

UNIVERZA V LJUBLJANI
BIOTEHNIŠKA FAKULTETA
ODDELEK ZA BIOLOGIJO

Luka AUSEC

**PRIPRAVA IN ANALIZA KNJIŽNICE GENOV ZA 16S rRNA IZ
ŠOTNIH TAL LJUBLJANSKEGA BARJA**

DIPLOMSKO DELO

Univerzitetni študij

**CONSTRUCTION AND ANALYSIS OF 16S rRNA CLONE LIBRARY
FROM *Sphagnum* BOG IN THE LJUBLJANA MARSH**

GRADUATION THESIS

University studies

Ljubljana, 2008

Diplomsko delo je zaključek univerzitetnega študija biologije. Opravljeno je bilo opravljeno na Katedri za mikrobiologijo Oddelka za živilstvo Biotehniške fakultete v Ljubljani.

Komisija za študijske zadeve univerzitetnega dodiplomskega študija biologije je dne 26.5.2006 za mentorico imenovala prof. dr. Ines Mandić-Mulec.

Mentorica: prof. dr. Ines Mandić-Mulec

Recenzent: prof. dr. Darja Žgur Bertok

Komisija za oceno in zagovor:

Predsednik: prof. dr. Kristina Sepčič

Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo

Članica: prof. dr. Ines Mandić-Mulec

Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za živilstvo

Članica: prof. dr. Darja Žgur Bertok

Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo

Datum zagovora: 5.6.2008

Podpisani se strinjam z objavo svoje naloge v polnem tekstu na spletni strani Digitalne knjižnice Biotehniške fakultete. Izjavljam, da je naloga, ki sem jo oddal v elektronski obliki, identična tiskani verziji.

Delo je rezultat lastnega raziskovalnega dela.

Luka Ausec

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

- ŠD Dn
- UDK 579.66: 575.86: 556.56 (043.2) = 163.6
- KG Ljubljansko barje, visoko barje, filogenija, struktura bakterijske združbe, genska knjižnica, 16S rRNA
- AV AUSEC, Luka
- SA MANDIĆ-MULEC, Ines (mentor)/ŽGUR-BERTOK, Darja (recenzent)
- KZ SI-1000 Ljubljana, Večna pot 111
- ZA Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za biologijo
- LI 2008
- IN PRIPRAVA IN ANALIZA KNJIŽNICE GENOV ZA 16S rRNA IZ ŠOTNIH TAL LJUBLJANSKEGA BARJA
- TD Diplomsko delo (univerzitetni študij)
- OP IX, 57 str., 7 pregl., 11 sl., 48 vir
- IJ sl
- JI sl/en
- AI Barje je pomemben ekosistem, v katerem so mikroorganizmi ključni člen pri kroženju snovi in energije. Raznovrstnost bakterij v barjanskih tleh je zelo slabo poznana. Pričujoče delo predstavlja vpogled v filogenetsko sestavo bakterijske združbe v kisljih (pH 4,5) šotnih tleh Ljubljanskega barja (Kozlarjeva gošča) s pomočjo priprave knjižnice genov za 16S rRNA. 154 delnih sekvenc smo s programom ARB filogenetsko uvrstili med *Acidobacteria* (41,6%) in *Proteobacteria* (40,2%), ostale skupine so bile bolj skromno zastopane: *Actinobacteria* (7,1%), *Plactomycetes* (5,2%), *Verrucomicrobia* (5,2%) in *Spirochaetes* (0,7%). Neobičajno velik je delež skupine *Acidobacteria*, katere zaporedja večinoma (skupno 92%) pripadajo skupinam 1, 2, 3, nekaj sekvenc tudi skupinam 5, 7 in 13. V knjižnici ni bilo genov za 16S rRNA predstavnikov skupin *Firmicutes*, *Chloroflexi* in *Bacteroides*, ki so jih Kraigher in sod. (2006) pridobili v knjižnici genov iz tal nizkega barja na Ljubljanskem barju, zato pa se je dodatno pojavila skupina *Spirochaetes*.

KEY WORDS DOCUMENTATION

- DN Dn
- UDC 579.66: 575.86: 556.56 (043.2) = 163.6
- CX Ljubljana Marsh, Sphagnum bog, phylogeny, bacteria community structure, gene library, 16S rRNA
- AU AUSEC, Luka
- AA MANDIĆ-MULEC, Ines (supervisor)/ŽGUR-BERTOK, Darja (reviewer)
- PP SI-1000 Ljubljana, Vecna pot 111
- PB University of Ljubljana, Biotechnical Faculty, Department of Biology
- PY 2008
- TI CONSTRUCTION AND ANALYSIS OF 16S rRNA CLONE LIBRARY FROM
Sphagnum BOG IN THE LJUBLJANA MARSH
- DT Gradation Thesis (Univesity studies)
- NO IX, 57 p., 7 tab., 11 fig., 48 ref.
- LA sl
- AL sl/en
- AB Marshlands are important ecosystems in which microorganisms play a vital role in the cycling of energy and matter. In spite of that, little information about bacterial diversity in marshland soils is available. We investigated acid (pH 4,5) *Sphagnum* bog soil in the Ljubljana marsh and constructed the 16S rRNA gene library from the soil sample. 154 partial sequences were retrieved and phylogenetic trees were inferred using the ARB program. *Acidobacteria* (41,6%) and *Proteobacteria* (40,2%) dominate the library while other groups were more scarcely represented: *Actinobacteria* (7,1%), *Plactomycetes* (5,2%), *Verrucomicrobia* (5,2%) in *Spirochaetes* (0,7%). Unusually numerous sequences belonging to the division Acidobacteria were placed mainly in groups 1, 2 and 3 (together 92% of the cloned sequences), the remainder of the sequences belonged to groups 5, 7 and 13. We did not find 16S rRNA sequences of *Firmicutes*, *Chloroflexi* or *Bacteroides*, which were previously retrieved from the Ljubljana Marsh fen soil by Kraigher et al. (2006). Group *Spirochaetes*, however, was present in the bog library but absent in the fen library.

KAZALO VSEBINE

KLJUČNA DOKUMENTACIJSKA INFORMACIJA	III
KEY WORDS DOCUMENTATION	IV
KAZALO VSEBINE	V
KAZALO PREGLEDNIC	VII
KAZALO SLIK	VIII
SEZNAM OKRAJŠAV	IX
1 UVOD	1
2 PREGLED OBJAV	3
2.1 2.1 KLONSKE KNJIŽNICE TALNIH EKOSISTEMOV	3
2.2 RAZISKAVE MIKROBNE ZDRUŽBE V BARJANSKIH TLEH.....	7
2.3 NAMEN DELA IN HIPOTEZA	9
2.3.1 Namen dela	9
2.3.2 Hipoteza	9
3 METODE IN MATERIALI	10
3.1 PRIPRAVA VZORCA TAL	10
3.2 IZOLACIJA DNA IN PREVERJANJE UČINKOVITOSTI IZOLACIJE	10
3.3 POMNOŽEVANJE DNA V VERIŽNI REAKCIJI S POLIMERAZO (PCR)...	12
3.4 ČIŠČENJE PRODUKTA VERIŽNE REAKCIJE S POLIMERAZO	13
3.5 LIGACIJA, TRANSFORMACIJA IN SELEKCIJA	14
3.6 PREGLEDOVANJE VSTAVLJENIH FRAGMENTOV Z VERIŽNO REAKCIJO S POLIMERAZO	15
3.7 SEKVENIRANJE IN PRIPRAVA PRIPRAVA SEKVENC	16
3.8 FILOGENETSKE ANALIZE SEKVENC DELOV GENOV ZA 16S RRNA... 17	
3.9 MATERIALI	18
3.9.1 Reagenti	18
3.9.2 Gojišča in dodatki	19
3.9.3 Kompleti in encimi	20
3.9.4 Pufri in raztopine	20
3.9.5 Uporabljeni začetni oligonukleotidi	21

4	REZULTATI.....	22
4.1	IZOLACIJA DNA, OCENA ČISTOSTI IN KONCENTRACIJE IZOLIRANE DNA	22
4.2	VERIŽNA REAKCIJA S POLIMERAZO IN ČIŠČENJE PRODUKTA	23
4.3	KLONIRANJE, IZBOR KLONOV IN SEKVENC.....	24
4.4	FILOGENETSKO DREVO	26
4.5	OPERACIJSKE TAKSONOMSKE ENOTE.....	28
4.6	PRIMERJAVA DVEH KNJIŽNIC GENOV ZA 16S RNA.....	30
4.7	ANALIZA SEKVENC IZ SKUPINE <i>Acidobacteria</i>	34
5	RAZPRAVA IN SKLEPI.....	35
5.1	RAZPRAVA.....	35
5.1.1	Artefakti in vzroki pristranskosti verižne reakcije s polimerazo (PCR)..	35
5.1.1.1	Pistranskost in težave pri prepoznavanju vrst po zaporedjih, pridobljenih s PCR	35
5.1.1.2	Artefakti, ki nastanejo pri verižni reakciji s polimerazo	39
5.1.1.2.1	Heterodupleksi.....	39
5.1.1.2.2	Himerne sekvence.....	40
5.1.1.2.3	Naključni dogodki	43
5.1.2	Druge napake in omejitve	45
5.1.3	Diskusija rezultatov	47
5.2	SKLEPI.....	49
6	POVZETEK.....	51
7	VIRI	53

KAZALO PREGLEDNIC

Preglednica 1: Sestava reakcijske mešanice (uporabili smo Taq polimerazo proizvajalca Promega).....	12
Preglednica 2: Pogoji verižne reakcije s polimerazo.....	13
Preglednica 3: Sestava reakcijske mešanice (uporabili smo Taq polimerazo proizvajalca Biotools)	15
Preglednica 4: pogoji verižne reakcije s polimerazo	16
Preglednica 5: Sekvence in tarčna mesta začetnih oligonukleotidov	21
Preglednica 6: število sekvenc v klonskih knjižnicah visokega (KG) in nizkega (HC) barja po skupinah.....	30
Preglednica 7: Število sekvenc v posamezni skupini acidobakterij (Barns in sod., 2007) ter deleži teh skupin v knjižnicah z visokega (KG) in nizkega (HC) barja. Ene sekvence iz knjižnice HC nismo mogli nedvoumno uvrstiti v nobeno od 26 predlaganih skupin.	34

KAZALO SLIK

Slika 1: Fotografija elektroforetskega gela, na katerem smo preverjali velikost izolirane DNA	22
Slika 2: Preverjanje produkta PCR in negativne kontrole na agaroznem gelu.....	24
Slika 3 – Primer rezultata verižne reakcije s polimerazo, kjer smo testirali 24 klonov. V prvi vrstici obeh stolpcev je velikostni standard (S), v zadnji vrstici desnega stolpca pa je negativna kontrola (N). Za nadaljnje delo so primerni vsi razen treh klonov (1, 2, 3).	25
Slika 4: Deleži skupin bakterij v klonski knjižnici visokega barja.....	26
Slika 5: Filogenetsko drevo knjižnice 16S rRNA iz vzorca tal visokega barja (KG), izračunano po metodi varčnosti.....	27
Slika 6: Spreminjanje števila operacijskih taksonomskih enot (OTU) pri različnih genetskih razdaljah sekvenc v naši knjižnici v odvisnosti od števila sekvenc.	29
Slika 7: Deleži bakterijskih skupin v klonskih knjižnicah visokega (KG) in nizkega (HC) barja.	31
Slika 8 – Skupno drevo obeh knjižnic – samo skupina <i>Proteobacteria</i> . Drevo je narejeno po metodi povezovanja sosedov (ang. <i>neighbour joining</i>). Sekvence HC so označene z arabskimi številkami, sekvence KG pa z rimskimi in črko ter z rumeno barvo.....	32
Slika 9 – Skupno drevo obeh knjižnic – skupine razen <i>Proteobacteria</i> . Drevo je narejeno po metodi povezovanja sosedov (ang. <i>neighbour joining</i>). Sekvence HC so označene z arabskimi številkami, sekvence KG pa z rimskimi in črko ter z rumeno barvo.....	33
Slika 10 – Shematičen prikaz nastanka himernih molekul DNA. (A) Vsaka veriga s prostim 3' koncem, ki je vsaj delno homologen kateri koli matrici iz vzorca, lahko deluje kot začetnik za polimerizacijo. (B), (C) Nastanek himer z zamenjavo matrične DNA med polimerizacijo.	41
Slika 11: Deleži, ki jih v knjižnici zavzemajo večje skupine bakterij, če upoštevamo surove sekvence (modri stolpci) oziroma če sekvence, ki so si med seboj manj kot 1% različne, združimo v 132 operacijskih taksonomskih enot (rdeči stolpci).	45

SEZNAM OKRAJŠAV

PCR	verižna reakcija s polimerazo (ang. <i>polymerase chain reaction</i>)
colony PCR	PCR, kjer namesto izolirane DNA dodamo suspendirano celično kulturo
HC	eksperimentalno polje na Ljubljanskem barju, kjer je v tleh veliko organskega ogljika in predstavlja nizko barje; to je predmet raziskovanja Kraigher in sod. (2006) in za nas referenčna študija; knjižnica HC pomeni klonsko knjižnico genov za 16S rRNA iz vzorca teh tal
KG	eksperimentalno področje Kozlarjeve gošče, od koder smo odvzeli vzorec tal za pričujoče delo; ta habitat predstavlja ostanek visokega barja na Ljubljanskem barju; knjižnica KG pomeni klonsko knjižnico genov za 16S rRNA iz vzorca teh tal
bp	bazni par, običajen način podajanja dolžine nukleotidnega zaporedja; kbp pomeni kilobaza, to je 1000 baznih parov
OTU	operacijska taksonomska enota (ang. <i>operational taxonomic unit</i>)

1 UVOD

Barja, kjer prevladuje šotni mah (*Sphagnum* sp.) in nizek pH, predstavljajo enega najboljšežnejših tipov mokrišč v Severni Ameriki in Evraziji (Dedysh in sod., 2005) in zavzemajo približno 3% kopne površine Zemlje (ponekod v Sibiriji celo do 80% površine). Barja zaradi nizke biorazgradnje skladiščijo približno 30% organskega ogljika (SOC, Dedysh in sod., 2005) in zato predstavljajo pomemben ponor ogljikovega dioksida. Barja vplivajo tudi na lokalno podobo in funkcije pokrajine, saj delujejo kot hidrološki regulatorji in predstavljajo pomembne zaloge vode. Barja so vroča točka biodiverzitete (Ljubljansko barje je vključeno v evropsko omrežje varstvenih območij Natura 2000) in imajo tudi neposredni pomen za človeka kot površine za športno udejstvovanje, lov in ribolov, ter so nam v estetsko zadovoljstvo.

Od nekdanj mogočnega Ljubljanskega barja, ki se je raztezalo na skoraj 20 tisoč hektarjih, je danes ostalo le še malo (Martinčič, 2003). Rezanje šote, gradnja prekopov in cest ter izsuševanje so ga do začetka 20. stoletja skrčili na manj kot 1400 hektarov, do prepovedi rezanja šote sredi 20. stoletja pa je ostalo le še 135 raztresenih hektarov šotnih površin. Danes o prvotnem barju skorajda ne moremo več govoriti (Martinčič, 2003), so se pa ohranili ostanki barjanskih tal nizkega barja, ki so danes večinoma spremenjena v rodovitne obdelovalne površine, in nekaj ostankov visokega barja. Ključna značilnost visokega barja je ombrotrofnost: barje vso vodo dobi s padavinami in ni povezano s podtalno vodo ali hudourniško vodo z okoliških bregov. Za visoka barja so značilne različne vrste šotnih mahov (*Sphagnum* sp.) in kisel pH tal.

Klub pomembni vlogi barja je o sestavi mikrobnih vrst tega ekosistema malo znanega. Redke študije, ki uporabljajo gojitvene in/ali molekularne metode, se osredotočajo na posamezne filogenetske skupine (na primer *Actinobacteria* in *Acidobacteria*) ali skupine mikrobov s skupno fiziološko oziroma metabolno značilnostjo, predvsem metanogene arheje. V tej nalogi smo zato skušali dodati košček sestavljanke s pripravo knjižnice genov za 16S rRNA iz šotnih tal.

Predmet naše raziskave predstavlja vzorec gozdnatih tal v Kozlarjevi gošči, ki je ostanek visokega barja z zelo visoko vsebnostjo nerazgrajene organske snovi in pH približno 4,8. Izolirali smo DNA ter pomnožili gene za 16S ribosomalno RNA (rRNA) z verižno reakcijo s polimerazo (PCR). Pomnožene fragmente smo vstavili v plazmidne vektorje ter z njimi transformirali kompetentne celice bakterije *E. coli*. Pridobili smo zaporedja vstavljenih fragmentov primernih transformant (klonov) ter s filogenetskimi metodami uvrstili te sekvence v glavne bakterijske skupine.

Rezultati predstavljajo vpogled v raznolikost bakterijskega sveta v vzorcu tal visokega barja in dajejo iztočnice za nadaljnje raziskovanje.

2 PREGLED OBJAV

2.1 2.1 KLONSKA KNJIŽNICA TALNIH EKOSISTEMOV

Preučevanje biotske raznovrstnosti organizmov ter njeno ohranjanje v naravnih ekosistemih postaja vedno pomembnejša tema znanstvenih raziskav, saj se danes zavedamo pomena raznovrstnosti odnosov med organizmi za stabilnost ekosistemov. Zaradi svoje številčnosti in metabolne (fiziološke) raznolikosti so mikrobi velikega pomena za kroženje snovi in energije, vendar jih zaradi njihove majhnosti in težavnega gojenja ne poznamo dovolj dobro. Molekularne metode so v zadnjih dveh desetletjih pokazale na izjemno veliko raznolikost bakterij, ki je bila s klasičnimi mikrobiološkimi gojitvenimi metodami povsem nedosegljiva. V čisti kulturi v laboratoriju lahko namreč vzgojimo le majhen delež bakterijskih vrst (pogosta ocena je manj kot 1%, Amann in sod., 1995). V barjanskih tleh, ki so v primerjavi z dobro prezračenimi tlemi glede bakterijske številčnosti nekoliko bolj skromne, je ocenjeno število bakterijskih celic med 10^8 in 10^9 na gram tal (Morales in sod. 2006), v različnih vzorcih tal pa je glede na teoretične modele med 2000 in 18000 različnih genomov (Dunbar in sod., 2002).

Molekularne metode so od sredine 80.-ih let 20. stoletja standardno orodje mikrobni ekologov pri njihovem spopadanju z neverjetno raznolikostjo mikrobnega sveta. Med glavne metode sodijo fluorescentna in situ hibridizacija (FISH), polimorfizem dolžine restrikcijskih fragmentov (T-RFLP), elektroforeza v denaturirajočem gradientu (DGGE) ter pomnoževanje s PCR, kloniranje in sekveniranje (Head in sod., 1998). Pri naši raziskavi smo uporabili slednjo metodo, ki je zaradi zapletenih in naključju podvrženih korakov podvržena številnim možnostim za napake in pristranskosti. Na nekatere opozarjajo raziskovalci že dlje časa (Head in sod., 1998), na primer na morebitne težave pri izolaciji DNA iz mirujočih stadijev bakterij, na dejstvo, da pristranska izbira začetnih oligonukleotidov za PCR nujno prinese pristransko pomnoževanje DNA in podobno.

V zadnjem času pa je nekaj raziskovalcev z dobro načrtovanimi eksperimenti osvetlilo nekatere hipotetične probleme, ki dotlej še niso bili sistematično podvrženi preiskovanju. Tako so na primer pokazali, da v splošnem ne moremo trditi, da razmerja klonov v klonski knjižnici odražajo razmerja izvorne DNA v vzorcu (Polz in Cavanaugh, 1998; Kurata in sod., 2004) in da je mahne klonske knjižnice (nekaj sto klonov) tudi zaradi tega težko primerjati (Schloss in Handelsman, 2005). Običajno lahko reproduciramo veliko vrstno bogastvo, če naredimo na primer dve knjižnici iz istega vzorca, ne moremo pa pričakovati ponovljivosti posameznih vrst oziroma sekvenc (Dunbar in sod., 2002). Natančneje so se tudi lotili odkrivanja molekularnih mehanizmov nastanka artefaktov pri verižni reakciji s polimerazo, na primer tvorbe heterodupleksov in himernih molekul, ter možnosti za njihovo odkrivanje in preprečevanje (Kanagawa, 2003; Hugenholtz in Huber, 2003, Acinas in sod., 2005). Tovrstne napake in pristranskosti metode nadrobneje obravnavamo v Razpravi.

Kljub omenjenim problemom je metoda pomnoževanja s PCR, kloniranja in sekveniranja dosti v rabi, saj v primerjavi z drugimi metodami daje natančnejše informacije o kloniranih fragmentih DNA, zato je njihova napovedna moč večja. Prve klonske knjižnice različnih talnih habitatov so iz 90-ih let prejšnjega stoletja. Borneman in sod. (1996) so preučevali vzorec tal iz pašnika v ZDA in izolirali 124 klonov. Večinoma so jih uvrstili med *Proteobacteria* (16,1%) in širši skupini *Cytophaga-Flexibacter-Bacteroidetes* (22%) in bakterije z nizkim deležem G-C (22%). Čeprav je bilo takrat v bazah podatkov že 6500 zaporedij gena za 16S rRNA (večinoma je šlo za kratke in nepopolne sekvence), avtorji številnih sekvenc niso mogli uvrstiti v večje in dobro poznane skupine bakterij (Borneman in sod., 1996), saj se je takrat razkrivanje biotske raznovrstnosti bakterij šele dobro začelo. McCaig in sod. (1999) so naredili šest manjših knjižnic, s katerimi so želeli pokazati morebitne razlike med negnojenim in gnojenim pašnikom na Škotskem. Med skupno 275 delnimi sekvencami so prepoznali nekatere, ki so jih pripisali na novo porajajočim se skupinam *Acidobacteria* in *Verrucomicrobia*, približno 50% pa so jih opredelili kot *Proteobacteria*. Avtorji zaključujejo, da je raznovrstnost bakterij v tleh zelo velika in da jo s tako majhno klonsko knjižnico ni možno zaobjeti, zato tudi niso mogli statistično razlikovati med vrstnim bogastvom obeh tipov pašnika. V klonski knjižnici iz tal v Amazonskem nižavju so našli manj kot 20% *Proteobacteria*, zato pa so toliko več sekvenc

uvrstili v rodova *Bacillus* in *Clostridium* ter v novo skupino *Verrucomicrobia* (Borneman in Triplett, 1997). Nekaj zaporedjem niso mogli določiti filogenetske pripadnosti nobeni izmed poznanih bakterijskih skupin, saj so bile takrat baze še veliko bolj nepopolne kot danes. Z metodo RISA (ang. RNA intergenic spacer analysis) so uspeli celo pokazati, da se bakterijski združbi v zrelem gozdu in pašniku ob njem razlikujeta. Schloss in Handelsman (2005) sta kasneje pokazala, da je klonske knjižnice s po približno 100 kloni izredno težko primerjati med sabo; tako na primer nista mogla statistično pokazati razlike v vrstnem bogastvu »bogate« zemlje iz amazonskega pragozda in »navadnega« škotskega pašnika, ker je za to enostavno premalo podatkov.

Pozna devetdeseta leta prejšnjega stoletja so čas, ko so v različnih študijah kar naprej opazali nove, dotlej neznane skupine bakterij. Še en tak primer so plitvi in rahlo kisli sladkovodni ribniki v Karolini (Wise in sod., 1997), kjer so v sicer dokaj skromni knjižnici 35 sekvenc našli 11 iz skupine *Proteobacteria*, 8 iz skupine *Acidobacteria* in 7 iz skupine *Verrucomicrobia*. Kopičenje tovrstnih podatkov je pripomoglo k natančnejši filogenetski karakterizaciji teh skupin (Hugenholtz in sod., 1998). *Acidobacteria* so leta 1999 prepoznali za raznoliko in široko razširjeno novo bakterijsko deblo (Barns in sod., 1999), kasnejša revizija pa je pokazala (Quaiser in sod., 2003), da kljub naraščajoči količini podatkov in posledično razraščajočem filogenetskem drevesu o življenju in funkciji teh organizmov vemo zelo malo. Logična posledica je bila tako izboljšava metod izolacije in kultivacije teh bakterij (Janssen in sod., 2002, Sait in sod., 2005, Dedysch in sod., 2006), nekaj namigov o njihovih fizioloških zmožnostih in ekološkem pomenu pa smo dobili tudi s študijami različnih ekstremnih okolij, na primer vročih vrelcev, bioreaktorjev čistilnih naprav in rudnikih radioaktivnih elementov (Quaiser in sod., 2003, Barns in sod., 2007), kjer se bakterije te skupine množično pojavljajo. Ker so se *Acidobacteria* izkazale za pomembno skupino tudi v pričujoči raziskavi barjanskih tal, o njih nekoliko podrobneje pišemo v Razpravi.

Obstaja še nekaj študij celotne združbe bakterij v tleh ali sladkovodnih habitatih. Graff in Conrad (2005) sta ugotavljala filogenetsko sestavo bakterijske združbe v večkrat poplavljenih tleh v sestoji topola. Pripravila sta šest majhnih knjižnic (po približno 55 klonov) iz vzorca zemlje, področja korenin (ang. *rhizosphere*) ter površine korenin (ang.

rhizoplane), in sicer ločeno za poplavljenja in nepoplavljenja tla. Rezultati nakazujejo, da se šest knjižnic med seboj razlikuje, kar nakazujejo tudi različni profili vzporedne T-RFLP analize. V poplavljenih tleh se število restrikcijskih fragmentov zmanjša. Skupno predstavljajo polovico vseh izoliranih klonov *Proteobacteria*, sledijo jih skupine *Bacillales*, *Actinobacteria* in *Acidobacteria*. V klonski knjižnici iz nizkega evtrofnega jezera na Japonskem (Tamaki in sod., 2005) pa so denimo našli 47% *Proteobacteria*, sledijo jim skupine *Nitrospira* (13%), *Acidobacteria* (8%) in *Bacteroidetes* (6%). Avtorji poudarjajo, da so potrebne tudi študije, ki so zasnovane na gojitvenih metodah, saj nam zgolj filogenetska uvrstitev ne da informacije o fiziologiji organizma (Tamaki in sod., 2005). Poročajo o uspešnejši izolaciji skupin *Planctomyces* in *Bacteroidetes* na alternativnem gojišču (»gellan gum« namesto agarja), acidobakterij pa denimo niso uspeli vzgojiti (Tamaki in sod., 2005).

V številnih študijah tako prevladujejo *Proteobacteria* (povprečno 39%) in *Acidobacteria* (povprečno 20%), pogoste so še *Actinobacteria*, *Firmicutes*, *Verrucomicrobia*, *Bacteroidetes*, *Chloroflexi*, *Planctomycetes* in *Gemmatimonadetes* (Janssen, 2006). Teh devet debel predstavlja povprečno 92% vseh sekvenc različnih klonskih knjižnic (Janssen, 2006). Splošno znani pomislek zadeva pristranskost PCR reakcije, ki nastane zaradi izbora začetnih oligonukleotidov (Head in sod., 1998); proteobakterije so v bazah podatkov najbolj zastopane, zato imajo večjo težo pri njihovem načrtovanju. Zato so zanimivi precej drugačni rezultati nizozemskih raziskovalcev (Felske in sod., 1998), ki so se lotili raziskovanja vzorca tal travnika z izolacijo ribosomov, da bi pokazali, katere bakterije so najbolj aktivne v tleh. Iz ribosomov so nato z reverzno transkriptazo pomnožili dele genov za 16S rRNA ter pregledali še klonsko knjižnico. 65% ribosomov je izviralo iz skupine *Firmicutes*, tudi v klonski knjižnici so bili predstavniki rodu *Bacillus* najštevilčnejše zastopani, zato avtorji ugotavljajo, da so to verjetno najbolj aktivne bakterije v preučevanih rahlo kislih travniških tleh (Felske in sod., 1998).

Sodoben pristop k raziskovanju bakterijskih združb je metagenomika (Handelsman, 2004). V umetne bakterijske kromosome (BAC) lahko vstavimo zelo velike kose DNA (tipično velike 100 kbp), kar nam poleg filogenetske umestitve izvornega organizma omogočajo tudi dostop do nekaterih drugih funkcionalnih genov. Heterologna ekspresija genov je

zadosti pogosta, da opravičuje testiranje velikega števila klonov pri iskanju antibiotikov in biotehnološko pomembnih encimov (Handelsman, 2004). Tako so tudi razkrili drobec fiziologije dotlej nekultivirane acidobakterije (Liles in sod., 2003). Kljub izboljšanju metod gojenja mikrobov in novim pristopom k pridobivanju molekularnih podatkov pa smo še daleč od možnosti, da bi lahko dostojno ocenili razsežnost bakterijske raznolikosti (Schloss in Handelsman, 2004). Nekatere skupine so v bazah podatkov dokaj dobro zastopane, na primer *Gammaproteobacteria*, pri številnih skupinah pa ne moremo določiti niti, koliko sekvenc bi še potrebovali, da bi lahko vsaj ocenili vrstno bogastvo te skupine (Schloss in Handelsman, 2004). Naraščanje baz podatkov razkriva vedno nove skupine bakterij, število debel je glede na molekularne podatke že večje od 50, čeprav jih polovica še nima gojenih predstavnikov (Schloss in Handelsman, 2004). Zaradi obsežnosti je raznolikost bakterij težko raziskovati in trenutno smo še daleč od popisa vrst v posamezni skupini bakterij ali v katerem od ekosistemov.

2.2 RAZISKAVE MIKROBNE ZDRUŽBE V BARJANSKIH TLEH

O bakterijski združbi v barjanskih tleh je malo znanega. Nemška raziskava (Rheims in sod., 1996) se osredotoča predvsem na nekaj novih skupin aktinomicet, sicer pa je bilo nekaj fizioloških (in deloma filogenetskih) raziskav opravljenih na metanotrofnih bakterijah in metanogenih arhejah. Celovito so se tega vprašanja lotili v Rusiji (Dedysh in sod., 2006), ki so preučevali kislja (pH 3,9 do 4,5) barjanska tla s šotnim mahom iz zahodne Sibirije. Poleg 16S rRNA klonske knjižnice so tudi preštevali bakterije s fluorescentno in situ hibridizacijo (FISH) ter uprabili nove gojitvene metode (bogatitev v biofilmih). Največ sekvenc iz njihove knjižnice (skupaj 84 klonov) so pripisali *Acidobacteria* (28,6%) in *Protobacteria* (28,6%), sledijo *Verrucomicrobia* (15,5%) ter *Actinobacteria*, *Planctomyces*, *Bacteroides* in *Chlorobi* s po nekaj sekvencami. *Acidobacteria* so razvrstili v skupine 1 (najštevilčnejša), 3, 4 in 8, s svojo gojitveno metodo pa so vzgojili tudi nekaj kokultur, ki vsebujejo predstavnike *Acidobacteria* in *Planctomyces*; ta metoda lahko pomeni korak naprej pri spoznavanju fiziologije acidobakterij. Zanimivi so rezultati štetja celic: bakterijske celice so predstavljale največ 18% ($1,0-1,5 \times 10^8$ na gram vlažnih

tal), arheje pa do 8% celotnega števila celic. Ugotavljajo, da glavnino celic tvorijo zelo majhne celice (<0,5 nm), ki imajo 16S rRNA podobno tisti pri arhejah. Skupina Alphaproteobacteria je bila najštevičnejša ($1,1 \times 10^7$ na gram vlažnih tal), ostale proteobakterije ter Firmicutes in Bacteroidetes pa imajo v šotnih tleh majhne populacije (velikostni red 10^5 celic na gram vlažnih tal), kar se sklada z rezultati njihove klonske knjižnice. Nasprotno pa je Acidobacteria, ki so v knjižnici najbolj zastopane, v tleh le 10^6 na gram vlažnih tal, pri čemer se zaporedje označevalca popolnoma prilega vsem sekvencam. Razlog je lahko v preferenčnem pomnoževanju genov Acidobacteria, v zelo dobri lizi teh celic med izolacijo DNA iz vzorca tal ali pa je glavnina celic te skupine prisotna v dormantnem stadiju. Avtorji se do teh hipotez ne morejo opredeliti (Dedysh in sod., 2006).

V raziskavi 24 šotnih barj v Ameriki je raziskovalce (Morales in sod., 2006) zanimala struktura bakterijske združbe v barjanskih tleh ter morebitne razlike glede na geografsko lokacijo in/ali profil tal. V majhni klonski knjižnici (64 sekvenc) so prevladovali Proteobacteria (55%), večji skupini sta bili še Acidobacteria in Firmicutes (po 11%). Raziskovalci ugotavljajo, da v tako skromni raziskavi ne morejo zajeti vrstnega bogastva barja. Zanimivi so tudi njihovi rezultati T-RFLP analize, ki kažejo na medsebojno podobnost vseh 24 površinskih vzorcev. Nasprotno pa so opazili razlike med površinskimi vzorci ter tistimi z globine enega metra, kar kaže na to, da je bakterijska združba v tleh stratificirana.

Kraigher in sod. (2006) so na Ljubljanskem barju raziskovali mikrobo aktivnost in bakterijsko združbo na ostankih nizkega barja. Preučevali so tla z visoko (HC) in nizko vsebnostjo ogljika (LC) in med njima niso zaznali večjih razlik s T-RFLP profiliranjem, zato pa so s tehniko substratno inducirane respiracije (SIR) pokazali, da je respiracija v vzorcu HC tal vedno višja. Polovico sekvenc iz knjižnice (skupno 114 sekvenc) so pripisali Proteobacteria, približno četrtino Acidobacteria, druge skupine so bile skromneje zastopane (Preglednica 6). Izsledke te raziskave podrobneje predstavljamo v poglavjih Rezultati in Razprava, saj nam je pri našem delu služila za izhodišče in je bila primerjava med to raziskavo in našimi podatki eden glavnih namenov tega dela.

2.3 NAMEN DELA IN HIPOTEZA

2.3.1 Namen dela

Barja so pomemben ekosistem in mikrobi igrajo v njem ključno vlogo. Kljub temu je o sestavi bakterijske združbe v barjanskih tleh malo znanega. Želeli smo razširiti dognanja Kraigher in sod. (2006), ki so z molekularnimi metodami preiskovali sestavo bakterijske združbe na nizkem barju, s pripravo nove klonske knjižnice iz vzorca tal iz Kozlarjeve gošče, ki predstavlja ostanek visokega barja. Hoteli smo primerjati oba filogenetska profila, zato smo se pri konstrukciji klonske knjižnice držali uporabljenih postopkov dela. S teoretično obravnavo uporabljene metode smo želeli oceniti smiselnost takega raziskovanja ter podati morebitne izboljšave za prihodnje delo. S temeljito filogenetsko analizo pa smo dokazali prisotnost nekaterih novih in zanimivih skupin bakterij, ki so prisotne v preiskovanih tleh in ki so lahko dobra iztočnica za nadaljnje raziskovanje.

2.3.2 Hipoteza

Skušali smo potrditi osnovno hipotezo, da se bo združba v obravnavanih tleh visokega barja v Kozlarjevi gošči razlikovala od tiste iz nizkega barja zaradi drugačnih fizikalno-kemijskih pogojev tal.

Osnovna vprašanja raziskave:

1. Kakšna je filogenetska sestava bakterijske združbe v tleh visokega barja?
2. So v tleh prisotne kake zanimive skupine bakterij, ki bi jih kazalo v prihodnje podrobneje raziskati?
3. Je splošno uveljavljeni protokol izdelave klonske knjižnice, ki smo ga uporabili, ustrezen, ali bi ga v prihodnje kazalo kako modificirati?
4. Kako majhna klonska knjižnica odraža resnično bakterijsko združbo v tleh?

3 METODE IN MATERIALI

3.1 PRIPRAVA VZORCA TAL

Klonsko knjižnico smo pripravili iz vzorca tal, ki ga je 25. avgusta 2003 pripravila dr. Barbara Kraigher v sklopu svojih raziskovanj na Ljubljanskem barju (Kraigher in sod., 2006): v Kozlarjevi gošči je na področju približno 3 x 4 metre odvzela 20 talnih sredic (podvzorcev), ki jih je v laboratoriju homogenizirala skozi z etanolom sterilizirano sito (premer luknjic 4 mm) in jih temeljito premešala. Sterilne mikrocentrifugirke je napolnila s po 0,5g homogeniziranih tal in jih shranila na -20 °C.

Septembra 2005, ko smo z delom začeli, smo za izolacijo DNA porabili štiri izmed tako shranjenih vzorcev.

3.2 IZOLACIJA DNA IN PREVERJANJE UČINKOVITOSTI IZOLACIJE

Skupno mikrobno DNA smo izolirali s kompletom UltraClean Soil DNA Isolation Kit (proizvajalec MO BIO Laboratories, Solana Beach, California). Delali smo v štirih paralelkah s po 0,5g vzorca tal po navodilih proizvajalca. Po končanem protokolu smo tako dobili 4 mikrocentrifugirke z izolirano DNA, ki smo jo nato združili, premešali, ponovno razdelili v 4 mikrocentrifugirke (v vsaki je 40 µl raztopine DNA) ter jih shranili na -20 °C.

Izolirano DNA smo preverili na 1% (wt/vol) agaroznem gelu. Želeli smo oceniti količino in velikost izoliranih fragmentov DNA.

Agarozne gele smo vedno pripravljali na enak način. V 30 ml 1x TAE pufru smo zatehtali 0,30 g agaroze ter segrevali v mikrovalovni pečici, dokler se agarosa ni raztopila. Nato

smo dodali 1,5 μ l etidijevnega bromida (založna koncentracija 0,4 μ g/ml) in razlili gel v elektroforetsko kadičko. Najprej smo na gel nanegli 2 kilobazno lestvico standardov (GeneRulerTM DNA Ladder Mix, MBI Fermentas, Litva), nato vse vzorce ter na koncu negativno kontrolo (produkt PCR, ki mu nismo dodali vzorčne DNA). Po 3 μ l vzorcev in kontrole smo pred nanosom na gel obarvali s po 1 μ l pufru za nanašanje na gel.

Oceno čistosti izolirane DNA lahko določimo na več načinov (Promega Protocols and Applications guide, chapter 9-5: DNA Purification):

- z merjenjem optične gostote (absorbance),
- z elektroforezo na agaroznem gelu,
- z vezavo fluorescentnih barvil na DNA,
- z merjenjem sklopitve luciferaze in pirofosfataze.

Mi smo zaradi enostavnosti in razpoložljivosti reagentov in aparatur izbrali prvo metodo, ki je tudi sicer najpogosteje uporabljena. Rastopino DNA iz ene izmed zamrznjenih mikrocentrifugirk smo razdelili v dve paralelki in jo redčili za faktor 3,5 (končni volumen vzorca v merilni kivetki spektrofotometra je bil tako 70 μ l). Izmerili smo absorbanco rastopine DNA pri 230, 260 ter 280 nm (A_{230} , A_{260} , A_{280}). Meriti moramo pri več valovnih dolžinah, ker pri 260 nm, kjer DNA sicer absorbira največ, absorbira tudi RNA, znaten delež pa lahko prispevajo tudi aromatske aminokisliline v prisotnih beljakovinah, ki imajo sicer vrh absorbcije pri 280 nm. Zato je najobičajnejše merilo čistosti razmerje med absorbcijo pri 260 in 280 nm (A_{260}/A_{280}); idealna vrednost tega razmerja je med 1,7 in 2,0. Razmerje A_{260}/A_{230} pa nam služi kot ocena prisotnosti kaotropnih soli in organskih spojin (npr. gvanidin) v rastopini DNA (optimalno je vrednost nad 1,5).

Koncentracijo izolirane DNA smo ocenili s pomočjo enačbe (Promega Protocols and Applications guide, chapter 9-5: DNA Purification):

koncentracija [μ g/ml] = ($A_{260} - A_{320}$) x faktor redčenja x 50 μ g/ml

Vrednost A320 je merilo turbidnosti vzorca in podaja absorbanco, ki ne nastane zaradi nukleinskih kislin. Ker tega podatka nismo izmerili, sem ga pri zgornji formuli izpustil, zato je izračunana koncentracija DNA najbrž rahlo precenjena.

3.3 POMNOŽEVANJE DNA V VERIŽNI REAKCIJI S POLIMERAZO (PCR)

Celokupna bakterijska DNA iz vzorca tal nam je služila kot matrika za pomnoževanje genov za 16S ribosomalno ribonukleinsko kislino (rRNA). Uporabili smo univerzalne začetne oligonukleotide, in sicer 27f in R1495, katerih tarča so geni za 16S rRNA bakterij (Weisburg s sod., 1991).

Reagenti za pripravo reakcijske mešanice, njihove založne koncentracije ter končne koncentracije, so razvidni iz spodnje tabele.

Preglednica 1: Sestava reakcijske mešanice (uporabili smo Taq polimerazo proizvajalca Promega)

	založna konc.	končna konc.	skupni volumen [μ l]	volumen na reakcijo [μ l]
MiliQ			117,95	16,85
PCR pufer	10x	1x	17,50	2,50
MgCl₂	25mM	3,5 mM	14,00	2,00
Formamid	100%	1%	1,75	0,25
BSA	20 mg/ml	0,40 mg/ml	3,50	0,50
dNTP	10 mM	200 μ M	3,50	0,50
27f	10 μ M	200 nM	3,50	0,50
1495R	10 μ M	200 nM	3,50	0,50
taq polimeraza	5000 U/ml	1 U	2,80	0,40
		skupaj:	168,00	24,00

Pripravili smo 7 reakcij, to je 6 pomnoževalnih, v katere smo 24 μ l mešanice dodali po 1 μ l vzorčne DNA, ter eno negativno kontrolo brez dodane DNA. Pomnoževali smo po sledečem protokolu:

Preglednica 2: Pogoji verižne reakcije s polimerazo

		temperatura [°C]	čas [s]
začetna denaturacija		94	300
pripravljalni cikli (10x)	denaturacija	94	60
	vezava olig. začetnikov	58-53	60
	pomnoževanje	72	90
	denaturacija	94	60
cikli (25 x)	vezava olig. začetnikov	53	60
	pomnoževanje	72	90
	končno podaljševanje	72	600
	ohlajevanje	4	~

Uspešnost pomnoževanja smo ugotavljali z gelsko elektroforezo. Na 1% (wt/vol) agarozni gel z dodanim etidijevim bromidom smo nanegli lestvico standardov, po 3 μ l vsakega od obarvanih vzorcev (jamice 2-7) ter negativno kontrolo v 8. jamico.

3.4 ČIŠČENJE PRODUKTA VERIŽNE REAKCIJE S POLIMERAZO

Produkt pomnoževanja smo očistili s kompletom Wizard SV Gel and PCR Clean-up System (cat. # A9281, Promega, Madison, Wisconsin) po navodilih proizvajalca. Uporabili smo 1% gel iz agaroze z znižano temperaturo tališča (ang. *low-melt agarose*), celotni produkt smo razdelili v 4 jamice na gelu (približno 30 μ l produkta in 6 μ l barvila na

jamico). Ker so bili izrezani koščki gela z DNA dovolj majhni, smo lahko po dva čistili hkrati (skupna masa dveh koščkov ni presegala 350 mg). Obe očiščeni frakciji smo nato preverili na 1% (wt/vol) agaroznem gelu (nanesli smo po 3 μ l vzorca). Po čiščenju smo tako dobili skupno 70 μ l v vodi raztopljene DNA.

3.5 LIGACIJA, TRANSFORMACIJA IN SELEKCIJA

Pomnožene fragmente DNA smo vstavili v pGEM-T Easy plazmidne vektorje po navodilih proizvajalca (Promega, Madison, Wisconsin; pGEM-T and pGEM-T Easy Vector Systems). V ligacijsko mešanico smo dali 3 μ l očiščenega produkta pomnoževanja ter inkubirali preko noči na 4 °C.

Uporabili smo visoko kompetentne celice *E. coli*, prilagojene za ampicilinsko ter modro-belo selekcijo. Transformacijo smo izvajali v dveh paralelkah, v vsako smo dodali po 90 μ l kompetentnih celic ter po 2 μ l ligacijske mešanice. Transformacijo smo izvedli po protokolu proizvajalca plazmidnega vektorja (Promega, Madison, Wisconsin; pGEM-T and pGEM-T Easy Vector Systems) s temperaturnim šokom pri 45 °C. Po navodilih smo pripravili tekoči hranilni medij za transformacijo (SOC).

Po končani transformaciji smo celice nacepili na gojišče za selekcijo transformiranih kolonij. Pripravili smo približno 15 plošč (0,5 litra) LB gojišča z ampicilinom, IPTG in X-Gal; pomembno je, da dodamo slednje tri sestavine šele po avtoklaviranju. Plošče smo inkubirali preko noči na 37 °C, nato pa izvršili selekcijo. Vse transformirane celice, ki so prejele plazmid, so z njim prejele tudi rezistenco na antibiotik penicilin v gojišču, zato so sploh lahko zrasle na selektivnem gojišču z ampicilinom. Take celice dopolnijo svoj okvarjeni genom s še enim genom, to je lacZ (β -galaktozidaza). Tako transformirane celice postanejo sposobne izrabljati X-Gal (5-bromo-4-kloro-3-indolil- β -D-galaktozid), ki se pretvori v modro obarvan produkt. V primeru, da je v prejetem plazmidnem vektorju

tudi vstavljen fragment pomnožene DNA, pa se gen za lacZ prekine in take kolonije ostanejo bele (modro-bela selekcija).

Bele kolonije smo precepili na nove plošče z enakim gojiščem ter jih ponovno inkubirali preko noči na 37°C. Nato smo kolonije nacepili v mikrotitrne plošče s tekočim gojiščem LB z ampicilinom (po 150 µl gojišča v vsaki luknjici) ter jih preko noči inkubirali na 37°C. S sterilnim »glavnikom« smo nato podvojili plošče; izvornim ploščam smo v luknjice dodali po 60 µl 50% sterilnega glicerola ter jih shranili pri -20°C, kopije pa smo pustili rasti preko noči na 37°C. Kolonije na teh ploščah smo uporabili za preverjanje velikosti vstavljenih fragmentov z verižno reakcijo s polimerazo (ang. *colony PCR*). Do eksperimenta smo jih prav tako hranili na -20°C, prekrite z glicerolom.

3.6 PREGLEDOVANJE VSTAVLJENIH FRAGMENTOV Z VERIŽNO REAKCIJO S POLIMERAZO

Velikost vstavljenih fragmentov smo pregledovali z verižno reakcijo s polimerazo, kjer smo za matrično DNA dodali kar po 2 µl suspendirane kulture transformant (ang. *colony PCR*). Hkrati smo preiskovali po 24 kolonij. Sestava reakcijske mešanice in pogoji verižne reakcije so podani v tabelah 3 in 4.

Preglednica 3: Sestava reakcijske mešanice (uporabili smo Taq polimerazo proizvajalca Biotools)

	založna konc.	končna konc.	skupen volumen [µL]	volumen na reakcijo [µL]
MiliQ			447,50	17,90
PCR pufer	10x	1x	62,50	2,50
MgCl ₂	25mM	3,5 mM	25,00	1,00
dNTP	10 mM	200 µM	12,50	0,50
T7	10 µM	200 nM	12,50	0,50
SP6	10 µM	200 nM	12,50	0,50
taq polimeraza	5000 U/ml	0,5 U	2,50	0,10
		skupaj:	575,00	23,00

Preglednica 4: pogoji verižne reakcije s polimerazo

		temperatura [°C]	čas [s]
začetna denaturacija		94	300
cikli (25 x)	denaturacija	94	60
	vezava olig. začetnikov	53	60
	pomnoževanje	72	90
zaključno pomnoževanje		72	600
ohlajevanje		4	~

Uporabili smo oligonukleotidDNA začetnika T7 in SP6, ki nalegata na začetke poliklonskega mesta na plazmidnem vektorju pGEM-T Easy in tako pomnožijo vstavljeni fragment DNA.

Velikost vstavljenih fragmentov smo ocenili z gelsko elektroforezo. Na 1% agarozni gel smo nanašali preiskovane vzorce v dveh stolpcih, vsakokrat najprej velikostni standard (GeneRuler™ DNA Ladder Mix, MBI Fermentas, Litva), nato vzorce, kot zadnjo v desnem stolpcu smo nanegli negativno kontrolo.

Pričakovana velikost kloniranih odsekov gena za 16S rRNA je bila 1500 baznih parov. Klone smo na koncu zbrali na dveh mikrotitrskih ploščah ter jih poslali v sekveniranje v podjetje Macrogen.

3.7 SEKVENIRANJE IN PRIPRAVA PRIPRAVA SEKVENC

Sekveniranje so izvedli v korejskem podjetju Macrogen z začetnim oligonukleotidom R1495. Dobljene sekvence smo preverili s programom Chromas (Version 2.3, dostopno na <http://www.technelysium.com.au/chromas.html>). Odstranili smo po nekaj deset baznih parov na začetku in koncu sekvenc, kjer je bilo sekveniranje nezanesljivo, ter shranili sekvence v

enotni datoteki v FASTA formatu za nadaljnjo obdelavo. Sekvence so bile dolge med 760 in 800 bp.

Iz nadaljnje obravnave smo najprej izključili tiste sekvence, za katere smo menili, da so lahko himerne in kot take artefakt, ki nastane pri verižni reakciji s polimerazo. Sekvence smo najprej pregledali s spletnim programom Chimera_Check verzija 2.7 (Chimera detection ..., 2004). Potencialne himere smo preverili s programom Pintail (Ashelford s sod., 2005), kjer smo za primerjalno zaporedje vzeli kar tisto, ki ga je BLAST (<http://www.ncbi.nlm.nih.gov/blast/>) našel kot najbližjega preiskovanemu zaporedju. Poleg tega smo poravnane sekvence tudi prerezali na pol in iz polovic naredili filogenetski drevesi (ang. *partial treeing analysis*, Hugenholtz in Huber, 2003). Če sta se dela iste sekvence na drevesih drugače razporejala, je bil to dodatni pokazatelj, da gre morda za himero.

3.8 FILOGENETSKE ANALIZE SEKVENC DELOV GENOV ZA 16S RRNA

Filogenetske analize smo naredili v programu ARB (Ludwig in sod., 2004). Sekvence smo poravnali z integriranim programom FastAlign, pri čemer smo uporabljali dobro urejeno bazo približno 55 tisoč sekvenc, ki smo jo dobili na uradni strani programa ARB (<http://www.arb-home.de/>). Pri preučevanju acidobakterij smo uporabili bazo 2222 dobro urejenih zaporedij 16S rRNA te skupine, ki smo jo dobili v bazi RDP (Ribosomal Database Project, Cole in sod., 2007). Filogenetska drevesa smo naredili z metodo varčnosti (ang. *maximum parsimony*) in metodo povezovanja sosedov (ang. *neighbour joining*). Za korenino drevesa smo uporabili sekvenco *Aquifex aeolicus*, saj večina filogenetskih analiz kaže, da se se *Aquificales* razvejijo zelo globoko v filogenetskem drevesu bakterij (Hugenholtz in sod., 1998).

Pri pregledovanju knjižnice nas je zanimalo tudi, kako so si sekvence med seboj podobne, zato smo jih združili v operacijske taksonomske enote (OTU). Uporabili smo program DOTUR (Schloss in Handelsman, 2005) in vanj vnesli distančno matriko, ki smo jo

izračunali v programu ARB; pri računanju OTU smo uporabili metodo najbolj oddaljenih sosedov (ang. *furthest neighbour*), kot to predlagajo avtorji članka (Schloss in Handelsman, 2005). Najbolj nas je zanimalo, koliko različnih OTU bomo dobili pri razdaljah 0,005 in 0,03 (slednja razdalja tipično pomeni razliko med vrstami, Gevers in sod., 2005).

3.9 MATERIALI

3.9.1 Reagenti

Agaroz	Sigma-Aldrich, Steinheim, Nemčija
Bromfenol modro	Sigma-Aldrich, Steinheim, Nemčija
BSA	Sigma-Aldrich, Steinheim, Nemčija
dNTP set	MBI Fermentas, Litva
EDTA	Fluka, St. Gallen, Švica
Etanol	Merck KgaA, Darmstadt, Nemčija
Etidijev bromid	Sigma-Aldrich, Steinheim, Nemčija
Formamid	Merck KgaA, Darmstadt, Nemčija
GeneRuler™ DNA Ladder Mix	MBI Fermentas, Litva
Smart Ladder 0,2-10 kbp	Nippon Gene, Japonska
Glicerol	Merck KgaA, Darmstadt, Nemčija
Glukoza	Kemika, Zagreb, Hrvaška
KCl	Merck KgaA, Darmstadt, Nemčija
Ledocetna kislina	Merck KgaA, Darmstadt, Nemčija
Tris Base	Merck KgaA, Darmstadt, Nemčija

3.9.2 Gojišča in dodatki

Bakteriološki agar	Biolife, Milano, Italija
Bacto tripton	Biolife, Milano, Italija
Kvasni ekstrakt	Biolife, Milano, Italija
Ampicilin	Sigma-Aldrich, Steinheim, Nemčija
X-gal	Eppendorf, Hamburg, Nemčija

Trdno LB gojišče z Amp, X-gal, IPTG:	Bacto tripton	5 g
	Kvasni ekstrakt	2,5 g
	NaCl	2,5 g
	agar	7,5 g
	Ampicilin (100 mg/ml)	0,5 ml
	X-gal (50 mg/ml)	0,8 ml
	IPTG (0,1M)	2,5 ml
	dH ₂ O	do 500 ml

Tekoče LB gojišče z Amp:	Bacto tripton	10 g
	Kvasni ekstrakt	5 g
	NaCl	5 g
	Ampicilin (100 mg/ml)	1 ml
	dH ₂ O	do 1000 ml

SOC gojišče (100 ml):	Bacto tripton	2 g
	Kvasni ekstrakt	0,5 g
	1M NaCl	1 ml
	1M KCl	0,25 ml
	2M raztopina Mg ²⁺	1 ml
	2M glukoza	1 ml
	dH ₂ O	do 100 ml

2M raztopina Mg^{2+} (100 ml)
20,33 g $MgCl_2 \cdot 6H_2O$
24,65 g $MgSO_4 \cdot 7H_2O$
dH₂O do 100 ml

3.9.3 Kompleti in encimi

Ultra Clean Soil DNA Isolation Kit	MOBIO, CA, ZDA
Wizard [®] SV Gel and PCR Clean-Up System	Promega, Madison, WI, ZDA
pGEM [®] -T Easy Vector Cloning Kit	Promega, Madison, WI, ZDA
<i>Taq</i> DNA polimeraza	Promega, Madison, WI, ZDA
<i>Taq</i> DNA polimeraza	Bitools B&M Labs, S.A., Španija

3.9.4 Pufri in raztopine

TAE puffer 50×:

242 g Tris Base
37,1 ml ledocetne kisline
100 ml 0,5 M EDTA pH 8
dH₂O do 1000 ml; pH 8,5

Nalagalni puffer za agarozno gelsko elektroforezo (6×):

50% (wt/vol) saharoza
0,05% (wt/vol) bromfenolmodro

Etidijev bromid (10×): 5 mg/l (wt/vol) raztopina v dH₂O

3.9.5 Uporabljeni začetni oligonukleotidi

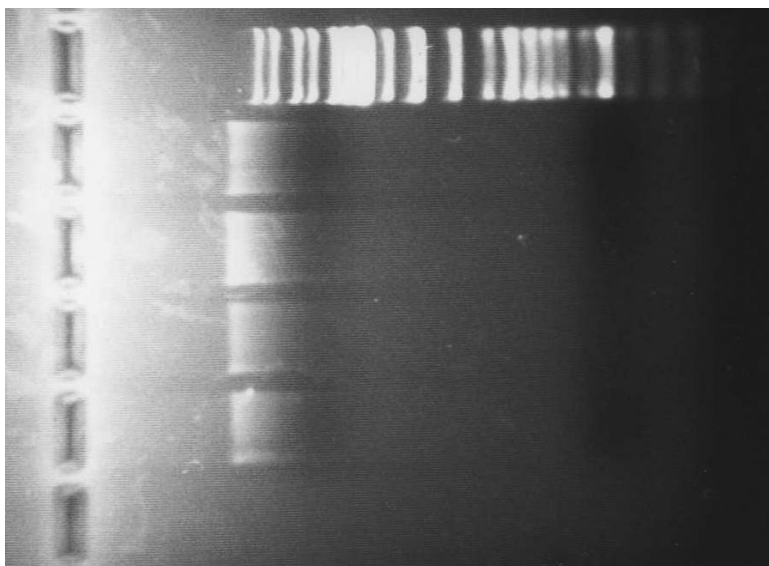
Preglednica 5: Sekvence in tarčna mesta začetnih oligonukleotidov

Začetni oligonukleotid	Sekvenca (5' proti 3')	Tarča
27f	AGAGTTTGATCCTGGCTCAG	16S rRNA bakterij
R1495	CTACGGCTACCTTGTTACGA	16S rRNA bakterij
T7	GTAATACGACTCACTATAGGGC	pGEM-T vector
SP6	ATTTAGGTGACACTATAG	pGEM-T vector

4 REZULTATI

4.1 IZOLACIJA DNA, OCENA ČISTOSTI IN KONCENTRACIJE IZOLIRANE DNA

Po izolaciji celotne bakterijske DNA iz vzorca tal smo na agaroznem gelu preverili uspešnost izolacije (Slika 1). Zanimala nas je velikost fragmentov ter prva ocena količine izolirane DNA, ki jo razberemo z jakosti sevanja DNA na gelu, ko ga osvetlimo pod UV lučjo. V vseh štirih paralelkah smo dobili zadovoljivo količino zadosti velikih fragmentov DNA (nad 20 kilobaznih parov – kbp), da smo lahko nadaljevali z raziskovanjem.



Slika 1: Fotografija elektroforetskega gela, na katerem smo preverjali velikost izolirane DNA

Čistost izolirane DNA smo ocenili z metodo merjenja optične gostote (absorbance) raztopine DNA. Delali smo v dveh paralelkah, vsako od njih smo dvakrat pomerili pri treh valovnih dolžinah (230, 260 in 280 nm). Tabela 3 podaja izmerjene absorbance.

Preglednica 6: Absorbance raztopine izolirane DNA pri različnih valovnih dolžinah

valovne dolžine [nm]:		230	260	280
A	1.	0,55	0,25	0,16
	2.	0,5	0,29	0,17
B	1.	0,53	0,26	0,19
	2.	0,57	0,31	0,18
povprečna absorbanca:		0,5375	0,2775	0,175
razmerja	260/280	1,588571		optimum 1,7-2,0
	260/230	0,516279		optimum >1,5

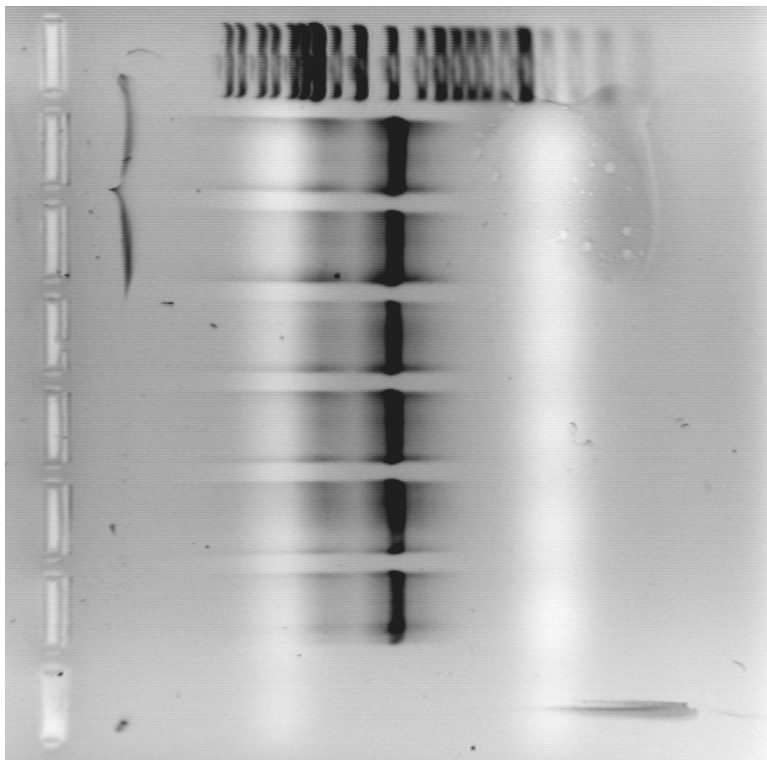
Razmerje A260/A280 je 1,59, kar nam pove, da ima naša raztopina DNA precej nečistoč, saj vrednost leži izven intervala [1,7-2,0]. To ne pomeni, da je taka DNA neuporabna za nadaljnje raziskovanje, lahko pa je to razlog težav v nadaljevanju (Promega Protocols and Applications Guide, p. 9-4). Tudi razmerje A260/A230 nakazuje prisotnost znatne količine kaotropnih soli in organskih molekul v raztopini, saj je izračunana vrednost 0,51 daleč od optimalne vrednosti, ki je nad 1,5.

Izračunana koncentracija DNA je med 40 in 55 $\mu\text{g/ml}$. Te vrednosti so nekoliko precenjene, ker nismo upoštevali popravka absorbance pri 320nm (Promega Protocols and Applications guide, chapter 9-5: DNA Purification).

4.2 VERIŽNA REAKCIJA S POLIMERAZO IN ČIŠČENJE PRODUKTA

Pri verižni reakciji s polimerazo (PCR) smo uporabili univerzalne bakterijske začetne oligonukleotide za pomnoževanje gena za ribosomalno RNA male podenote ribosoma (16S rRNA), in sicer 27f in 1495R. Pričakovali smo torej, da bo naš produkt polimerizacije dolg približno 1470 baznih parov (od 27. do 1495. baznega para gena glede na *E. coli*). Uspešnost polimerizacije smo preverili na 1% (wt/vol) agaroznem gelu (Slika 2), kjer smo

poleg velikostne lestvice in šestih reakcij preverjali tudi negativno kontrolo PCR (8. jamica). Rezlutat smo ocenili kot ustrezen.



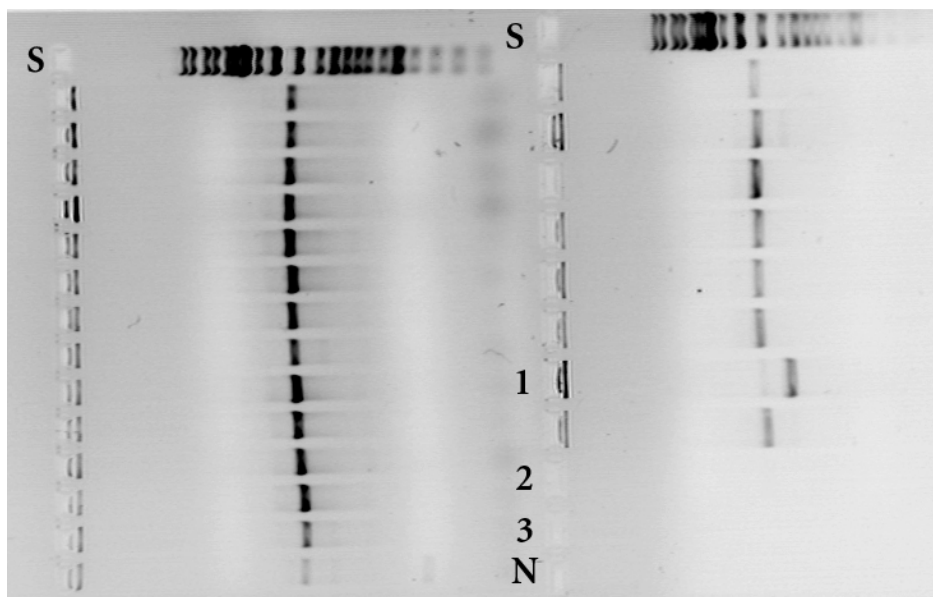
Slika 2: Preverjanje produkta PCR in negativne kontrole na agaroznem gelu

Vzorci smo nato združili ter jih čistili iz gela s kompletom po navodilih proizvajalca (Wizard SV Gel and PCR Clean-up System, cat. # A9281; Promega, Madison, Wisconsin). Uporabili smo 1% gel iz agaroze z znižanim tališčem (ang. *low-melting agarose*). Uspeh čiščenja smo preverili z elektroforezo na običajnem 1% agaroznem gelu in ga ocenili kot ustreznega.

4.3 KLONIRANJE, IZBOR KLONOV IN SEKVENC

Kloniranje smo izvedli s pGEM-T vektorskim sistemom in na gojiščih za modro-belo selekcijo izbrali transformante. Ker velikih belih kolonij ni bilo zelo veliko, smo k izbranim klonom dodali še 30 rahlo modrih kolonij. Skupno smo tako izbrali 312 kolonij.

Verižno reakcijo s polimerazo (PCR) smo uporabili tudi za pregledovanje velikosti vstavljenih fragmentov izbranih kolonij. V reakciji smo uporabili začetna oligonukleotida, katerih tarčno mesto je na plazmidu pGEM-T Easy Vector na robovih vstavljenega fragmenta, in sicer SP6 in T7. Naenkrat smo testirali 24 klonov, reakcija je opisana v razdelku 3.6.



Slika 3 – Primer rezultata verižne reakcije s polimerazo, kjer smo testirali 24 klonov. V prvi vrstici obeh stolpcev je velikostni standard (S), v zadnji vrstici desnega stolpca pa je negativna kontrola (N). Za nadaljnje delo so primerni vsi razen treh klonov (1, 2, 3).

192 klonov od 312 je imelo vstavljene fragmente, ki so po velikosti ustrezali želenemu genu (približno 1,5 kbp). Te smo poslali sekvenirat.

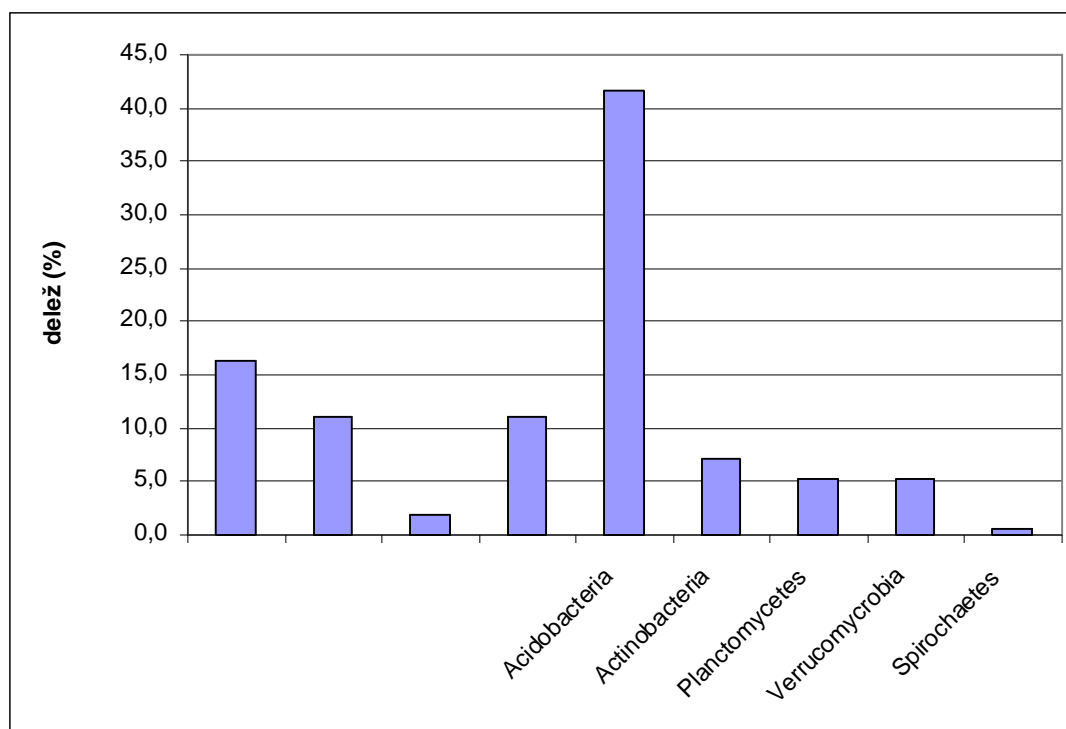
Sekvence smo ročno pregledali in odstranili nezanesljive začetne in končne dele (po približno 30 baznih parov na vsaki strani zaporedij). Odstranili smo tudi 23 sekvenc, katerih zaporedja ni bilo mogoče nedvoumno določiti.

Nato smo skušali ugotoviti, pri katerih sekvencah gre za artefakte reakcije (himerne sekvence), da bi jih izločili pred nadaljevanjem dela. Uporabili smo metode, ki so opisane v razdelku 3.7. Precej zanesljivo lahko trdimo, da je 15 sekvenc himernih ($15/169 = 0,0887$, torej himerne sekvence predstavljajo približno 9% klonov).

V klonski knjižnici nam je po odstranitvi nezanesljivih in himernih sekvenc tako preostalo 154 sekvenc.

4.4 FILOGENETSKO DREVO

Filogenetsko drevo naše knjižnice 16S rRNA bakterijskih genov (Slika 5), narejeno po metodi varčnosti (ang. *maximum parsimony*) v programu ARB, razdeli sekvence v več skupin: 62 sekvenc pripada proteobakterijam (od tega 25 v skupino *Alphaproteobacteria*, 17 med *Betaproteobacteria*, 3 med *Gammaproteobacteria*, 17 med *Deltaproteobacteria*), 64 sekvenc pripada skupini *Acidobacteria*, 11 *Actinobacteria*, po 8 *Planctomycetes* in *Verrucomicrobia* ter ena, ki se na drevesih venomer pojavlja ločeno in smo jo s pomočjo baz podatkov določili za *Spirochaetes*. Delež sekvenc po skupinah predstavlja Slika 4.

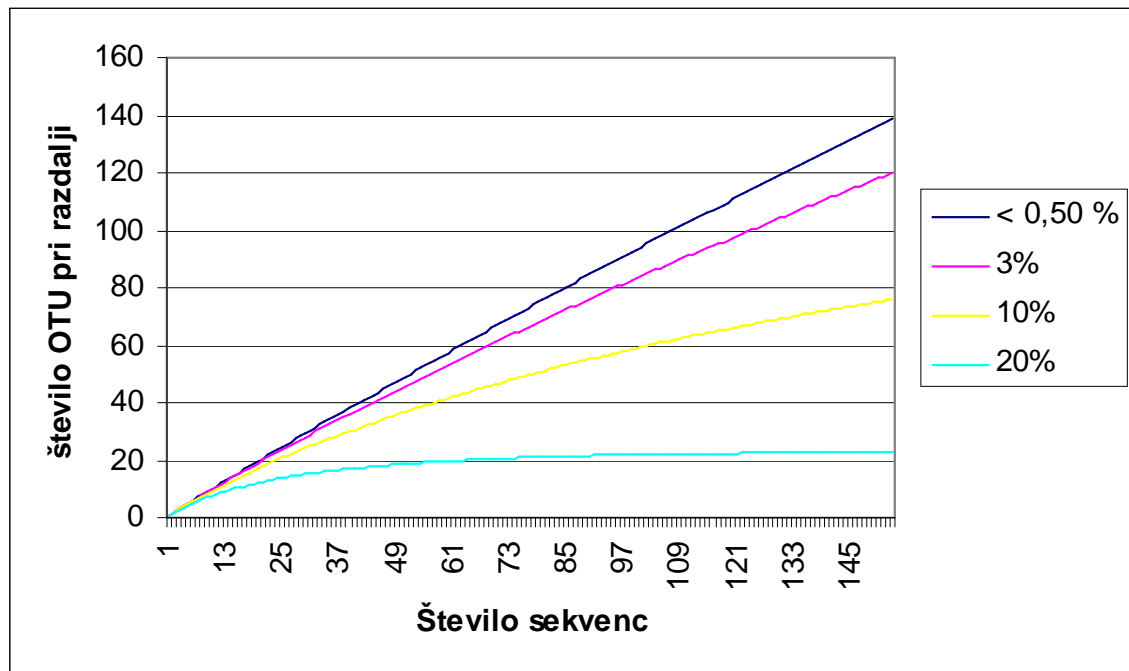


Slika 4: Deleži skupin bakterij v klonski knjižnici visokega barja.

4.5 OPERACIJSKE TAKSONOMSKE ENOTE

Pri pregledovanju knjižnice nas je zanimalo tudi, kako so si sekvence med seboj podobne, zato smo jih združili v operacijske taksonomske enote (OTU). Uporabili smo program DOTUR (Schloss in Handelsman, 2005), saj je ročno pregledovanje tako velike distančne matrike zelo zamudno in se pri tem zlahka zmotimo. Prednost programa DOTUR je tudi v tem, da hkrati izračuna število OTU in njihove predstavnike za vse možne razdalje (Schloss in Handelsman, 2005). Zanimivo je pogledati število OTU pri razdalji 3% (genetska razdalja med 0,0250 in 0,0349), saj slednja največkrat označuje mejo med vrstami (Gevers in sod., 2005). 154 sekvenc program razdeli glede na distančno matriko v 120 OTU, in sicer v dve s po šestimi sekvencami, v pet s po tremi sekvencami, v 14 OTU s po dvema sekvencama, ostale OTU (99) pa imajo po eno sekvenco naše genske knjižnice. Polovica OTU s po več kot eno sekvenco (10 od 21, med njimi tudi obe enoti s po šestimi sekvencami) vključuje acidobakterijske predstavnike. V primeru naše knjižnice velja, da število OTU upada s povečevanjem števila sekvenc v knjižnici pri razdalji 0,03 zaradi podobnosti med pridobljenimi sekvencami *Acidobacteria*. Število OTU pri različnih razdaljah prikazuje Slika 6. Izbrali smo razdalje, ki največkrat razmejujejo enake sekvence (<0,5%), vrste (< 3%), rodove (<5%) in debla (<20 - 25%) (Schloss in Handelsman, 2005; Hugenholtz in sod., 1998). Pri tem se zavedamo, da so to arbitrarne vrednosti in so le približek izsledkov klasične taksonomije in modernih metod sekveniranja DNA. Iz slike 8 vidimo, da smo pri razdalji 20% dobili kar 23 OTU, vendar že pri razdalji 24% število operacijskih taksonomskih enot pade na 12, in sicer na tiste velike skupine, ki smo jih navedli v diagramu 2.

Drugih parametrov v zvezi s pokritostjo knjižnice in resnično biotsko raznovrstnostjo nismo računali, saj je za to naša knjižnica premajhna (Schloss in Handelsman, 2005).



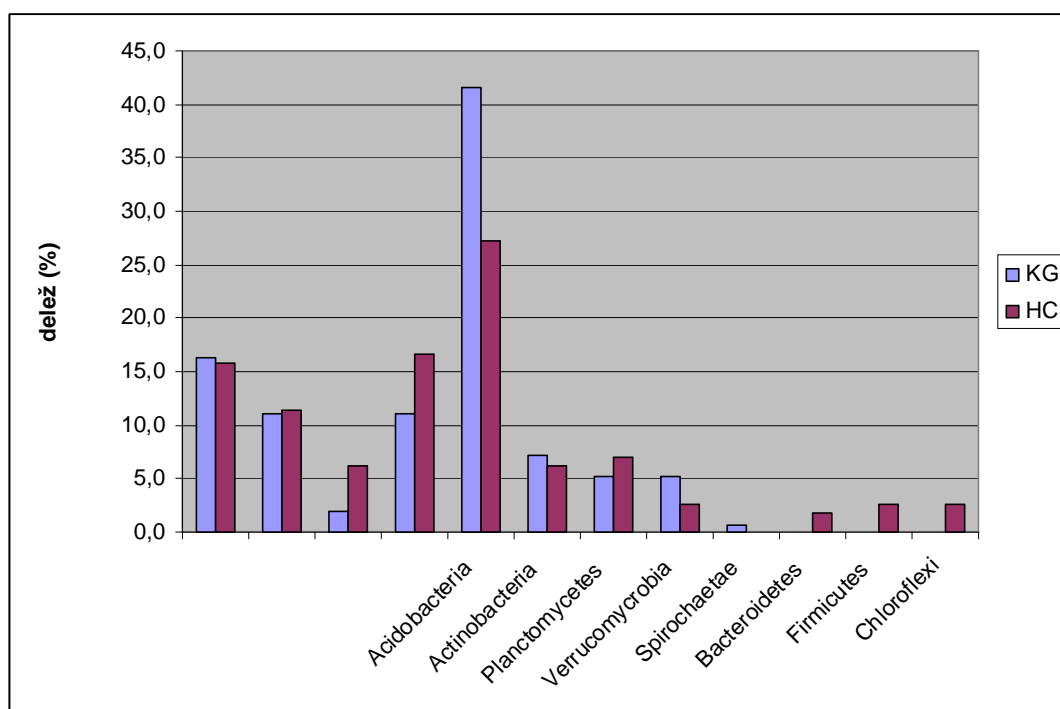
Slika 6: Spreminjanje števila operacijskih taksonomskih enot (OTU) pri različnih genetskih razdaljah sekvenc v naši knjižnici v odvisnosti od števila sekvenc.

4.6 PRIMERJAVA DVEH KNJIŽNIC GENOV ZA 16S RNA

Knjižnico genov 16S rRNA iz tal Kozlarjeve gošče (KG) smo primerjali s knjižnico, ki so jo iz vzorca tal nizkega barja z visoko vsebnostjo organskih snovi (HC) pripravili Kraigher in sod. (2006). Deleži bakterijskih skupin v obravnavanih knjižnicah so predstavljeni v Preglednici 6, grafično predstavljamo te podatke na Sliki 6.

Preglednica 6: število sekvenc v klonskih knjižnicah visokega (KG) in nizkega (HC) barja po skupinah.

	KG		HC	
	število	%	število	%
α -Proteobacteria	25	16,2	18	15,8
β -Proteobacteria	17	11,0	13	11,4
γ -Proteobacteria	3	1,9	7	6,1
δ -Proteobacteria	17	11,0	19	16,7
Acidobacteria	64	41,6	31	27,2
Actinobacteria	11	7,1	7	6,1
Planctomycetes	8	5,2	8	7,0
Verrucomycrobia	8	5,2	3	2,6
Spirochaetes	1	0,6		0,0
Bacteroidetes			2	1,8
Firmicutes			3	2,6
Chloroflexi			3	2,6
SKUPAJ:	154	100	114	100

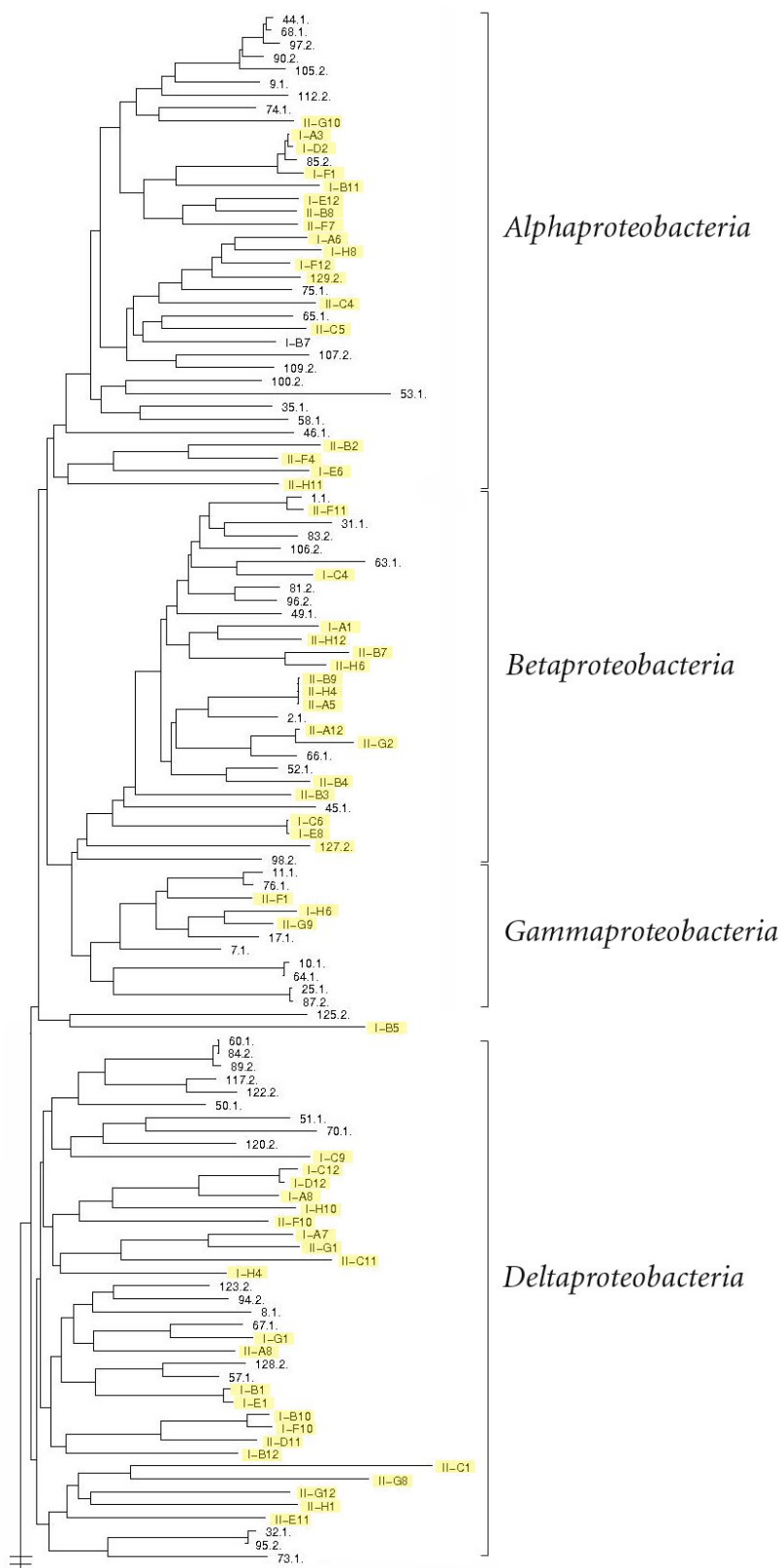


Slika 7: Deleži bakterijskih skupin v klonskih knjižnicah visokega (KG) in nizkega (HC) barja.

V obeh knjižnicah predstavljajo proteobakterije pomemben delež; v knjižnici z visokega barja (KG) jih je približno 40 %, v tisti z nizkega barja (HC) pa celo 50 %. Zdi se, da je raznolikost slednje knjižnice nekoliko večja, saj so prisotne 3 skupine bakterij, ki jih v prvi ne zasledimo. Očitna pa je tudi dominacija skupine *Acidobacteria* v knjižnici iz visokega barja, saj predstavlja skoraj 42% celotne knjižnice.

Za vtis o podobnosti zaporedij obeh knjižnic jih je nujno prikazati na skupnem filogenetskem drevesu. Zaradi velikega števila sekvenc je metoda varčnosti zelo zamudna, zato smo uporabili metodo povezovanja sosedov (ang. *neighbour joining*). Rezultat prikazuje Slika 7.

Kljub temu, da gre za v osnovi popolnoma drugačno filogenetsko metodo, so filogenetske skupine razen nekaj izjem enako velike in vsebujejo enake sekvence kot v ločenih drevesih (drevo, narejeno po metodi varčnosti, za knjižnico iz nizkega barja – HC – je prikazano v članku Kraigher in sod., 2006). Vidimo, da so pri vseh skupinah, razen seveda pri tistih, ki so prisotne le v eni knjižnici, sekvence obeh knjižnic med seboj premešane. To pa ne velja za skupino *Acidobacteria*, kjer so sekvence obeh knjižnic dokaj dobro razmejene.



Slika 8 – Skupno drevo obeh knjižnic – samo skupina *Proteobacteria*. Drevo je narejeno po metodi povezovanja sosedov (ang. *neighbour joining*). Sekvence HC so označene z arabskimi številkami, sekvence KG pa z rimskimi in črko ter z rumeno barvo.

4.7 ANALIZA SEKVENC IZ SKUPINE *Acidobacteria*

Zaradi naraščajočega zanimanja za skupino *Acidobacteria* med raziskovalci (Quaiser in sod., 2003) smo želeli ugotoviti, če zares obstajajo razlike med kloni obeh knjižnic, ki smo jih uvrstili med acidobakterije. Naredili smo filogenetsko analizo teh sekvenc v programu ARB, kjer smo sekvence poravnavali ob posebno bazo 2222 dobro poravnanih acidobakterijskih sekvenc, ki smo jo dobili v internetni bazi RDP (Cole in sod., 2007) Tabela 6 povzema filogenetsko drevo (ni prikazano), ki smo ga naredili po metodi varčnosti (ang. maximum parsimony).

Preglednica 7: Število sekvenc v posamezni skupini acidobakterij (Barns in sod., 2007) ter deleži teh skupin v knjižnicah z visokega (KG) in nizkega (HC) barja. Ene sekvence iz knjižnice HC nismo mogli nedvoumno uvrstiti v nobeno od 26 predlaganih skupin.

	KG	delež	HC	delež
GP1	26	0.41	5	0.19
GP2	22	0.34		
GP3	11	0.17		
GP4			5	0.19
GP5	1	0.02		
GP6			12	0.44
GP7	2	0.03		
GP10			1	0.04
GP11			1	0.04
GP13	2	0.03		
GP25			2	0.07
neuvrščeno			1	0.04
SKUPAJ:	64	1	27	1

Ugotovili smo, da razen 1. podskupine nimata knjižnici skupne nobene druge podskupine. Pri KG je v skupinah 1, 2, in 3 kar 92% vseh sekvenc, pri HC pa v skupinah 1, 4 in 6 skupno 82% vseh sekvenc iz knjižnice.

5 RAZPRAVA IN SKLEPI

5.1 RAZPRAVA

Oceniti želimo postopek pridobivanja klonske knjižnice in teoretično opredeliti vzroke in velikosti napak ter omejitve pri posameznih korakih. Nadrobnejši pregled literature teoretičnih raziskav metode lahko pomaga pri interpretaciji rezultatov in načrtovanju prihodnjih raziskav. Sledi komentar rezultatov, ki smo jih primerjali z nekaterimi drugimi študijami in odgovorili na vprašanja in hipotezo iz poglavja 2.3.

5.1.1 Artefakti in vzroki pristranskosti verižne reakcije s polimerazo (PCR)

5.1.1.1 Pristranskost in težave pri prepoznavanju vrst po zaporedjih, pridobljenih s PCR

Kadar želimo pomnožiti nek gen iz zelo raznolikih in nesorodnih genomov, naletimo na raznovrstne težave. Ne glede na to, da pomnožujemo en gen, katerega zaporedje je precej ohranjeno (konzervativno) v različnih evlucijskih linijah, razlike med posameznimi kopijami vendarle obstajajo. Univerzalni oligonukleotidni začetniki so sicer izbrani tako, da ustrezajo čim bolj ohranjenim regijam gena, ki je torej skupen čim več predstavnikom, ki jih želimo zaznati, a pri tem obstaja več pasti.

Naše poznavanje bakterijske diverzitete je nepopolno, saj je do približno 90-ih let 20. stol. temeljilo na študiju izolatov, glede na ocene pa znamo gojiti manj kot odstotek vrst (Amann in sod., 1995). Največja zbirka zaporedij bakterijskih rRNA genov - Ribosomal Database Project (Cole s sod. 2007) - ima ob zadnji dopolnitvi približno 440 tisoč vnosov (izdaja 9.55, oktober 2007). Resda se število hitro večja, v zadnjem letu se je skoraj podvojilo, vendar vsebuje veliko ponovljenih sekvenc, delnih in nepopolnih sekvenc (le približno 35% zaporedij je daljših od 1200 baznih parov), na kar kaže hierarhično drevo, ki ga spletni servis ponuja.

Prepoznavanje vrst po sekvencah DNA pomembnih genov (na primer gena za 16S rRNA) oteži tudi dejstvo, da imajo nekatere bakterijske vrste po več kopij operonov z geni za ribosomalno RNA (rrn operoni). Počasi rastoče bakterije imajo običajno manj kopij, *Bradyrhizobium japonicum* ima celo zgolj eno samo, hitro rastoče pa več kopij: različne vrste *Bacillus sp.* med 9 in 12, *Clostridium paradoxum* celo 15 kopij (povzeto v Wintzingerode in sod., 1997). Raziskave kažejo, da variabilnost med geni znotraj vrste presega pričakovanja; raziskava Claytona in sodelavcev (1995), ki so preiskovali zaporedja 16S rRNA iz baze GenBank, je pokazala, da variabilnost teh genov znotraj enega seva dostikrat tudi za dvajsetkrat presegajo ocenjeno napako sekveniranja in so torej najverjetneje znak diverzitete gena celo znotraj seva ali celo posamezne bakterije. Avtorji zato svarijo pred napačnimi taksonomskimi uvrstitvami organizmov s pomočjo 16S rRNA gena, saj se razlike med geni znotraj enega seva po velikosti lahko približajo razlikam med rodovi. Nekateri pa svarijo, da je zaradi tega lahko predpostavljena raznolikost bakterij v naravi precenjena (Schloss in Handelsman, 2005).

Navedeno nas vzpodbudi k razmišljanju v dveh smereh: bodisi pripravimo degeneriran začetni oligonukleotid ali pa se moramo sprijazniti, da afiniteta začetnega nukleotida ne bo enaka za vse matrične DNA. Ker vse matrične DNA, predvidene in nepredvidene, pač nimajo skupnega tarčnega mesta, lahko to povzroči pristranskost pri pomnoževanju. Uporaba degeneriranih začetnih oligonukleotidov lahko tako pomeni izboljšavo: če uporabimo mešanico začetnih oligonukleotidov, ki se med seboj razlikujejo na mestih, kjer pričakujemo razlike med različnimi tarčami, lahko pričakujemo manj pristranskosti pri pomnoževanju, kar pomeni, da bomo v pomnožku verjetneje dobili vse sekvence, ki so bile zastopane v vzorcu, in v enakih razmerjih. Polz in Cavanaugh (1998) sta pokazala, da to drži v enostavnih primerih, ko sta uporabila le dve sekvenci in za njiju dva začetna oligonukleotida, a še to le, če sta uporabila zelo mile pogoje pri vezavi začetnih oligonukleotidov na DNA. Ob zaostritvi pogojev (ang. high stringency) sta venomer dobila produkt v prid tisti sekvenci, ki je imela na degeneriranem mestu G-C par. Zato je priporočljivo preveriti pristranskost vsakič, ko uporabimo degenerirane začetne oligonukleotide, oziroma se jim, kjer je to mogoče, izogniti (Kanagawa, 2003). Wintzingerode in sodelavci (1997) pa svarijo pred drugačno pristranskostjo: verige tistih

genov, ki imajo več G-C parov, se bodo v fazi denaturacije težje ločile, zato bo med matrikami za pomnoževanje več takih z večjo vsebnostjo A-T baz in te se bodo preferenčno pomnoževale. Da bi se temu izognili, je potrebno prilagoditi protokol PCR reakcije.

Večkrat so že opazili, da se med verižno reakcijo s polimerazo ruši razmerje med sekvencami glede na tisto v izhodnem vzorcu. Razmerje različnih sekvenc se z naraščanjem števila ciklov približuje 1:1, v produktu se količina različnih PCR produktov z naraščanjem števila ciklov izenačuje. Pojav sta skušala pojasniti Suzuki in Giovannoni (1996) s svojim kinetičnim modelom: v PCR produktu, ki nastane s pomnoževanjem različnih izvornih matrik, hitreje narašča število kopij tiste molekule DNA, ki se pomnožuje z večjo učinkovitostjo (če je recimo malo krajša, če se začetni oligonukleotid bolje prilega); ko je veliko takih molekul, te v fazi vezave začetnih oligonukleotidov ponovno hibridizirajo in onemogočijo vezavo začetnega nukleotida ali nadaljevanje polimerizacije. Zato se v kasnejših ciklih PCR stopnja pomnoževanja zanje hitreje zmanjšuje. Nasprotno pa se stopnja pomnoževanja tistih molekul DNA, ki se sicer pomnožujejo manj učinkovito, ne spremeni, kar privede do zmanjševanja razmerja med sekvencami glede na začetni vzorec (Suzuki in Giovannoni, 1996).

Model so preverili (Kurata in sod., 2004) s serijo poskusov, pri katerih so tri sekvence v znanih začetnih razmerjih pomnoževali s PCR po različnih protokolih; spreminjali so trajanje različnih faz cikla in hitrost menjavanja temperature med njimi. Ugotovili so, da je bistvena hitrost spreminjanja temperature med denaturacijo in fazo naleganja začetnih oligonukleotidov. Če se to primeri hitro, bodo vsi trije možni hibridi (v termodinamskem vrstnem redu so to homodupleks, heterodupleks ter matrika in začetni nukleotid) imeli enako časa za nastanek, tedaj bo homo- in heterodupleksov relativno malo in celokupna reakcija ne bo tako pristranska (Kurata in sod., 2004). Če pa je nasprotno ta čas dolg, bodo termodinamsko ugodnejši homodupleksi imeli dlje časa za nastanek, zato bo pristranskost večja. Kurata in sodelavci (2004) tako popravljajo model Suzukija in Giovannonija (1996) v tem, da je bistveno trajanje spreminjanja temperature med fazami cikla pomnoževanja; ni težava v ponovni hibridizaciji PCR produktov v fazi vezave začetnih oligonukleotidov,

ampak nastanek homodupleksov v času (počasnega) ohlajanja na prehodu med fazama denaturacije in vezave začetnega nukleotida.

Farrelly in sodelavci (1995) so pomnoževali 16S rRNA gene iz dveh bakterijskih vrst naenkrat, za katere so poznali velikost genoma in število rrm operonov. Ponekod so iz teh podatkov pravilno napovedali razmerje genov obeh vrst v produktu PCR reakcije, ponekod pa ne. Avtorji zaključujejo, da številčno vrednotenje količine posameznih matrik (vrst) v izvornem vzorcu ni možno, če ne poznamo velikosti in števila operonov z rRNA geni; sklepanje o številu vrst v vzorcu na podlagi klonske knjižnice ni mogoče (Farrelly in sod., 1995). Tega se moramo zavedati, ko pripravljamo majhno klonsko knjižnico, kot je naša, iz tako kompleksnega vzorca, kot so tla. Pokazali so na primer (Schloss in Handelsman, 2005), da je klonske knjižnice s po približno 100 kloni izredno težko primerjati med sabo; tako na primer niso mogli statistično pokazati razlike v vrstnem bogastvu »bogate« zemlje iz amazonskega pragozda in »navadnega« škotskega pašnika, ker je bilo za to enostavno premalo podatkov.

Na splošno lahko rečemo, da bodo v produktu ohranjena razmerja genov iz vzorca le, če se bodo vsi geni pomnoževali enako učinkovito. To pa pomeni sprejeti naslednje predpostavke (Wintzingerode in sod. 1997):

- vse matrične molekule so ves čas enako dosegljive za začetne oligonukleotide;
- na vse molekule se začetni oligonukleotid veže enako učinkovito (z enako vezavno energijo);
- DNA polimeraza pomnožuje na vseh matrikah enako učinkovito;
- ko zmanjka reagentov, se to enako pozna pri pomnoževanju vseh matričnih DNA.

Glede na vse prej povedano je jasno, da je te predpostavke težko popolnoma sprejeti, zato je enako težko trditi, da razmerja med velikostmi večjih skupin bakterij v naši knjižnici ustrezajo realnim razmerjem v tleh na visokem barju, od koder smo vzorec odvzeli.

5.1.1.2 Artefakti, ki nastanejo pri verižni reakciji s polimerazo

5.1.1.2.1 Heterodupleksi

V vsakem ciklu verižne reakcije s polimerazo DNA denaturiramo, tako da je vsa DNA enoverižna. Nato se temperatura zniža, da pride do vezave homolognih delov DNA. Načeloma pride lahko do treh vezav: v začetnih ciklih je reagentov veliko in produkta malo, zato se večinoma povežejo začetni oligonukleotidi in njihove tarčne sekvence (matrična DNA). Ko pa je pomnoženega produkta veliko, se lahko povežejo komplementarne verige produktov (homodupleks) ali celo nehomologni deli DNA (heterodupleks). V slednjem primeru dobimo artefakt, na katerega je potrebno misliti ob nadaljnjem raziskovanju.

Če pregledujemo produkte pomnoževanja s PCR z elektroforezo, bomo heterodupleks zaznali kot novo linijo (ang. *band*) na elektroforetskem gelu: hitrost potovanja heterodupleksa v gelu je zaradi nepopolne homologije upočasnjena. Zmanjšanje mobilnosti heterodupleksa je obratno sorazmerna homolognosti sekvenc, ki ga tvorita (Kanagawa, 2003). Tako nam heterodupleksi dajo lažno sliko o diverziteti produkta.

Na številne težave naletimo, ko produkt naprej analiziramo z metodami, ki temeljijo na PCR, na primer pri T-RFLP, kloniranju in sekveniranju (Kanagawa, 2003). Zavedati se moramo popravljalnih mehanizmov, ki jih ima gostiteljski organizem, v katerega kloniramo produkte (največkrat je to *E. coli*), ker lahko spremeni heterodupleks v homodupleks. In ker ni nobenega načina, kako bi popravljalni sistem določil izvorno zaporedje, naključje odloči, katera veriga se bo popravila. Taka odločitev lahko poteče na vsakem nehomolognem mestu posebej, tako da se enkrat popravi ena, drugič pa druga veriga, kar pripelje do nastanka umetnih zaporedij, ki jih ni v začetnem vzorcu (Kanagawa, 2003). Speksnijder in sodelavci (2001) so v eksperimentu, v katerem so pomnoževali 7 ozko sorodnih genov za 16S rRNA, pokazali, da so heterodupleksi lahko problem in da je v klonski knjižnici lahko nekaj odstotkov umetnih sekvenc (pri njih 2 od 66, kar je 3%).

Obstaja nekaj metod, s katerimi se znebimo heterodupleksov (Kanagawa, 2003): lahko ponovimo PCR z zelo razredčenim produktom za zelo malo ciklov ali pa heteroduplekse detektiramo z elektroforezo (PAGE) in jih ročno odstranimo, vendar je to zelo težavno, ko ne poznamo sekvenc ali so si te ozko sorodne. Glede na to, da se heterodupleksi tvorijo v poznih ciklih PCR, je smiselno skrajšati reakcijo in se jim tako izogniti (Kanagawa, 2003).

V našem primeru je nemogoče oceniti, če imamo v knjižnici prisotne artefakte te vrste in koliko jih je. Smiselno bi bilo skrajšati protokol PCR reakcije za nekaj ciklov, vendar smo uporabili uveljavljen protokol, ki je bil že prej uporabljen pri podobnih raziskavah (Kraigher in sod., 2006).

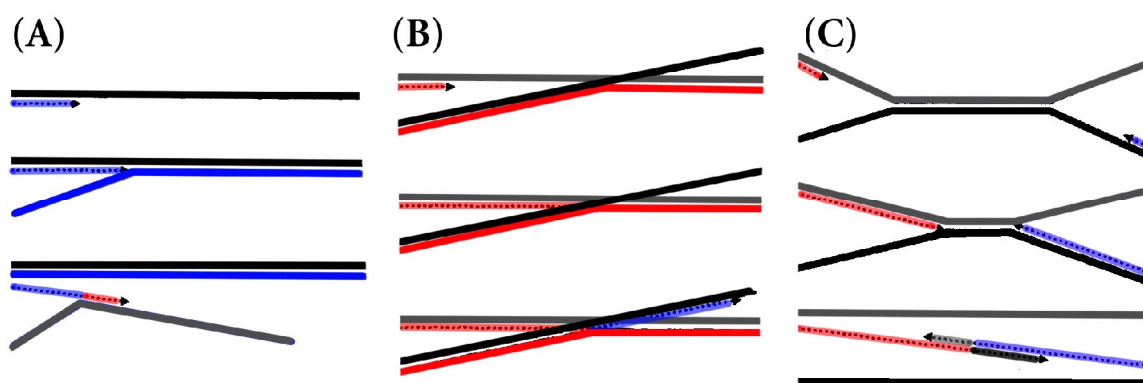
5.1.1.2.2 Himerne sekvence

Himere so artefakti PCR, saj nastanejo iz dveh (različnih) izvornih verig DNA. Rezultat je torej povsem umetna tvorba, ki je ni v izvornem vzorcu, in daje napačen vtis o diverziteti vzorca. Do sedaj so predlagali dva molekularna mehanizma nastanka himer (Kanagawa, 2003).

Vsak košček DNA, ki ima prost 3'-konec in je na tem koncu vsaj delno homologen DNA iz vzorca, lahko deluje kot začetni oligonukleotid. Če je tak košček dolg nekaj sto baznih parov in je pravzaprav produkt pomnoževanja z neke druge verige DNA (ali iz drugega mesta iste), potem to vodi v nastanek himerne sekvence (Slika 10.A). Himere nastanejo po tem mehanizmu večinoma takrat, ko je dovolj takih kosov DNA, ki zato uspešno tekmujejo z dodanimi začetnimi oligonukleotidi, kar pa se verjetneje dogaja bolj proti koncu PCR reakcije. Frekvenco pojavljanja takih himer lahko tako znatno zmanjšamo z omejitvijo števila ciklov verižne reakcije s polimerazo (Kanagawa, 2003).

Drug način nastanka himernih sekvenc je zamenjava matrične DNA med samo polimerizacijo (ang. *template switching*), torej lahko himere nastanejo v enem samem ciklu PCR (Kanagawa, 2003). Ko v vsakem ciklu hibridizirajo homologni deli enoverižne DNA, se lahko zgodi, da se poleg začetega oligonukleotida nekje v smeri polimerizacije (ang.

downstream) veže še kakšna druga veriga DNA, ki ji je na tistem delu homologna (Slika 10.B). Polimeraza tako začne sintezo komplementarne verige na mestu vezave začetnega oligonukleotida, a nato ob razvejitvi preskoči na drugo verigo in nadaljuje sintezo ob njej. Da se to zgodi, se mora nastajajoča veriga na 3'-koncu odlepiti od matrike, kar se lažje zgodi pri višjih temperaturah polimerizacije; to so potrdili tudi Patel in sodelavci (1996), ki so preučevali nastanek himer v enem ciklu PCR in dobili več himer pri višji temperaturi. Možno je tudi, da verigi, ki nastajata ob obeh vrvicah heterodupleksa, nadaljujeta sintezo druga ob drugi (slika 10.C), lahko pa homologna veriga prepreči nadaljevanje sinteze (Slika 10.A) in zato polimerizacijski aparat razpade (Kanagawa, 2003); slednji primer je možen začetek nastanka himere v naslednjih ciklih po mehanizmu, ki smo ga opisali v prejšnjem odstavku. Čeprav tovrstne himere lahko nastanejo v enem samem ciklu PCR, pa načeloma potrebujemo dovolj homolognih sekvenc, ki bi se vezale na matrike in zmotile polimerazo, to pa se kot rečeno, bolj pogosto dogaja v zadnjih ciklih PCR. Acinas in sodelavci (2005) so pokazali velik pomen števila ciklov PCR na tvorbo himer: z zmanjšanjem števila ciklov iz 35 na 18 so zmanjšali delež himernih sekvenc s 13 na 3% ($N \sim 1000$), pri čemer se struktura obeh knjižnic (deleži sekvenc, ki so jih uvrstili v večje bakterijske skupine) ni značilno razlikovala.



Slika 10 – Shematičen prikaz nastanka himernih molekul DNA. (A) Vsaka veriga s prostim 3' koncem, ki je vsaj delno homologna kateri koli matriki iz vzorca, lahko deluje kot začetnik za polimerizacijo. (B), (C) Nastanek himer z zamenjavo matrične DNA med polimerizacijo.

Wintzingerode in sodelavci (1997) poleg omenjenih dveh ključnih razlogov za nastanek himer (homologne sekvence in veliko število ciklov PCR reakcije) poudarjajo še dva dejavnika: čas elongacije in poškodbe DNA. Obstaja obratno sorazmerje med številom nastalih himer in elongacijskim časom (daljša kot je elongacija, manj je himer). Pääbo in

sodelavci (1990) so preverjali povezavo med poškodbami DNA in nastankom himer; DNA so poškodovali na različne načine (z ultrazvokom, UV sevanjem, prelamljanjem z restrikcijskimi encimi, depurinacijo) in vselej pokazali povečano produkcijo himernih molekul. Pokazali so, da lahko s pomočjo nastalih himernih sekvenc do neke mere ugotovimo, kako poškodovana je bila matrična DNA. Menijo, da je glede na ostre pogoje, s katerimi izoliramo DNA iz okoljskih vzorcev, zelo verjetno, da DNA poškodujemo in tako pripomoremo k nastanku artefaktov (Wintzingerode in sod., 1997).

Glede na to, da se himere redno pojavljajo v različnih raziskavah ter da imamo nekaj orodij z njihovo odkrivanjem (opisane in citirane so v razdelku 2.8), smo jih uporabili tudi v tej raziskavi in odstranili 15 sekvenc, kar predstavlja približno 9% naše klonske knjižnice, kar je podobno primerljivim raziskavam (Acinas in sod., 2005). Kot smo že zapisali, bi tudi tu nekoliko skrajšana reakcija PCR utegnila zmanjšati število himer. Delež himer v naši knjižnici je bil primerljiv z rezultati drugih raziskav (13% - Morales in sod., 2006; 10,5% - Graff in sod., 2005; 17% - Dedysh in sod., 2006).

Za odkrivanje himernih sekvenc, ki smo jih kot artefakt morda pomnožili iz kompleksnega okoljskega vzorca, vendarle nimamo nedvoumnega postopka. Odkrivanje himer je enostavno le v primeru, ko natančno vemo, kaj je v vzorcu, ki ga pomnožujemo. Če so v vzorcu le znane molekule, lahko himero enostavno dokažemo s primerjanjem dobljene sekvence s tistimi v vzorcu (Wang in Wang, 1997). Vrsta programske opreme, ki lahko naredi poravnave sekvenc (ang. *alignment*), je sposobna nedvoumne določitve sekvence za himero. Zgodba pa se zelo zaplete, ko je vzorec celokupna izolirana DNA iz kompleksnega naravnega vzorca; v tem primeru ne moremo zagotovo trditi, da je sekvenca himerna in še vedno drži, da takega orodja še nimamo in ga najbrž prav kmalu tudi ne bomo imeli (Robinson-Cox in sod., 1995). V nadaljevanju zato sledi kratka diskusija uporabljenih metod.

Baza Ribosomal Database Project (Maidak in sod., 2000) je že sredi devetdesetih ponudila program CHIMERA_CHECK, ki primerja dano sekvenco z ostalimi v bazi in na podlagi podobnosti med posameznimi odseki skuša določiti, če je sekvenca himera. Program sicer deluje na zastareli bazi, ki se ne obnavlja več (<http://rdp8.cme.msu.edu/html/>). Podobnost

program računa po modelu najbližjega soseda (ang. *nearest neighbour*), kar je učinkovito, če sta izvorni sekvenci, iz katerih je nastala himera, dovolj nesorodni (različni). Pokazali so (Robinson-Cox in sodelavci, 1995), da je program nezanesljiv pri odkrivanju himer, ki so nastale iz podobnih sekvenc (podobnost >84%), čeprav je znano, da himere lažje nastajajo prav iz sorodnih matrik. Težava je tudi v tem, da je v bazi zelo veliko himernih sekvenc, ki jih ta program enostavno ne vidi (Hugenholtz in Huber, 2003). Vseeno je program ugoden za prvo testiranje sekvenc, saj je hiter, za referenco uporablja celo bazo podatkov, ni potrebno narediti poravnave sekvenc (ang. alignment) in ponudi prvo iztočnico o morebitni točki preloma. Istočasno nam program sporoči, katerim sekvencam iz baze je dana sekvenca najbolj podobna, kar je pomembna informacija za nadaljnje preizkušanje.

Naslednji korak lahko vključuje primerjanje testne sekvence z morebitnimi "starševskimi" sekvencami, iz katerih bi lahko nastala himerna sekvenca. Eden od učinkovitih programov je Pintail, ki so ga razvili na univerzi na Cardiffu (Ashelford in sod., 2005). Program primerja evolucijske razdalje podanih sekvenc in izračuna predvideno deviacijo od te sekvence, ki bi jo še lahko razložili z evolucijo. Če testirana sekvenca odstopa za več kot predvideno mero, je to močan znak, da je s sekvenco nekaj narobe. S tem programom so testirali 1400 sekvenc iz baze RDP in ugotovili, da vsaj 5% vsebuje znatne nepravilnosti in od teh naj bi bila več kot polovica sekvenc himernih (Ashelford in sod., 2005). Avtorji tega preizkusa zato trdijo, da je potrebno natančno pregledati sekvence za tovrstne napake, preden jih vključimo v bazo podatkov. To sta pokazala tudi Hugenholtz in Huber (2003), ki spodbujata rutinsko odkrivanje himer pred objavo v bazah podatkov; poleg programa CHIMERA_CHECK naj bi naredili še metodo s filogenetskimi drevesi z delnimi sekvencami (ang. *partial treeing analysis*): sekvence prerežemo na polovico in iz vsakega seta podatkov naredimo filogenetsko drevo. Primerjava obeh dreves oz. različno razporejanje delov iste sekvence v obeh drevesih nam služi kot dodaten kriterij, da neko sekvenco označimo za himero.

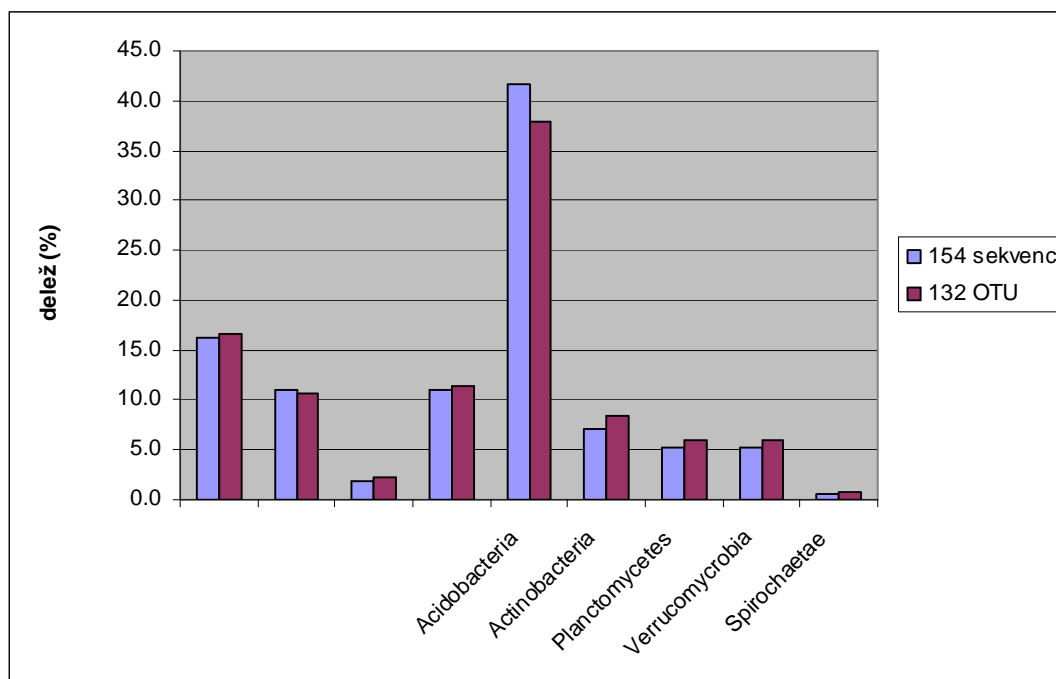
5.1.1.2.3 Naključni dogodki

Naključni dogodki predstavljajo napačno naleganje začetnih oligonukleotidov in primere, ko polimeraza vgradi napačno bazo. Polimeraze različnih proizvajalcev imajo zelo različno napako. Rekombinantna Taq DNA (<http://www.biotoools.net/eng/productos/pdf/a1.pdf>) polimeraza podjetja BIOTOOLS, na primer, vgradi napačen nukleotid na vsakih deset tisoč baz (10^{-5}), čeprav so pokazali, da je to odvisno tudi od reakcijskih pogojev (povzeto v Wintzingerode in sod. 1997). Napake so bolj usodne, če se primerijo na začetku verižne reakcije oziroma če je izhodne DNA malo. V vsakem primeru pa gre za naključne dogodke, ki jih ni možno reproducirati. Torej napako zmanjšamo, če naredimo več reakcij in zmešamo produkte oziroma uporabimo polimerazo, ki popravlja svoje napake (ang. proofreading function); uporaba Pfu polimeraze, na primer, zmanjša število napak za desetkrat (Wintzingerode in sod., 1997)

Acinas in sodelavci (2005) so zasnovali zelo obsežen eksperiment, s katerim so želeli preveriti nastajanje napak v PCR, zaradi česar naj bi napačno ocenjevali diverziteto. Iz istega naravnega vzorca bakterioplanktona so naredili dve veliki knjižnici 16S rRNA genov (vsaka s približno tisoč sekvencami); eno so naredili po bolj utečenem protokolu za PCR s 35 cikli, drugo pa s 15 cikli in nato še s 3, kar je močno zmanjšalo prisotnost ostalih napak (himere in heterodupleksi). S statistično primerjavo so ugotovili, da zaradi napake polimeraze, ki so jo v njihovem eksperimentu ocenili na $3,3 \times 10^{-5}$, lahko znatno precenimo diverziteto v originalnem vzorcu. Če so primerjali med seboj le sekvence, katerih razlike so bile večje od enega odstotka, sta si bili knjižnici presenetljivo enaki (80%). To je relevantno tudi za naše delo, saj bi lahko v prihodnje skrajšali PCR in se tako izognili nekaterim napakam. Avtorji poleg tega trdijo, da napaka polimeraze ni več pomembna, če sekvence združimo v skupine z 99% podobnostjo, in zato naj bi bil to najprimernejši način podajanja informacij v bodočih tovrstnih raziskavah (Acinas in sod., 2005).

Tudi naše podatke smo analizirali tako, da smo sekvence z 99% podobnostjo združili v skupne (OTU) in jih obravnavali kot enake. Ta pristop nas je privedel do ugotovitve, da se v knjižnici zmanjša število acidobakterij na 50 OTU, število proteoakterij pa poveča na 54 OTU. To je posledica medsebojne podobnosti acidobakterijskih zaporedij, kar smo omenili

že v rezultatih. Primerjavo med surovimi podatki (154 sekvenc) in v 132 OTU združenimi zaporedji prikazuje slika 10.



Slika 11: Deleži, ki jih v knjižnici zavzemajo večje skupine bakterij, če upoštevamo surove sekvence (modri stolpci) oziroma če sekvence, ki so si med seboj manj kot 1% različne, združimo v 132 operacijskih taksonomskih enot (rdeči stolpci).

5.1.2 Druge napake in omejitve

Iz nekaterih bakterij (predvsem po Gramu pozitivnih) je s celično lizo zelo težko dobiti DNA. Pokazali so (Felske in sod., 1998), da so v tleh zelo aktivni in številčno dobro zastopani predstavniki debla *Firmicutes*, na primer rod *Bacillus*, a jih je težko dobiti v klonskih knjižnicah gena 16S rDNA. Za razbitje njihovih viabilnih celic in odpornih stadijev (spor) je potrebno veliko energije, kar pa poškoduje DNA, ki smo jo izolirali iz drugih celic (Wintzingerode in sod., 1997; Liles in sod., 2005). Proizvajalec komercialnega kompleta, s katerim smo mi izolirali DNA iz vzorca tal (MO BIO Laboratories, Solana Beach, California), sicer trdi, da njihov komplet učinkovito lizira tudi spore, saj za razbitje celic uporabljajo kombinacijo toplote, detergenta in mehanske sile (www.mobio.com/products/productdetail.php?pid=12). Kljub temu predstavnikov

Firmicutes nismo našli, čeprav predstavljajo pomembno skupino bakterij v raziskavah podobnih tal (Graff in sod., 2005, Morales in sod., 2006). Tudi v tleh nizkega barja na Ljubljanskem barju je bil delež skupine *Firmicutes* manj kot 3% (Kraigher in sod., 2006). Pri kloniranju pomnoženih fragmentov DNA v gostiteljske organizme (še vedno je to najpogosteje bakterija *E. coli*), pride do napak in sicer v več korakih. Težko je sprejeti predpostavko, da imajo vsi fragmenti enako verjetnost, da se vključijo v vektor in prenesejo v organizem, še težje pa je verjeti, da se pri tem natančno ohranjajo razmerja iz vzorca. Poleg tega smo že omenili možnost, da lahko popravljalni mehanizmi v gostiteljskem organizmu spremenijo vstavljen fragment DNA v artefakt (Kanagawa 2003).

Če se omejimo le na modro-belo selekcijo, ki smo jo uporabili v našem primeru, lahko razmišljamo o korakih, ki privedejo do pristranskosti pri izboru klonov. Glede na obseg raziskave bomo morda izbrali le majhen delež, tipično nekaj sto primernih klonov, ki jih bomo tudi sekvenirali. Ogromna množica primernih klonov je za nas tako izgubljena. Modro-bela selekcija temelji na predpostavki, da bo uspešno vstavljeni fragment v plazmidnem vektorju prekinil kodirajoče zaporedje β -galaktozidaze in posledično produkcijo tega encima, zato bo taka kolonija na selekcijskem gojišču bela. Transformante pa so lahko tudi modre, če je fragment vstavljen v bralni okvir tega gena (dolga nekajkrat 3 bazne pare). Proizvajalci komercialnega kompleta za kloniranje ugotavljajo, da je tak fragment lahko dolg tudi 2000 bp (Promega 2003 – pGEM-T in pGEM-T Easy Vector Systems, Technical Manual No. 042).

Ena možnost, da se izognemo izkrivljanju rezultatov, je, da korak kloniranja v gostiteljski organizem izpustimo in pomnoženo DNA takoj sekveniramo. V zadnjih letih je na voljo nova metoda sekveniranja (Margulies in sod., 2005), imenovana 454. Omogoča izjemno hitro (do 100 Gbp v 7 urah in pol), cenovno dostopno in natančno sekveniranje (napaka pri enkratnem sekveniranju je manjša od 99,5%), s to tehnologijo so postali dosegljivi številni cilji, na primer sekveniranje genomov, uporaba v metagenomiki ter pri razkrivanju mikrobne raznolikosti (<http://www.454.com/enabling-technology/the-system.asp>). Glavna pomanjkljivost metode so kratke sekvence (tipično okoli 300 bp), ki jih tako dobimo, in to predstavlja nov izziv za računalniško obdelavo.

5.1.3 Diskusija rezultatov

Iz vzorca tal iz Kozlarjeve gošče na Ljubljanskem barju smo izolirali DNA, z verižno reakcijo (PCR) pomnožene dele genov za 16S rRNA smo klonirali v *E. coli* ter primerne klone sekvenirali. Po pregledu sekvenc in odstranitvi himer smo tako dobili 154 delnih sekvenc genov za ribosomalno RNA. Pridobljene sekvence so bile dolge med 700 in 800 baznimi pari in so pokrivalo drugo polovico gena.

Sekvence smo nato pripisali večjim bakterijskim skupinam (deblom, v primeru *Proteobacteria* in *Acidobacteria* tudi poddeblom) na podlagi filogenetskih dreves ter baz podatkov, kjer smo iskali njim podobna zaporedja. Filogenetske analize sekvenc smo naredili s programskim paketom ARB (Ludwig in sod., 2004), kjer smo naše sekvence poravnali ob veliko bazo (>50000) dobro urejenih sekvenc genov za 16S ribosomalno RNA. Ugotovili smo, da se sekvence z različnimi filogenetskimi metodami konsistentno razporejajo v spodaj našteje skupine. V tem delu smo prikazali filogenetsko drevo varčnosti za set sekvenc iz Kozlarjeve gošče in skupno drevo dveh knjižnic (KG in HC), ki je narejeno z distančno metodo. Glede na to, da se ti rezultati ujemajo uvrstitvijo s programom BLAST v bazah podatkov, lahko pripadnosti sekvenc v večje taksonomske skupine v veliki meri zaupamo.

V naši knjižnici prevladujeta debli *Proteobacteria* (40,3%) in *Acidobacteria*. (41,6%), deleži ostalih skupin so relativno majhni: *Actinobacteria* (7,1%), *Planctomycetes* (5,2%) in *Verrucomicrobia* (5,2%) ter *Spirochaetes* (1 sekvenca oz. 0,6%).

Velik delež predstavnikov debela *Proteobacteria* ni presenetljiv, saj primerljive študije običajno najdejo v tleh veliko teh sekvenc (Morales in sod., 2006; Graff in sod., 2005; Dedysh in sod., 2006). Proteobakterije so morfološko in metabolno zelo raznolika skupina. Z mikroskopskim štetjem s pomočjo fluorescentne *in situ* hibridizacije (FISH) so pokazali, da so *Alphaproteobacteria* najštevilčnejša skupina bakterij v kisljih šotnih tleh v Rusiji (Dedysh in sod., 2006), njihovo število na gram tal so ocenili na $1,3 \times 10^7$.

Nekoliko neznačilen pa je visok delež *Acidobacteria*. Gre za relativno novo skupino bakterij (Barns in sod., 1999), ki so jo odkrili po sekvencah rRNA. Predstavnike te skupine bakterij je zelo težko gojiti s tradicionalnimi metodami, zato so do sedaj predlagali že številne alternativne metode: gojenje na zelo razredčenih gojiščih (Jansen in sod., 2002), izolacija pri znižanem pH (Sait in sod., 2006), spodbujanje nastanka biofilmov (Dedysh in sod., 2006). Te bakterije rastejo zelo počasi, posamezne tehnike pa so primerne le za ožje skupine (Sait in sod., 2006), zato imamo še vedno zelo malo gojenih predstavnikov tega debla. Študije 16S rRNA so koristne pri prepoznavanju razširjenosti in filogenetske raznolikosti acidobakterij, ne moremo pa narediti nobenih zaključkov o njihovih fenotipskih ali fizioloških lastnostih. Quaiser in sod. (2003) so se zato prvi lotili odkrivanja vsebine acidobakterijskih genomov. Pokazali so podobnost z nekaterimi geni proteobakterij ter potrdili, da tudi drugi geni kažejo na koherentnost *Acidobacteria* kot samostojnega debla, trenutno pa poteka tudi sekveniranje genoma *Acidobacterium capsulatum* ATCC 51196 (TIGR, vir: <http://www.ncbi.nlm.nih.gov>). Acidobakterije so našli v sladkovodnih habitatih, v mikrobnih odejeh vročih vrelcev, v različnih tipih tal, v bioreaktorjih čistilnih naprav (povzeto v Quaiser in sod., 2003) ter drugih ekstremnih habitatih, na primer s težkimi kovinami, kislinami ali radioaktivnimi elementi kontaminiranimi habitatih (Barns in sod., 2007). Vedno novi molekularni podatki kažejo na veliko filogenetsko raznolikost teh bakterij, ki jih danes delimo v 26 podskupin, od katerih so najbolj razširjene skuine 1, 3, 4 in 6 (Barns in sod., 2007). Vse kaže, da so acidobakterije metabolno tako raznolike in ekološko tako pomembne kot na primer proteobakterije (Quaiser in sod., 2003).

Acidobacteria predstavljajo lahko 30 do 50% sekvenc v različnih klonskih knjižnicah (povzeto v Quaiser in sod., 2003). V knjižnici, narejeni iz peščenih tal v Arizoni, so na primer kar 49% od 766 sekvenc uvrstili med acidobakterije (Dunbar in sod., 2002). V naši knjižnici smo mednje uvrstili 64 od 154 sekvenc, vendar so si te med sabo dokaj podobne, kar smo opisali pri rezultatih (razdelek 4.5). Morda je to pokazatelj, da vrstna raznolikost *Acidobacteria* v preiskovanih tleh ni tako velika, pač pa je visoka njihova številčnost. Zaradi vseh obravnavanih pristranskosti tega ne moremo trditi, vendar so tla Ljubljanskega barja glede na biotehnološki potencial te nove skupine zanimiva za nadaljnje raziskovanje.

Z majhno klonsko knjižnico seveda ne moremo zajeti vse mikrobne raznolikosti v tleh. Ocenjujejo, da je v gramu tal tipično 10^9 bakterijskih celic, ki skupaj predstavljajo 2000 – 18000 različnih genomov (Dunbar in sod., 2002). Teoretični modeli napovedujejo (Dunbar in sod., 2002), da bi potrebovali med 16 in 50 tisoč sekvenc iz enega vzorca, da bi lahko natančno ocenili vrstno bogastvo. Ti modeli kažejo, da lahko z majhnimi knjižnicami reproduciramo vrstno bogastvo (pokažemo, da v vzorcu obstaja ogromno število različnih genomov, čeprav ne moremo dobro oceniti, kolikšno), ne moremo pa reproducirati zastopanosti vrst (nemogoče je v dveh knjižnicah s po nekaj sto kloni pričakovati enakih sekvenc, četudi jih naredimo iz istega vzorca). To so opazili tudi drugi raziskovalci (Morales in sod., 2006), ko so opazovali, koliko različnih OTU predstavljajo njihove sekvence (naredili so enako kot mi v diagramu 3). Kot oni tudi mi nismo opazili, da bi se pri genetski razdalji na primer 3% število OTU približevalo neki asimptotični vrednosti, kar kaže na nezadostno pokritost knjižnice (Morales in sod., 2006). Schloss in Handelsman (2004) sta na tak način poskušala sistematično oceniti, kako dobro so raziskovalci do sedaj uspeli pokriti posamezne skupine bakterij (debla), le da se nista omejevala na majhno knjižnico, temveč sta uporabila prek 56 tisoč sekvenc iz baze RDP. Z enako analizo, ko smo jo uporabili na sliki 6 (število OTU v odvisnosti od števila sekvenc), sta ugotovila, da smo še daleč od dobre slike raznolikosti mikrobov ter da je nemogoče tudi napovedati, koliko sekvenc bi še potrebovali. Zanimivi zaključki so na primer, da najboljše poznamo Gammaproteobacteria, Acidobacteria precej slabše, vendar kaže, da je raznolikost slednjih vsaj tolikšna kot raznolikost Cianobacteria. Avtorja pa opozarjata na pristranskost, ki je posledica izbire začetnih oligonukleotidov za PCR, zaradi katere smo morda prezrli pomembne skupine bakterij tudi znotraj dobro poznanih taksonomskih skupin, zato bo potrebno nadaljnje vzorčenje, da bomo lahko povečali zaupanje v te napovedi (Schloss in Handelsman, 2004).

5.2 SKLEPI

- S komercialnim kompletom izolirana DNA iz šotnih barjanskih tal je kljub čiščenju na gelu vsebovala precej primesi in je bila v nizki koncentraciji, kljub temu pri pripravi knjižnice nismo imeli težav.

- Uporabljeni (in sicer uveljavljeni) protokol PCR reakcije s 35 cikli ustvarja v poznih ciklih artefakte (na primer himere in močno popačena razmerja), ki se jih lahko vsaj teoretično z zmanjšanjem števila ciklov izognemo.
- Glede na tehnološko in cenovno dostopnost sekveniranja bi kazalo v prihodnje sekvenirati tudi prvo polovico gena za 16S rRNA (skoraj cel gen), saj več informacij pomeni zanesljivejšo filogenetsko analizo in boljšo primerljivost s sekvencami v bazah podatkov.
- Glede na napake, ki nastajajo pri kloniranju in sekveniranju, je smiselno obravnavati klonsko knjižnico kot skupek OTU, ki so si med seboj bolj kot 99% podobne (namesto surovih sekvenc). Poleg tega je pred objavo sekvenc v bazah potrebno odstraniti morebitne himere.
- Izolirali smo 154 klonov, ki smo jih pri razdalji 1% združili v 132 OTU. Najbolje sta zastopani skupini *Proteobacteria* (40,9% od 132 OTU oziroma 40,2 % od 154 sekvenc) in *Acidobacteria* (37,9% od 132 OTU oziroma 41,6% od 154 sekvenc).
- Številne sekvence acidobakterij večinoma (90%) sodijo v skupine 1, 2 in 3, ki so tudi v bazah podatkov najbolj zastopane. Nekaj sekvenc smo uvrstili tudi v skupine 5, 7 in 13.
- Od knjižnice iz tal nizkega barja (HC) se ta loči predvsem po visokem deležu acidobakterij (ki se razlikujejo od tistih v knjižnici HC) ter izostanku nekaterih skupin, predvsem *Firmicutes* (*Bacillus*).
- Knjižnica pa je premajhna, da bi lahko trdili, da razmerja sekvenc v knjižnici odražajo razmerja matričnih DNA v vzorcu tal visokega barja.

6 POVZETEK

Barja so zaradi svoje hidrološko regulatorne vloge kot ekosistem močno vpeta v lokalno okolje, zaradi velikega kopičenja nerazgrajenih organskih snovi pa njihov vpliv sega preko regionalnih meja. Za barja je značilna vsaj občasna poplavljenost z vodo, predvsem glede na izvor vode pa ločimo visoka in nizka barja. Slednja so povezana s talno vodo in precejšno količino vode dobijo tudi od površinskih pritokov, visoka barja pa so ombrotrofna in se napajajo pretežno z dežjem. Za visoka barja so značilne različne vrste šotnih mahov (*Sphagnum sp.*) in pa kisel pH tal. V pričujoči nalogi smo raziskovali bakterijsko združbo v tleh v Kozlerjevi gošči, ki predstavlja enega zadnjih ostankov visokega barja na sedaj že močno degradiranem Ljubljanskem barju.

V ta namen smo pripravili gensko knjižnico bakterijskih genov za 16S rRNA. S komercialnim kompletom smo izolirali DNA iz vzorca tal in z verižno reakcijo s polimerazo (PCR) pomnožili želene gene. Pomnožene fragmente smo vstavili v plazmidne vektorje ter z njimi transformirali kompetentne celice *E. coli*. Transformante smo izbrali z ampicilinsko in modro-belo selekcijo, velikost vstavljenih fragmentov pa smo preverjali z verižno reakcijo s polimerazo (colony PCR). Pridobili smo nukleotidna zaporedja 192 klonov. Sekvence smo temeljito pregledali in odstranili 15 zaporedij, ki so bila domnevno himerna in kot taka artefakt PCR. Po odstranitvi nekaj nezanesljivih sekvenc je naša knjižnica vsebovala 154 delnih zaporedij, dolgih med 760 in 800 bp, ki predstavljajo zadnjo polovico bakterijskega gena za 16S rRNA.

Filogenetsko analizo smo naredili s programom ARB. Sekvence smo poravnali ob veliko bazo (> 50000) dobro urejenih sekvenc in izračunali filogenetsko drevo po metodi varčnosti (ang. *maximum parsimony*). Sekvence smo uvrstili v devet skupin: 62 sekvenc pripada proteobakterijam, od tega 25 *Alphaproteobacteria* (16,2 %), 17 *Betaproteobacterie* (11%), 3 *Gammaproteobacteria* (1,9 %), 17 *Deltaproteobacteria* (11%), 64 sekvenc pripada skupini *Acidobacteria* (41,6%), 11 *Actinobacteria* (7,1%), po 8 *Planctomycetes* (5,2%) in *Verrucomicrobia* (5,2%) ter ena *Spirochaetes* (0,7%).

V klonski knjižnici je nepričakovano veliko sekvenc, ki spadajo v skupino *Acidobacteria*. Podrobnejša filogenetska analiza je sekvence uvrstila v skupine 1, 2, 3 (tem skupinam, ki so tudi v bazah podatkov najbolj zastopane, pripada 92% klonov te knjižnice), ter 5, 7 in 13. Delež te skupine je precej večji kot v knjižnici, ki so jo iz tal nizkega barja na Ljubljanskem barju pripravili Kraigher in sod. (2006). Poleg tega si sekvence acidobakterij obeh knjižnic niso podobne. V knjižnici tudi ni niti ene sekvence debla *Firmicutes* (rodova *Bacillus*, *Clostridium* in sorodstvo); vendar je bilo tudi v knjižnici iz nizkega barja teh sekvenc le malo. Menimo, da je morda kriva slaba izolacija DNA iz odpornih stadijev teh bakterij, na primer iz spor.

Metoda kloniranja in sekveniranja je zelo informativna in nam omogoča identificirati bakterije v kompleksnem talnem vzorcu, kar je denimo s klasičnimi gojitvenimi metodami povsem nemogoče. Ima pa tudi svoje slabosti in poleg težavne izolacije je predvsem pomnoževanje DNA tisti korak, ki ustvari največ pristranskosti. Ker večina artefaktov PCR nastane v zadnjih ciklih reakcije, bi bilo v prihodnje smotrno nekoliko zmanjšati število ciklov PCR. Poleg tega bi sekveniranje celotnega gena za 16S rRNA (namesto zgolj zadnje polovice) dalo več podatkov za zanesljivejšo filogenetsko določitev.

Ocenjena raznolikost bakterij v tleh je izredno visoka, zato pri tako majhni klonski knjižnici ne moremo trditi, da smo zajeli vse skupine, kar ugotavljajo tudi številni drugi raziskovalci. Enako tudi ne moremo trditi, da razmerja v klonski knjižnici odražajo razmerja celic v vzorcu tal. Lahko pa zaključimo, da so barja tudi glede bakterijske diverzitete vroča točka ter da je že zgolj prisotnost pred kratkim odkritih in slabo poznanih skupin bakterij, denimo *Acidobacteria* in *Verrucomicrobia*, dober razlog za nadaljevanje raziskovanja tega ekosistema.

7 VIRI

Acinas S.G., Sarma-Rupavtarm R., Klepac-Ceraj V., Polz M.F. 2005 PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Applied and Environmental Microbiology*, 71, 12:8966-8969.

Amann R.I., Ludwig W., Schleifer K.H. 1995. Phylogenetic Identification and in Situ Detection of Individual microbial Cells without Cultivation. *Microbial Reviews*, 59, 1:143-169

Ashelford K.E., Chuzhanova N.A., Fry J.C., Jones A.J., Weightman A.J. 2005. At Least 1 in 20 16S rRNA Sequence Records Currently Held in Public Repositories Is Estimated To Contain Substantial Anomalies. *Applied and Environmental Microbiology*, 71, 21:7724-7736

Barns S.M., Cain E.C., Sommerville L., Kuske C.R. 2007. *Acidobacteria* Phylum Sequences in Uranium-Contaminated Subsurface Sediments Greatly Expand the Known Diversity within the Phylum. *Applied and Environmental Microbiology*, 73,9: 3113-3116

Barns S.M., Takala S.L., Kuske C.R. 1999. Wide Distribution and Diversity of Members of the Bacterial Kingdom *Acidobacterium* in the Environment. *Applied and Environmental Microbiology*, 65,4:1731-1737

BIOTOOLS DNA Polymerase – recombinant, from *Thermus thermophilus* (Biotools, Madrid, Španija)
<http://www.biotools.net/eng/productos/pdf/a1.pdf> (zadnjič dostopano: 11.3.2008)

Borneman J. Skroch P.W., O'Sullivan K.M., Palus J.A., Rumjanek N.G., Jansen J.L., Nienhuis J., Triplett E.W. 1996. Molecular Microbial Diversity of an Agricultural Soil in Wisconsin. *Applied and Environmental Microbiology*, 62,6:1935-1943

Borneman J., Triplett E.W. 1997. Molecular Microbial Diversity in Soils from Eastern Amazonia: Evidence for Unusual Microorganisms and Microbial Population Shifts Associated with Deforestation. *Applied and Environmental Microbiology*, 63,7:2647-2653

Chimera detection against small subunit rRNAs (Ribosomal Database Project Release 8.1). 2004.
http://rdp8.cme.msu.edu/docs/chimera_doc.html (zadnjič dostopano: 11.3.2008)

Clayton R.A., Sutton G., Hinkle J.R., Bult C., Fields C. 1995. Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *International Journal of Systematic Bacteriology*, 45, 3:595-599.

Cole J.R., Chai B., Farris R.J., Wang Q., Kulam-Syed-Mohideen A.S., McGarrell D.M., Bandela A.M., Cardenas E., Garrity G.M., Tiedje J.M. 2007. The ribosomal database project (RDP-II): introducing *myRDP* space and quality controlled public data. *Nucleic Acids Research*, 35 (Database issue): D169-D172

Dedysh S.N., Pankratov T.A., Belova S.E., Kulichevskaya I.S., Liesack W. 2006. Phylogenetic Analysis and In Situ Identification of *Bacteria* Community Composition in an Acidic *Sphagnum* Peat Bog. *Applied and Environmental Microbiology*, 72,3:2110-2117

Dunbar J., Barns S. M., Ticknor L.O., Kuske C.R. 2002. Empirical and Theoretical Bacterial Diversity in Four Arizona Soils. *Applied and Environmental Microbiology*, 68,6:3035-3045

Farrelly V., Rainey F.A., Stackebrandt E. 1995. Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Applied and Environmental Microbiology*, 61,7:2798-2801

Felske A., Wolterink A., van Lis R., Akkermans A. D. L. 1998. Phylogeny of the Main Bacterial 16S rRNA Sequences in Drentse A Grassland Soils (The Netherlands). *Applied and Environmental Microbiology*, 64,3:871-879

Felske A., Wolterink A., van Lis R., Akkermans A.D.L. 1998. Phylogeny of Main Bacterial 16S rRNA Sequences in Drentse A Grassland Soils (The Netherlands). *Applied and Environmental Microbiology*, 64,3:871-879

Gevers D., Cohan F.M., Lawrence J.G., Spratt B.G., Coenye T., Feil E.J., Stackebrandt E., van de Peer Y., Vandamme P., Thompson F.L., Swings J. 2005. Opinion: Re-evaluating procaryotic species, *Nature Reviews Microbiology* 3,9:733-739

Graff A., Conrad R. 2005. Impact of flooding on soil bacterial communities associated with poplar (*Populus sp.*) trees. *FEMS Microbiology Ecology*, 53,3:401-415

Handelsman J. 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews*, 68,4:669-685

Hugenholtz P., Goebel B.M., and Pace N.R. 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. *Journal of Bacteriology*, 180, 18: 4765–4774

Hugenholtz P., Goebel B.M., Pace N.R. 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. *Journal of Bacteriology*, 180,18:4765-4774

Hugenholtz P., Huber T. 2003 Chimeric 16S sequences of diverse origin are accumulating in the public databases. *International Journal of Systematic and Evolutionary Microbiology*, 53:289-293

Janssen P.H. 2006. Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes. *Applied and Environmental Microbiology*, 72,3:1719-1728

Janssen P.H., Yates P.S., Grinton B.E., Taylor P.M., Sait M. 2002. Improved Culturability of Soil Bacteria and Isolation in Pure Culture of Novel Members of the Divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria* and *Verrucomicrobia*. *Applied and Environmental Microbiology*, 68,5:2391-2396

Kanagawa T. 2003. Bias and Artefacts in Multitemplate Polymerase Chain Reactions (PCR). *Journal of Bioscience and Bioengineering*, 96,4:317-323

Kraigher B., Stres B., Hacin J., Ausec L., Mahne I., van Elsas J.D., Mandic-Mulec I. 2006. Microbial activity and community structure in two fen soils in the Ljubljana Marsh. *Soil Biology and Biochemistry*, 38,9:2762-2771

Kurata S., Kanagawa T., Magariyama Y., Takatsu K., Yamada K., Yokomaku T., Kamagata Y. 2004. Reevaluation and Reduction of a PCR Bias Caused by Reannealing of Templates. *Applied and Environmental Microbiology* 70, 12:7545-7549

Liles M.R., Manske B.F., Bintrim S.B., Handelsman J., Goodman R.M. 2003. A Census of rRNA Genes and Linked Genomic Sequences within a Soil Metagenomic Library. *Applied and Environmental Microbiology*, 69,5:2684-2691

Ludwig W., Strunk O., Westram R., Richter L., Meier H., Yadhukumar, Buchner A., Ali T., Steppi S., Jobb G., Förster W., Brettske I., Gerber S., Ginhart A.W., Gross O., Grumann S., Hermann S., Jost R., König A., Liss T., Lüßmann R., May M., Nonhoff B., Reichel B., Strehlow R., Stamatakis A., Stuckmann N., Vilbig A., Lenke M., Ludwig T., Bode A., Schleifer K.H. 2004. ARB: a software environment for sequence data. *Nucleic Acids Research*, 32, 4:1363-1371

Maidak B.L., Cole J.R., Parker C.T.Jr., Garrity G.M., Larsen N., Li B., Lilburn T.G., McCaughey M.J., Olsen G.J., Overbeek R., Pramanik S., Schmidt T.M., Tiedje J.M., Woese C.R. 2000. The RDP (Ribosomal Database Project) continues. *Nucleic Acids Research*, 28, 1:173-174

Margulies M. in sod. 2005 Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057):376-380

Martinčič A. 2003. Barje na Ljubljanskem barju – nekdanj, včeraj, danes, jutri. *Proteus*, 65,6:246-256

McCaig A.E., Glover A., Prosser J.I. 1999. Molecular Analysis of Bacterial Community Structure and Diversity in Unimproved and Improved Upland Grass Pastures. *Applied and Environmental Microbiology*, 65,4:1721-1730

Morales S.E., Mouser P.J., Ward N., Hudman S., Gotelli N.J., Ross D.S., Lewis T.A.

2006. Comparison of Bacterial Communities in New England *Sphagnum* Bog Using Terminal Restriction Length Polymorphism (T-RFLP). *Microbial Ecology*, 52,1:34-44

Patel R., Lin C., Laney M., Kurn N., Rose S., Ullman E.F. 1996. Formation of chimeric DNA primer extension products by template switching onto an annealed downstream oligonucleotide. *Proceedings of National Academy of Science*, 93,7:2969-2974.

Polz M. F., Cavanaugh C. M. 1998. Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, 64,10:3724-3730.

Quaiser A., Ochsenreiter T., Lanz C., Schuster S. C., Treusch A. H., Eck J., Schleper C. 2003.

Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. *Molecular Microbiology* 50,2:563-575

Rheims H., Sproer C., Rainey F.A., Stackebrandt E. 1996. Molecular biological evidence for the occurrence of uncultured members of the actinomycete line of descent in different environments and geographical locations. *Microbiology*, 142: 2863-2870

Sait M., Davis K.E.R., Janssen P.H. 2006. Effect of pH on Isolation and Distribution of Members of Subdivision I of the Phylum *Acidobacteria* Occurring in Soil. *Applied and Environmental Microbiology*, 72,3:1852-1857

Schloss P. D., Handelsman J. 2004. Status of the Microbial Census. *Microbiology and Molecular Biology Reviews*, 68,4: 686 – 691

Schloss P.D., Handelsman J. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology*, 71, 3:1501-1506

Smit E., Leeflang P., Gommans S., Broek J., van Mil S., Wernars K. 2001. Diversity and Seasonal Fluctuations of the Dominant Members of the Bacterial Soil Community in a Wheat Field as Determined by Cultivation and Molecular Methods. *Applied and Environmental Microbiology*, 67, 5:2284-2291

Suzuki M.T., Giovannoni S.J. 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology*, 62,2:625-630.

Tamaki H., Sekiguchi Y., Hanada S., Nakamura K., Nomura N., Matsumura M., Kamagata Y. 2005. Comparative Analysis of Bacterial Diversity in Freshwater Sediment of a Shallow Eutrophic Lake by Molecular and Improved Cultivation-Based Techniques. *Applied and Environmental Microbiology*, 71,4:2162-2169

Weisburg W.G., Barns S.M., Pelletier D.A., Lane D.J. 1991. 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, 173: 697-703

Wintzingerode F., Goebel U.B., Stackebrandt E. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews*, 21,3:213-229.

Wise M.G., McArthur J. V., Shimkets L.J. 1997. Bacterial Diversity of a Carolina Bay as Determined by 16S rRNA Gene Analysis: Confirmation of Novel Taxa. *Applied and Environmental Microbiology*, 63,4:1505-1514

ZAHVALA

Zahvaljujem se mentorici prof. dr. Ines Mandić-Mulec za pomoč pri pripravi diplome ter vso spodbudo pred in po njej. Hvala dr. Barbari Kraigher, da mi je budno gledala čez ramo v laboratoriju. Hvala Jaku, da me je potegnil iz luknje. In hvala Maji in moji družini, da so me prenašali, ko sem bil horizontalno konstruktiven.