244 #FF1660 W10.9038192296248

K00249 #FF1660 W8.96538285225939

K00257 #FF1660 W6.66285140272822

K00260 #FF1660 W8.52567690926458

K00261 #002CE1 W8.57156119591373

K00265 #FF1660 W4.3520810105317



K00266 #FF1660 W6.75872268900138  
K00285 #FF1660 W12.4552557676155  
K00311 #FF1660 W10.5052430337545  
K00318 #002CE1 W14.6350614354452  
K00372 #FF1660 W11.9112747652015  
K00375 #FF1660 W8.80548641873465  
K00376 #002CE1 W15.9403393397795  
K00383 #FF1660 W9.81197534496085  
K00401 #002CE1 W17.9852809810808  
K00405 #FF1660 W9.82161507372668  
K00481 #FF1660 W12.466170763216  
K00483 #FF1660 W11.5936991257189  
K00528 #FF1660 W8.78390967873752  
K00547 #FF1660 W9.92541940921686  
K00558 #002CE1 W12.2810720924801  
K00571 #002CE1 W21.4192280370964  
K00574 #FF1660 W6.99500081115132  
K00590 #002CE1 W24.9725717854687  
K00612 #002CE1 W7.53070246800817  
K00626 #FF1660 W5.93951796431429  
K00632 #FF1660 W7.36158559114511  
K00666 #FF1660 W8.67058162779407  
K00696 #002CE1 W16.0535517021911  
K00697 #FF1660 W8.1831391120429  
K00754 #002CE1 W4.37221913090831  
K00782 #FF1660 W9.30748980881625

K00795 #FF1660 W9.43509668115791  
K00797 #FF1660 W7.78653052477818  
K00799 #FF1660 W9.005608307752  
K00806 #FF1660 W7.21620940835768  
K00809 #002CE1 W10.4143723750392  
K00812 #002CE1 W6.23286397942073  
K00870 #FF1660 W11.6580051003953  
K00901 #FF1660 W11.5936991257189  
K00906 #FF1660 W11.4778551949388  
K00924 #FF1660 W8.56285840294571  
K00955 #FF1660 W13.155644958285  
K00974 #FF1660 W9.69275058152771  
K00980 #002CE1 W18.9944426954081  
K00982 #FF1660 W6.37803066829227  
K00986 #002CE1 W15.8467115687235  
K00990 #FF1660 W8.69615285767567  
K01007 #FF1660 W7.78472718751962  
K01077 #002CE1 W12.4158349301859  
K01119 #002CE1 W16.2607332080616  
K01130 #FF1660 W7.17459810480368  
K01152 #002CE1 W6.19750096204428  
K01153 #002CE1 W4.06697723500607  
K01154 #002CE1 W14.5890027591039  
K01155 #002CE1 W23.8167873290999  
K01156 #002CE1 W11.444924409878  
K01176 #002CE1 W12.8062006329185

K01178 #FF1660 W11.5936991257189

K01238 #002CE1 W19.4347933186225

K01251 #002CE1 W9.55289147992924

K01277 #002CE1 W19.0509280259908

K01322 #002CE1 W7.4446489868366

K01354 #002CE1 W11.4283842804555

K01362 #002CE1 W4.88135217627749

K01372 #002CE1 W18.9944426954081

K01414 #FF1660 W6.99500081115132

K01417 #002CE1 W10.1293747150354

K01426 #FF1660 W11.8935331509498

K01451 #FF1660 W8.34131979058808

K01463 #FF1660 W11.5936991257189

K01467 #002CE1 W6.51003133088822

K01469 #FF1660 W8.38285064057799

K01473 #FF1660 W8.08871757216983

K01493 #002CE1 W12.416113121674

K01507 #002CE1 W5.87033537101437

K01525 #FF1660 W10.2429637723364

K01572 #002CE1 W16.3863966894795

K01599 #FF1660 W7.1385573635077

K01610 #002CE1 W9.48024600893181

K01623 #002CE1 W7.61515406215282

K01626 #FF1660 W8.53579286913677

K01628 #FF1660 W10.4408831952978

K01638 #FF1660 W8.78390967873752

K01640 #FF1660 W9.04784711882587

K01643 #002CE1 W22.2615977525309

K01652 #FF1660 W5.14092309943311

K01657 #FF1660 W5.79850012094641

K01676 #FF1660 W7.79887018938637

K01681 #002CE1 W6.0152947305702

K01692 #FF1660 W9.12765406993977

K01714 #FF1660 W7.00747399826832

K01719 #002CE1 W11.8372800518024

K01744 #002CE1 W10.751525466568

K01754 #FF1660 W6.0369750394665

K01755 #FF1660 W6.67079493753163

K01781 #FF1660 W11.5936991257189

K01782 #FF1660 W9.2836715404061

K01796 #FF1660 W11.7616924347289

K01802 #002CE1 W8.91285836900587

K01860 #002CE1 W14.8087050670903

K01872 #002CE1 W5.87033537101437

K01880 #002CE1 W14.3007881665047

K01894 #FF1660 W12.626594322086

K01897 #FF1660 W4.36958182260095

K01902 #002CE1 W6.51387119091811

K01912 #FF1660 W5.39623921682428

K01919 #FF1660 W8.86911519114704

K01934 #FF1660 W10.1112244583177

K01971 #FF1660 W6.56745274180436

K01992 #002CE1 W5.76554000055314

K01995 #FF1660 W9.17094044494132

K01996 #FF1660 W10.0123533035368

K01997 #FF1660 W9.24341962255369

K01998 #FF1660 W8.71234730639462

K02004 #FF1660 W2.90566712619344

K02011 #FF1660 W9.27768273539915

K02013 #FF1660 W7.69190137516365

K02014 #002CE1 W6.73552256640232

K02031 #FF1660 W5.53590251287339

K02032 #FF1660 W6.95756121563812

K02034 #FF1660 W4.55020227194299

K02049 #FF1660 W6.51440803770466

K02115 #002CE1 W10.9790489779646

K02256 #002CE1 W24.6013601166062

K02302 #FF1660 W11.4778551949388

K02319 #002CE1 W17.3083988403665

K02327 #002CE1 W13.0206338284829

K02336 #002CE1 W18.9944426954081

K02338 #002CE1 W6.21158567796945

K02356 #002CE1 W8.93013061931962

K02379 #FF1660 W7.80451070252274

K02390 #002CE1 W12.8337287777286

K02406 #002CE1 W12.0268137971509

K02407 #002CE1 W12.4154400876276

K02458 #FF1660 W11.5936991257189

K02461 #FF1660 W10.9038192296248  
K02470 #002CE1 W4.8831544963978  
K02471 #FF1660 W12.5496903466956  
K02483 #FF1660 W9.79952277419493  
K02485 #002CE1 W14.2096544809406  
K02487 #FF1660 W7.6056574295911  
K02495 #FF1660 W4.55835935185554  
K02518 #002CE1 W8.37170876056074  
K02519 #002CE1 W9.63388278657116  
K02520 #002CE1 W10.482968757425  
K02529 #FF1660 W6.36137000604422  
K02568 #002CE1 W13.4126364167212  
K02575 #002CE1 W7.2629534026851  
K02626 #002CE1 W13.9396823581315  
K02666 #002CE1 W8.57156119591373  
K02667 #FF1660 W8.02457039977729  
K02668 #FF1660 W10.4840182125867  
K02674 #002CE1 W11.9527197114263  
K02843 #002CE1 W8.95435464337343  
K02863 #002CE1 W14.1183751279016  
K02864 #002CE1 W12.9047752003698  
K02867 #002CE1 W16.3293043918078  
K02871 #002CE1 W13.6047721434949  
K02874 #002CE1 W13.0717280990931  
K02876 #002CE1 W18.4961976540674  
K02878 #002CE1 W17.7671766582548

K02879 #002CE1 W10.6217557394111

K02881 #002CE1 W17.081486700254

K02884 #002CE1 W11.6378754879457

K02886 #002CE1 W13.4959889366216

K02887 #002CE1 W20.6046640651658

K02888 #002CE1 W15.1877946613285

K02890 #002CE1 W15.8467115687235

K02892 #002CE1 W14.379679202338

K02895 #002CE1 W22.5347090406024

K02899 #002CE1 W21.5755027357934

K02902 #002CE1 W20.9248756766138

K02904 #002CE1 W16.0034740316953

K02906 #002CE1 W17.8942122447294

K02907 #002CE1 W22.8058218498296

K02909 #002CE1 W17.1108019812176

K02914 #002CE1 W12.0994564267635

K02916 #002CE1 W14.6350614354452

K02919 #002CE1 W20.1047571504611

K02926 #002CE1 W9.84063008536078

K02931 #002CE1 W20.1796576987866

K02933 #002CE1 W15.887297948113

K02935 #002CE1 W8.90595659973248

K02945 #002CE1 W5.52806218452948

K02946 #002CE1 W15.0535413242693

K02948 #002CE1 W9.44236762873503

K02950 #002CE1 W15.8467115687235

K02952 #002CE1 W16.3679617840544  
K02954 #002CE1 W10.5108434882067  
K02956 #002CE1 W18.7466662811947  
K02959 #002CE1 W9.55318238015085  
K02961 #002CE1 W14.7553103324491  
K02963 #002CE1 W14.0857186508945  
K02965 #002CE1 W14.379679202338  
K02967 #002CE1 W9.3604911947171  
K02968 #002CE1 W11.2658754926826  
K02970 #002CE1 W16.6836438310789  
K02982 #002CE1 W8.94358222714219  
K02988 #002CE1 W19.7537031439328  
K02992 #002CE1 W14.4086025748687  
K02994 #002CE1 W14.2397259140414  
K02996 #002CE1 W11.0473721090219  
K03040 #002CE1 W8.6389573691252  
K03046 #002CE1 W4.27693597770158  
K03073 #002CE1 W15.7379776457092  
K03075 #002CE1 W13.8318387275177  
K03076 #002CE1 W11.0572523204933  
K03086 #002CE1 W7.70537653314017  
K03088 #FF1660 W3.0444298899509  
K03111 #002CE1 W8.9657715309219  
K03116 #002CE1 W10.888995753  
K03162 #002CE1 W12.7395714428985  
K03164 #002CE1 W17.9852809810808



K03168 #002CE1 W9.40032431940148  
K03169 #FF1660 W7.48968182378342  
K03182 #FF1660 W5.71912120516866  
K03210 #002CE1 W9.65604212869401  
K03220 #FF1660 W7.41734236528235  
K03231 #002CE1 W21.1443867077454  
K03282 #002CE1 W7.90676986757187  
K03306 #002CE1 W7.84228715135613  
K03315 #002CE1 W15.1369486099919  
K03330 #002CE1 W18.9944426954081  
K03335 #FF1660 W12.1661171021462  
K03386 #002CE1 W8.37170876056074  
K03403 #002CE1 W21.1443867077454  
K03411 #002CE1 W14.486133723019  
K03413 #002CE1 W8.93013061931962  
K03427 #002CE1 W5.47593431300036  
K03455 #FF1660 W5.34436636874257  
K03473 #002CE1 W22.2615977525309  
K03519 #FF1660 W7.66016155391849  
K03530 #002CE1 W18.1702089427157  
K03553 #002CE1 W6.569763066745  
K03564 #002CE1 W9.37166735634427  
K03578 #FF1660 W6.93597889040263  
K03579 #FF1660 W8.10675072109518  
K03582 #FF1660 W12.0293750265172  
K03583 #FF1660 W10.9818460915042

K03587 #002CE1 W7.22149026377938

K03598 #FF1660 W12.1661171021462

K03611 #FF1660 W10.6079223622569

K03640 #002CE1 W11.2456980928191

K03656 #FF1660 W11.4778551949388

K03671 #002CE1 W11.6430070051011

K03672 #FF1660 W11.7616924347289

K03684 #FF1660 W10.1112244583177

K03688 #FF1660 W8.06004816408116

K03704 #002CE1 W13.9905009801007

K03717 #FF1660 W12.045315471788

K03724 #FF1660 W6.39051609402467

K03726 #002CE1 W19.7563121190063

K03737 #002CE1 W4.97811520245259

K03750 #FF1660 W5.00635528519589

K03768 #002CE1 W8.45455815485834

K03771 #002CE1 W11.2824334491548

K03821 #FF1660 W8.41573746945841

K03883 #002CE1 W16.0535517021911

K03918 #002CE1 W18.9944426954081

K03975 #002CE1 W12.0544403312413

K04043 #002CE1 W11.9704644993644

K04044 #FF1660 W10.4408831952978

K04077 #002CE1 W7.22716625139842

K04078 #002CE1 W14.4086025748687

K04079 #002CE1 W10.1521331851246

K04083 #FF1660 W11.8198315854359  
K04090 #FF1660 W11.4289018824497  
K04108 #FF1660 W10.1112244583177  
K04437 #002CE1 W17.9852809810808  
K04477 #FF1660 W9.40095192233736  
K04564 #002CE1 W9.35006955139564  
K04646 #002CE1 W17.9852809810808  
K04656 #FF1660 W6.95753045902222  
K04739 #002CE1 W19.7563121190063  
K04744 #FF1660 W7.41734236528235  
K04749 #002CE1 W11.6430070051011  
K04761 #FF1660 W9.29560797076534  
K04764 #002CE1 W14.486133723019  
K04780 #FF1660 W12.3632224611212  
K05083 #002CE1 W16.0535517021911  
K05524 #FF1660 W8.3362500695918  
K05559 #FF1660 W6.92728339390288  
K05790 #002CE1 W6.92936159952724  
K05799 #FF1660 W11.7616924347289  
K05802 #FF1660 W11.9112747652015  
K05812 #002CE1 W8.93013061931962  
K05896 #FF1660 W6.99500081115132  
K05903 #FF1660 W10.9038192296248  
K05973 #FF1660 W10.2588287681162  
K06016 #FF1660 W10.699187093652  
K06076 #002CE1 W11.5474756225062

xx

K06142 #002CE1 W11.1241683265269

K06174 #002CE1 W21.1443867077454

K06193 #002CE1 W14.8087050670903

K06223 #002CE1 W17.665592504131

K06326 #002CE1 W10.1888308806179

K06381 #002CE1 W13.9396823581315

K06871 #FF1660 W8.1831391120429

K06881 #002CE1 W11.9791749946093

K06895 #FF1660 W9.66195823916855

K06932 #002CE1 W19.7563121190063

K06943 #002CE1 W13.8625676197206

K06958 #FF1660 W8.17100626866384

K06960 #002CE1 W12.416113121674

K06979 #FF1660 W11.5936991257189

K07011 #002CE1 W13.0206338284829

K07012 #002CE1 W8.60099863325633

K07016 #002CE1 W6.86746361372653

K07027 #002CE1 W2.34119416635942

K07028 #FF1660 W12.9348439119983

K07091 #FF1660 W10.6079223622569

K07094 #002CE1 W17.6674657420728

K07102 #FF1660 W8.17100626866384

K07133 #002CE1 W9.45486874398996

K07151 #002CE1 W22.2615977525309

K07160 #FF1660 W12.275548767976

K07169 #FF1660 W7.13622611782928

K07171 #002CE1 W11.1418409118764

K07176 #FF1660 W8.80970294284488

K07182 #FF1660 W6.27388943468072

K07221 #002CE1 W11.3760933942934

K07234 #FF1660 W11.201534258603

K07245 #002CE1 W21.047467530448

K07259 #002CE1 W13.3576453830038

K07265 #002CE1 W7.4446489868366

K07270 #002CE1 W4.02865375619298

K07277 #002CE1 W7.69259897851975

K07315 #002CE1 W12.8337287777286

K07316 #002CE1 W7.54168282941403

K07334 #002CE1 W14.379679202338

K07339 #002CE1 W22.2615977525309

K07387 #002CE1 W12.8062006329185

K07402 #FF1660 W9.72315275462822

K07444 #002CE1 W9.81154669226354

K07445 #002CE1 W8.14264942861496

K07448 #002CE1 W20.5805316013556

K07453 #002CE1 W12.9047752003698

K07458 #002CE1 W19.5911131332046

K07459 #002CE1 W17.913142755557

K07461 #002CE1 W16.5312755530798

K07468 #002CE1 W20.2816521338497

K07494 #002CE1 W17.6674657420728

K07500 #002CE1 W9.67392229656386

K07552 #FF1660 W10.9818460915042

K07566 #FF1660 W5.95791385031745

K07572 #002CE1 W30

K07575 #002CE1 W4.73486506979632

K07637 #002CE1 W17.9852809810808

K07638 #FF1660 W12.3751435496252

K07647 #FF1660 W11.5936991257189

K07673 #FF1660 W13.155644958285

K07708 #FF1660 W11.4778551949388

K07712 #FF1660 W10.7432295711428

K07726 #002CE1 W14.486133723019

K07739 #002CE1 W12.0994564267635

K07814 #FF1660 W8.75495420963605

K08070 #002CE1 W7.7043857249433

K08083 #FF1660 W12.045315471788

K08087 #002CE1 W17.9852809810808

K08097 #002CE1 W19.7563121190063

K08100 #002CE1 W17.4935749442898

K08224 #FF1660 W12.9348439119983

K08252 #002CE1 W11.0872951877333

K08309 #FF1660 W8.03718436259023

K08311 #FF1660 W11.7616924347289

K08603 #002CE1 W17.9852809810808

K08651 #002CE1 W17.9852809810808

K08717 #002CE1 W17.4935749442898

K08884 #FF1660 W8.30202444038322

K08974 #002CE1 W16.2607332080616

K08984 #FF1660 W10.2588287681162

K09003 #002CE1 W17.9852809810808

K09006 #002CE1 W17.9852809810808

K09165 #002CE1 W13.9396823581315

K09181 #FF1660 W10.9709452573829

K09228 #002CE1 W17.9852809810808

K09421 #002CE1 W19.7563121190063

K09456 #FF1660 W11.1578180885554

K09487 #FF1660 W10.4408831952978

K09489 #002CE1 W17.9852809810808

K09497 #002CE1 W17.9852809810808

K09498 #002CE1 W17.9852809810808

K09585 #002CE1 W9.34032851348199

K09597 #FF1660 W6.35684585334472

K09705 #002CE1 W13.2986111874181

K09765 #002CE1 W17.9852809810808

K09767 #002CE1 W12.4158349301859

K09769 #002CE1 W11.0872951877333

K09800 #FF1660 W5.79850012094641

K09803 #002CE1 W12.6832001788037

K09883 #FF1660 W13.0752934782016

K09903 #002CE1 W7.39130044507293

K09919 #FF1660 W11.8198315854359

K09951 #002CE1 W7.7043857249433

K09952 #002CE1 W20.1649777351131

K09967 #002CE1 W14.8087050670903

K10441 #FF1660 W8.64210459841752

K10563 #FF1660 W6.42987567704161

K10726 #002CE1 W16.0535517021911

K10807 #002CE1 W16.0535517021911

K10817 #FF1660 W10.6079223622569

K10908 #002CE1 W17.9852809810808

K10941 #FF1660 W10.2301858869416

K10946 #002CE1 W17.6674657420728

K11105 #002CE1 W7.49847413460866

K11126 #002CE1 W21.485892954519

K11146 #FF1660 W4.43280517424905

K11263 #FF1660 W12.466170763216

K11609 #FF1660 W8.41573746945841

K11690 #FF1660 W10.2479457835523

K11799 #002CE1 W8.00740374816947

K11931 #FF1660 W11.5936991257189

K12072 #002CE1 W14.8087050670903

K12206 #002CE1 W19.7563121190063

K12508 #FF1660 W11.5713024159232

K12511 #FF1660 W10.2588287681162

K13010 #002CE1 W17.6674657420728

K13019 #002CE1 W15.5359665293238

K13069 #FF1660 W6.98370536233551

K13075 #FF1660 W12.3751435496252

K13282 #002CE1 W14.486133723019



K13299 #FF1660 W12.2084388701755  
K13525 #002CE1 W14.8087050670903  
K13581 #002CE1 W16.9988129371509  
K13788 #002CE1 W15.2488936458703  
K13821 #FF1660 W8.52567690926458  
K13841 #002CE1 W21.6885514590869  
K13955 #FF1660 W6.92728339390288  
K14165 #FF1660 W6.95753045902222  
K15977 #FF1660 W5.82658151301256  
K16039 #002CE1 W16.2607332080616  
K16183 #FF1660 W11.5936991257189  
K16557 #002CE1 W19.7563121190063  
K17602 #002CE1 W26.4550301149449  
K17910 #FF1660 W10.5878601727252  
K18166 #FF1660 W3.81595066762207  
K18320 #FF1660 W6.46353033986851  
K18360 #FF1660 W12.466170763216  
K19428 #002CE1 W9.4604310030224  
K19691 #FF1660 W3.26386246659254  
K20935 #002CE1 W23.8167873290999  
K23686 #002CE1 W17.9852809810808  
map00020 #002CE1 W1.83451525604621  
map00053 #FF1660 W3.58076774857042  
map00071 #FF1660 W6.49990342644256  
map00072 #FF1660 W6.75144996577613  
map00140 #FF1660 W8.88768839260772

map00190 #002CE1 W3.03279393472836  
map00195 #002CE1 W6.06100918028934  
map00240 #002CE1 W2.3718523365142  
map00261 #FF1660 W3.24497346656222  
map00280 #FF1660 W5.43580336130768  
map00281 #FF1660 W11.0932185260655  
map00310 #FF1660 W5.67302576583718  
map00330 #FF1660 W3.69034634799278  
map00332 #FF1660 W6.80172367252893  
map00350 #FF1660 W3.29742045420577  
map00360 #FF1660 W5.28884795594296  
map00362 #FF1660 W7.01380719237742  
map00380 #FF1660 W5.28729497656299  
map00400 #FF1660 W2.56429620233224  
map00405 #FF1660 W6.61838826321013  
map00410 #FF1660 W6.78616141551501  
map00471 #002CE1 W7.13996101057641  
map00480 #FF1660 W2.96192524629877  
map00510 #002CE1 W13.8971765361175  
map00513 #002CE1 W13.8508244340905  
map00520 #002CE1 W1.90831302609792  
map00550 #002CE1 W2.38845226149115  
map00561 #FF1660 W2.98861441215223  
map00600 #FF1660 W7.47214867656069  
map00620 #FF1660 W1.60302382767885  
map00624 #FF1660 W10.4877870997809

map00625 #FF1660 W4.38687875246633  
map00626 #FF1660 W4.89108071201999  
map00627 #FF1660 W10.9958867136367  
map00633 #FF1660 W5.91743218230574  
map00640 #FF1660 W3.82697577563508  
map00642 #FF1660 W9.52993135806556  
map00643 #FF1660 W7.62459472854184  
map00650 #FF1660 W4.09049985105904  
map00710 #002CE1 W4.01188760231544  
map00830 #FF1660 W4.98653131079864  
map00860 #FF1660 W2.01211041066854  
map00903 #FF1660 W8.94980824962097  
map00910 #002CE1 W3.01222875763965  
map00930 #FF1660 W10.5779792153561  
map00970 #002CE1 W2.92961142396953  
map00980 #FF1660 W7.50300180602018  
map00982 #FF1660 W6.65426165558956  
map01040 #FF1660 W4.19891978820305  
map01110 #FF1660 W0.903684079011128  
map01120 #FF1660 W1.15407414516491  
map01130 #FF1660 W0.861513643830983  
map01212 #FF1660 W4.32928164272632  
map01220 #FF1660 W5.53055055789114  
map01521 #002CE1 W30  
map01522 #002CE1 W21.7151224663697  
map02010 #FF1660 W3.68375587482675

map02020 #FF1660 W1.85344941349162  
map02024 #FF1660 W3.07026672689565  
map02025 #FF1660 W5.01475481343327  
map02040 #002CE1 W3.25213722932962  
map03010 #002CE1 W21.4076329393886  
map03013 #002CE1 W11.1723848166592  
map03015 #002CE1 W26.5003930686686  
map03018 #002CE1 W4.26781337475239  
map03020 #002CE1 W6.55178394529966  
map03022 #002CE1 W14.1015191580083  
map03030 #002CE1 W2.74907080553868  
map03040 #002CE1 W28.6397500875658  
map03050 #002CE1 W13.5794105756806  
map03060 #002CE1 W8.17179942861607  
map03070 #002CE1 W2.70938551758081  
map03320 #FF1660 W5.51108706733403  
map03430 #002CE1 W1.99818195825446  
map03460 #002CE1 W27.3394923297618  
map04010 #002CE1 W24.6477490942823  
map04011 #002CE1 W15.5824780784793  
map04012 #002CE1 W29.1168714672804  
map04013 #002CE1 W15.4500852739646  
map04014 #002CE1 W15.7248410571332  
map04015 #002CE1 W27.779960267893  
map04016 #002CE1 W10.2599613610977  
map04020 #002CE1 W23.2981292525935

map04022 #002CE1 W26.9949029914043  
map04024 #002CE1 W17.374041405462  
map04062 #002CE1 W26.2117769015786  
map04066 #002CE1 W4.46348049603708  
map04068 #002CE1 W5.72676985869596  
map04071 #002CE1 W9.93352709808911  
map04110 #002CE1 W13.0291359760731  
map04111 #002CE1 W19.847698369458  
map04112 #002CE1 W3.61649258282784  
map04114 #002CE1 W16.3745716598905  
map04120 #002CE1 W23.9840151054323  
map04136 #002CE1 W22.5625290457243  
map04138 #002CE1 W11.5630374479815  
map04140 #002CE1 W21.8083140238977  
map04141 #002CE1 W14.3064432780143  
map04144 #002CE1 W22.9853544902023  
map04145 #002CE1 W20.9020127118822  
map04150 #002CE1 W20.5847808973143  
map04151 #002CE1 W7.20726196904177  
map04210 #002CE1 W14.2525944201286  
map04211 #002CE1 W13.0610568365067  
map04212 #002CE1 W11.138852737722  
map04214 #002CE1 W13.8971765361175  
map04217 #002CE1 W5.13415816620001  
map04218 #002CE1 W16.7661172266702  
map04260 #002CE1 W10.6438474056263

XXX

map04261 #002CE1 W23.0007861373372  
map04310 #002CE1 W20.4198887286364  
map04360 #002CE1 W22.0374479836373  
map04370 #002CE1 W27.7571573062698  
map04371 #002CE1 W23.1277053934577  
map04390 #002CE1 W16.3745716598905  
map04510 #002CE1 W28.4659517601216  
map04512 #002CE1 W22.8243787223473  
map04514 #002CE1 W22.5625290457243  
map04520 #002CE1 W27.779960267893  
map04530 #002CE1 W18.7912357939718  
map04540 #002CE1 W21.2462279876488  
map04550 #002CE1 W29.1168714672804  
map04611 #002CE1 W17.9461623682507  
map04612 #002CE1 W16.8368280514594  
map04614 #002CE1 W9.31165532996119  
map04621 #002CE1 W17.2523533509685  
map04626 #002CE1 W5.77464371233632  
map04640 #002CE1 W24.5569655460851  
map04657 #002CE1 W10.9670855529213  
map04659 #002CE1 W15.3203843864696  
map04660 #002CE1 W20.5609926590514  
map04666 #002CE1 W18.8811445203853  
map04670 #002CE1 W26.5003930686686  
map04713 #002CE1 W24.368859875548  
map04720 #002CE1 W28.1259643748157

map04721 #002CE1 W25.730923093777  
map04722 #002CE1 W18.0142479659143  
map04723 #002CE1 W12.0571699948517  
map04725 #002CE1 W24.2650571688904  
map04730 #002CE1 W20.5609926590514  
map04750 #002CE1 W26.2117769015786  
map04810 #002CE1 W27.9423871530137  
map04912 #002CE1 W21.2583810805353  
map04913 #002CE1 W26.2117769015786  
map04914 #002CE1 W14.6159393234528  
map04915 #002CE1 W11.6517380156412  
map04916 #002CE1 W26.9949029914043  
map04920 #FF1660 W5.53242518293173  
map04921 #002CE1 W26.1090989984384  
map04922 #FF1660 W3.11766539415337  
map04924 #002CE1 W13.8997678545418  
map04925 #002CE1 W22.5625290457243  
map04926 #002CE1 W22.1333246646388  
map04927 #002CE1 W22.1333246646388  
map04932 #002CE1 W8.9316844001429  
map04933 #002CE1 W22.1333246646388  
map04934 #002CE1 W8.75751094228517  
map04940 #002CE1 W10.2147060276021  
map04960 #002CE1 W22.1333246646388  
map04961 #002CE1 W19.4584400347205  
map04966 #002CE1 W27.0365959217899

map04971 #002CE1 W24.2650571688904  
map04973 #002CE1 W18.3718904857711  
map04974 #002CE1 W12.8994713707725  
map04976 #002CE1 W17.2746683502825  
map04979 #002CE1 W28.9879616151852  
map05010 #002CE1 W15.007209854622  
map05012 #002CE1 W12.4617190290492  
map05016 #002CE1 W12.1824779724306  
map05100 #002CE1 W22.9134352753366  
map05110 #002CE1 W24.0951670963861  
map05130 #002CE1 W23.3581038656438  
map05131 #002CE1 W19.5011792060059  
map05132 #002CE1 W11.9453559296238  
map05134 #002CE1 W7.51329873390433  
map05142 #002CE1 W16.3121374180804  
map05144 #002CE1 W29.8916122091265  
map05152 #002CE1 W13.4995481560289  
map05160 #002CE1 W20.8347223680565  
map05165 #002CE1 W7.20726196904177  
map05166 #002CE1 W18.2227189612949  
map05167 #002CE1 W16.3745716598905  
map05169 #002CE1 W12.6181223857907  
map05200 #002CE1 W5.76895807214042  
map05204 #FF1660 W9.72537122122925  
map05205 #002CE1 W11.7159005354931  
map05206 #002CE1 W19.2495721103641



map05212 #002CE1 W18.8811445203853

map05213 #002CE1 W20.0235992715932

map05215 #002CE1 W12.0078961778926

map05216 #002CE1 W17.5156207508318

map05223 #002CE1 W17.5156207508318

map05224 #002CE1 W22.7296267167783

map05226 #002CE1 W27.9119299259814

map05322 #002CE1 W20.1189766483079

map05323 #002CE1 W27.0365959217899

map05412 #002CE1 W24.1992592570278

## 8. ŽIVOTOPIS

Rođena sam 1994. godine u Zagrebu, a 1995. godine preselila sam s obitelji u Split gdje sam završila osnovnu i srednju školu. Godine 2014. upisala sam Preddiplomski studij molekularne biologije na Prirodoslovno-matematičkom fakultetu u Zagrebu koji sam i završila 2019. godine sa završnim radom naslova „Porijeklo genetičkog koda“ pod vodstvom izv. prof. dr. sc. Damjana Franjevića. Laboratorijsku stručnu praksu odradila sam na Institutu Ruđer Bošković u Laboratoriju za biofiziku stanice dr. sc. Ive Tolić. Godine 2019. upisala sam i Diplomski studij molekularne biologije tijekom kojega sam se orijentirala na računalnu biologiju. Tijekom studiranja obavljala sam razne studentske poslove, od konobarenja i korisničke podrške do razvijanja genetičkih testova u laboratoriju.